# MOLECULAR EVOLUTION, SYSTEMATICS AND ECOLOGY OF MICROSPORIDIA FROM FISHES AND CRUSTACEANS 

by

## AMANDA MAY VIVIAN BROWN

B.Sc., University of British Columbia, 1996
A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in
THE FACULTY OF GRADUATE STUDIES
(Zoology)
THE UNIVERSITY OF BRITISH COLUMBIA
JUNE 2005
(C) Amanda May Vivian Brown, 2005


#### Abstract

Microsporidia are unicellular, intracellular parasites once considered primitive protists without mitochondria, but recently shown by DNA analysis to be related to Fungi. These pathogens have undergone evolution characterized by dramatic losses of genome complexity during adaptation and speciation within host cells. Microsporidian infections can cause mortality in fishes and crustaceans and significant commercial losses in wild, hatchery and fish farm stocks. This study examined microsporidia in salmon, trout, Atlantic cod, Pacific cod, lingcod, walleye pollock, haddock, sablefish, Pacific tomcod, smooth pink shrimp, and a copepod, from marine and freshwater habitats on the east and west coasts of Canada and elsewhere. The purpose was to examine problems in morphological and molecular classification, identify relationships in light of taxonomic and host-parasite co-evolution hypotheses, and combine these data towards description of new species. Ribosomal RNA (rDNA), elongation factor-1 alpha and RNA polymerase II genes were amplified. Phylogenies were reconstructed and tested statistically. Three bioinformatics software (Java) programs were developed to aid DNA analysis. Light and transmission electron microscopy (TEM) were used to characterize spores, developmental stages, and pathogenesis. Among 26 isolates of Loma salmonae, most genetic diversity occurred in farms in California, Colorado and Chile (up to 0.66 \% rDNA difference), while wild isolates from British Columbia were invariant. Ten other Loma species ( 1 to 8 isolates each) grouped in phylogenetic clusters often corresponding to species or host. Five new Loma species were described based on developmental timing, spore and xenoma features. Intraspecific polymorphism caused low genetic resolution of some species. Loma species displayed high evolutionary and developmental phenotypic plasticity relative to genetic variance. Data showed host-parasite co-evolution in family Gadidae, and separated previously synonymized taxa (L. branchialis and L. morhua, L. salmonae and Loma sp. in brook trout). Thelohania butleri from shrimp was closely related to microsporidians in marine decapods and grouped apart from other Thelohania species, suggesting this genus needs revision. A new genus and species was described from a copepod, based on production of a mucocalyx and 3 distinct spore forms - features suggesting a possible secondary host. Overall, host and ecology better predicted relationships than did morphology, emphasizing the importance of plasticity.


## Table of Contents

Abstract ..... ii
Table of Contents ..... iii
List of Tables ..... x
List of Figures ..... xi
Contributions by others ..... xv
Dedication ..... xvi
General Acknowledgements ..... xvii
Chapter 1: General Introduction ..... 1
Outline and structure of the thesis ..... 1
General features: unity and diversity of the microsporidia ..... 1
Historical problems in biology and systematics of the microsporidia ..... 5
Ubiquitous pests, pest controls, and curiosities ..... 5
Microsporidia as Fungi: the problem of parasitic reduction ..... 6
Towards a testable theory of parasitic reduction in the microsporidia ..... 6
Higher classification in flux: the problem of rapidly evolving morphology ..... 9
Systematics of closely related species ..... 10
Introduction to Loma spp. and microsporidia from fishes ..... 12
Introduction to microsporidia from crustaceans ..... 15
Introduction to tools of study ..... 17
Genes chosen ..... 17
Phylogenetic analysis methods and Modeltest (Posada \& Crandall, 1998) ..... 19
Monophyly constraints and hypothesis testing with AU and other tests (Shimodaira \& Hasegawa, 2001) ..... 20
Software design for analysis of gaps (indels), double signal, and rDNA stems ..... 21
Light and transmission electron microscopy (TEM) ..... 22
Objectives ..... 23
Literature cited ..... 24
Figures and Tables ..... 36
Chapter 2: DNA distinguishes populations of Loma salmonae (Microsporidia) in salmon and trout (Oncorhynchus spp.) and identifies a cryptic species in brook trout (Salvelinus fontinalis) ..... 40
Introduction ..... 40
Materials and Methods ..... 44
Specimen collection ..... 44
Spore isolation ..... 45
DNA isolation ..... 45
Polymerase chain reaction test (PCR-test) ..... 45
PCR for sequencing and cloning ..... 46
Cloning ..... 47
DNA sequencing ..... 47
Sequence analysis ..... 47
Results ..... 49
PCR-test for Loma salmonae in wild salmon ..... 49
Intraspecific variation in the rDNA genes ..... 49
Intraspecific variation in the EF-1 $\alpha$ gene ..... 51
Variation between $L$. salmonae isolates and " $L$. salmonae SV" from brook trout ..... 51
Discussion ..... 52
Intraspecific rDNA sequence variation in $L$. salmonae ..... 52
Intraspecific differences in the EF-1 $\alpha$ gene ..... 55
Prevalence of $L$. salmonae in five wild salmon species ..... 55
The question of strain "SV" ..... 57
Conclusions ..... 60
Further Investigation ..... 61
Acknowledgements ..... 61
Literature Cited ..... 62
Figures and Tables ..... 66
Chapter 3: Description of five new species of Loma (Microsporidia) parasitizing Pacific fishes. ..... 74
Introduction ..... 74
Materials and Methods ..... 78
Specimen collection ..... 78
Tissue fixation and preparation ..... 78
Prevalence and intensity ..... 79
Xenoma size ..... 79
Measurements using transmission electron microscopy (TEM) ..... 80
Spore measurement and conversion factors ..... 81
Polymerase chain reaction (PCR) and DNA sequencing ..... 82
Results ..... 84
Definitions of developmental stages ..... 84
Descriptions ..... 85
Loma pacificodae n. sp. ..... 85
Taxonomic summary ..... 86
Ecology ..... 87
Pathology ..... 87
Remarks ..... 87
Comparison of $L$. pacificodae n. sp. to other species examined in this study ..... 88
Comparison of L. pacificodae n . sp. to previously described Loma species ..... 90
Loma wallae n . sp . ..... 91
Taxonomic summary ..... 92
Ecology ..... 93
Pathology ..... 93
Remarks ..... 93
Comparison of $L$. wallae n. sp. to other species examined in this study ..... 93
Comparison of $L$. wallae n . sp . to previously described Loma species ..... 95
Loma kenti n. sp. ..... 95
Taxonomic summary ..... 96
Ecology ..... 96
Pathology ..... 96
Remarks ..... 96
Comparison of L. kenti n. sp. to other species examined in this study ..... 97
Comparison of $L$. kenti n . sp. to previously described Loma species ..... 98
Loma lingcodae n . sp ..... 98
Taxonomic summary ..... 99
Ecology ..... 99
Pathology ..... 99
Remarks ..... 99
Comparison of $L$. lingcodae n . sp. to other species examined in this study ..... 100
Comparison of $L$. lingcodae n . sp. to previously described Loma species ..... 101
Loma richardi n . sp . ..... 101
Taxonomic summary ..... 102
Ecology ..... 102
Pathology ..... 103
Remarks ..... 103
Comparison of $L$. richardi n . sp . to other species examined in this study ..... 103
Comparison of $L$. richardi n . sp. to previously described Loma species ..... 104
Inferred developmental sequences of new Loma species ..... 104
Development in L. pacificodae n. sp. ..... 105
Development in $L$. wallae n . sp . ..... 105
Development in L. kenti n. sp. ..... 106
Development in L. lingcodae n. sp. ..... 106
Development in $L$. richardi n. sp. ..... 107
Discussion ..... 108
Comparative ecology and pathology of new Loma species ..... 108
Prevalence ..... 108
Pathology ..... 109
New features ..... 110
Considerations regarding character interpretation ..... 110
Relative character variability in new Loma species ..... 110
Characters that may be correlated ..... 110
Spore shrinkage factors ..... 111
Characters potentially determined by the host ..... 112
Questions answered by transmission studies ..... 112
Ultrastructure vs. DNA data and species concepts in Loma ..... 113
Further Investigation ..... 117
Acknowledgements ..... 117
Literature Cited ..... 118
Figures and Tables ..... 122
Chapter 4: Phylogenetic species- and genus-boundary tests in Loma (Microsporidia) using three genetic loci: rDNA, EF-1 $\alpha$, and RPB1. ..... 151
Introduction ..... 151
Materials and Methods ..... 159
Specimen collection ..... 159
Spore size measurements for Loma spp. in Atlantic cod and haddock ..... 160
DNA isolation ..... 161
Polymerase chain reaction (PCR) ..... 162
Cloning ..... 162
DNA sequencing ..... 162
Alignment ..... 163
Flip algorithm ..... 163
Polymorphic sites ..... 163
Phylogenetic analysis ..... 164
Monophyly constraints and AU tests ..... 165
Tests for recombination ..... 166
Results ..... 167
Spore size comparisons for Loma spp. in Atlantic cod and haddock ..... 167
Ribosomal DNA (SSU, ITS and LSU) sequence characteristics ..... 167
Elongation factor 1-alpha (EF-1 $\alpha$ ) sequence characteristics ..... 168
RNA polymerase largest subunit (RPB1) sequence characteristics ..... 169
Phylogenetic relationships among Loma species for rDNA Region 1 ( $5^{\prime}$ half of SSU) ..... 169
Phylogenetic relationships among Loma species for rDNA Region 2 (3' half of SSU) ..... 170
Phylogenetic relationships among Loma species for rDNA Region 3 (3' end of SSU, ITS and 5' end of LSU) ..... 170
Phylogenetic relationships among Loma species for partial EF-1 $\alpha$ ..... 172
Monophyly tests (AU-tests) for Loma species for rDNA Region 1 (5' half of SSU) ..... 172
Monophyly tests (AU-tests) for Loma species for rDNA Region 2 (3' half of SSU) ..... 173
Monophyly tests (AU-tests) for Loma species for rDNA Region 3 (3' end of SSU, ITS and 5' end of LSU) ..... 173
Monophyly tests (AU-tests) for Loma species for partial EF-1 $\alpha$ ..... 174
Recombination tests ..... 175
Phylogenetic relationships among genera for SSU rDNA ..... 175
Phylogenetic relationships among genera for partial RPB1 ..... 176
Monophyly tests (AU-tests) for genera for SSU rDNA ..... 176
Monophyly tests (AU-tests) for genera for partial RPB1 ..... 177
Discussion ..... 178
Overview of major results ..... 178
Identity of spores with L. branchialis and L. morhua (Morrison \& Sprague, 1981) ..... 179
Phylogenetic support for separate species L. branchialis and L. morhua (Morrison \& Sprague, 1981) ..... 180
Phylogenetic placement and support for seven undescribed Loma spp. ..... 181
Reticulate speciation in the " G " clade? ..... 186
Possibility 1: Genes underwent reticulate evolution. ..... 186
Possibility 2: Organisms underwent reticulate evolution. ..... 187
Improving resolution in clade " G " ..... 189
Integrity of genus Loma ..... 189
Co-evolution and phylogeography of Loma spp. in gadids ..... 191
Conclusions ..... 193
Further Investigation ..... 194
Acknowledgements ..... 194
Literature Cited ..... 195
Figures and Tables ..... 200
Chapter 5: Phylogenetic distance of Thelohania butleri Johnston, Vernick \& Sprague, 1978 (Microsporidia; Thelohaniidae) from the smooth pink shrimp Pandalus jordani Rathbun, 1902 from other Thelohania species: implications for genus Thelohania Henneguy, 1892 and family Thelohaniidae Hazard \& Oldacre 1975. ..... 257
Introduction ..... 257
Materials and Methods ..... 261
Specimen collection ..... 261
Light and transmission electron microscopy (TEM) ..... 261
DNA isolation ..... 262
Polymerase chain reaction (PCR) for cloning and sequencing ..... 262
Cloning ..... 263
DNA sequencing ..... 263
Flip algorithm ..... 263
Polymorphic sites ..... 263
Alignment ..... 264
Phylogenetic analysis ..... 264
Monophyly constraints and approximately unbiased (AU) tests ..... 265
Results ..... 266
Morphological features ..... 266
Ribosomal DNA sequence characteristics ..... 267
Phylogenetic relationships ..... 267
Monophyly constraints and AU-tests ..... 268
Discussion ..... 269
Morphological identity with T. butleri Johnston, Vernick \& Sprague, 1978 ..... 269
Phylogenetic placement of T. butleri ..... 269
Implications for genus Thelohania Henneguy, 1892 ..... 272
Implications for family Thelohaniidae Hazard \& Oldacre, 1975 ..... 273
Further Investigation ..... 274
Acknowledgements ..... 274
Literature Cited ..... 276
Figures and Tables ..... 279
Chapter 6: A new microsporidian, Vossbrinckus richardi n. gen., n. sp. (Microsporidia; Golbergiidae) from a copepod: ultrastructural features and phylogenetic placement. ..... 294
Introduction ..... 294
Materials and Methods ..... 297
Specimen collection ..... 297
Light and transmission electron microscopy (TEM) ..... 297
DNA isolation ..... 297
Polymerase chain reaction (PCR) ..... 298
Cloning ..... 298
DNA sequencing ..... 299
Flip algorithm ..... 299
Alignment ..... 299
Phylogenetic analysis ..... 299
Monophyly constraints and AU tests ..... 300
Results ..... 301
Description ..... 301
Genus Vossbrinckus n. gen. ..... 301
Diagnosis ..... 301
Vossbrinckus richardi n. gen., n. sp. ..... 301
Taxonomic summary ..... 302
Ecology and pathology ..... 303
Ribosomal DNA sequence characteristics ..... 303
Phylogenetic analysis of rDNA ..... 304
Monophyly constraints and AU-tests of rDNA region ..... 304
Remarks ..... 305
Morphological justification for $V$. richardi n. gen., n. sp ..... 306
Molecular and phylogenetic justification for $V$. richardi n . gen., n. sp. ..... 309
Justification for placement in family Golbergiidae Issi, 1986 ..... 310
Life-cycle evolution in the "Aquatic Outgroup" ..... 311
Usefulness and design of a molecular-ecological PCR-probe ..... 312
Conclusion ..... 313
Further Investigation ..... 314
Acknowledgements ..... 314
Literature Cited ..... 316
Figures and Tables ..... 320
Chapter 7: General Conclusions ..... 332
Overview ..... 332
Introduction ..... 332
Taxonomic Conclusions ..... 335
Defining and evaluating species ..... 335
New species ..... 338
Field-diagnostic characters ..... 340
Plastic of convergent characters ..... 341
Biological Conclusions ..... 343
Co-evolution and biogeographic patterns ..... 343
Epidemiology and pathology ..... 345
DNA Peculiarities ..... 349
Indels and paralogs ..... 349
Ribosomal DNA evolutionary rate ..... 351
New Questions ..... 352
Summary ..... 354
Literature Cited ..... 357
Figures and Tables ..... 363
Appendices ..... 365
Appendix 1: Explanation of three software programs designed in this study. ..... 366
Introduction ..... 366
Software language ..... 367
Design of Flip Analyzer program ..... 367
Demonstration of Flip Analyzer ..... 370
Design of Gap Matrix program ..... 371
Demonstration of Gap Matrix ..... 372
Design of Stem State program ..... 372
Demonstration of Stem State ..... 376
Help files ..... 376
Discussion of program applications ..... 376
Further applications of Flip Analyzer ..... 377
Further applications of Gap Matrix ..... 377
Further applications of Stem State ..... 379
Acknowledgements ..... 380
Literature Cited ..... 380
Figures and Tables ..... 382
Appendix 2: Example of Modeltest instructions ..... 400
Appendix 3: Example of Modeltest output ..... 401
Appendix 4: Sample PAUP* input commands ..... 404
Appendix 5: Explanation of monophyly tests and formulation of species-boundaries hypotheses ..... 406
Appendix 6: Justification for use of the "AU test" of Shimodaira \& Hasegawa (2001) ..... 407
Appendix 7: Instructions used for creation of monophyly-forced clades ..... 410
Appendix 8: Example of likelihood scores used in CONSEL ..... 413
Appendix 9: Example of CONSEL output file ..... 414
Appendix 10: Dot-blot results. ..... 415
Appendix 11: Features of previously described and new Loma species ..... 416
Appendix 12: Name equivalents across chapters ..... 417
Appendix 13: Key to reading isolate names ..... 418
Appendix 14: Pairwise distance matrix for rDNA reference sequences from Loma species. ..... 419
Appendix 15: Partial elongation factor 1-alpha DNA reference sequences. ..... 420
Appendix 16: Partial elongation factor 1 -alpha amino acid reference sequences ..... 422
Appendix 17: Partial RNA polymerase largest subunit II sequences. ..... 423
Appendix 18: Phylograms of Loma spp. for rDNA Region 3 ..... 424
Appendix 19: Consensus parsimony tree with inclusion of polymorphic characters ..... 427
Appendix 20: Technique improvements ..... 428

## List of Tables

Table 1.1: Loma species, hosts, and localities. ..... 39
Table 2.1: Number of fish examined and prevalence of $L$. salmonae ..... 67
Table 2.2: List of isolates and rDNA and EF-1 $\alpha$ sequence differences. ..... 69
Table 2.3: Percent differences in rDNA sequences among isolates ..... 71
Table 3.1: Fishes examined for Loma infection ..... 122
Table 3.2: Prevalence estimates for Loma species. ..... 123
Table 3.3: Xenoma features of new Loma species ..... 125
Table 3.4: Xenoma wall features of new Loma species. ..... 126
Table 4.1: Isolates of various Loma species for which DNA was sequenced ..... 200
Table 4.2: Primers used in PCR and sequencing. ..... 202
Table 4.3: Spore sizes and converted spore sizes for $L$. branchialis and L. morhua. ..... 203
Table 4.4: Length of DNA sequenced for 5 gene regions for Loma species. ..... 204
Table 4.5: Intraspecific sequence differences within Loma species. ..... 205
Table 4.6: Insertions or deletions (indels) in Loma rDNA. ..... 206
Table 4.7: Polymorphic sites in Loma rDNA. ..... 207
Table 4.8: Statistical test results from CONSEL for rDNA Region 1 ..... 219
Table 4.9: Statistical test results from CONSEL for rDNA Region 2. ..... 227
Table 4.10: Statistical test results from CONSEL for rDNA Region 3 ..... 241
Table 4.11: Statistical test results from CONSEL for the partial EF-1 $\alpha$ gene. ..... 243
Table 4.12: Summary of monophyly test results for all gene regions and species of Loma. ..... 244
Table 4.13: Sequences obtained from Genbank for all gene regions. ..... 245
Table 4.14: Test results from CONSEL for SSU rDNA and RPB1 for Loma and other genera. ..... 255
Table 5.1: Ribosomal DNA sequences from Genbank. ..... 279
Table 5.2: Bootstrap results for various input data sets. ..... 284
Table 5.3: Statistical test results for alternate trees using CONSEL ..... 290
Table 6.1: Ribosomal DNA sequences from Genbank. ..... 320
Table 6.2: Statistical test results for alternate tree topologies using CONSEL ..... 330

## List of Figures

Figure 1.1: Microsporidian spore features ..... 36
Figure 1.2: Microsporidian life cycles and development. ..... 37
Figure 2.1: Polymerase reaction test for L. salmonae in wild salmon. ..... 66
Figure 2.2: Inhibitors in PCR. ..... 68
Figure 2.3: Ribosomal DNA sequence showing sequence variants. ..... 72
Figure 2.4: EF-1 $\alpha$ DNA and amino acid sequence showing variants among isolates. ..... 73
Figure 3.1: Prevalence differences of $L$. pacificodae n . sp. ..... 124
Figure 3.2: Mean xenoma sizes. ..... 127
Figure 3.3: Mean vesicle sizes. ..... 128
Figure 3.4: Mean number of vesicles. ..... 129
Figure 3.5: Mean number of tubules. ..... 130
Figure 3.6: Mean spore sizes (resin). ..... 131
Figure 3.7: Mean spore sizes (converted). ..... 132
Figure 3.8: Mean number of polar filament turns. ..... 133
Figure 3.9: L. pacificodae n . sp. xenomas in secondary lamellae. ..... 134
Figure 3.10: L. pacificodae n. sp. xenomas in central venous sinus ..... 134
Figure 3.11: L. pacificodae n. sp. xenoma in gonads. ..... 134
Figure 3.12: L. pacificodae n. sp. xenoma periphery. ..... 134
Figure 3.13: L. pacificodae n. sp. xenoma wall. ..... 135
Figure 3.14: L. pacificodae n. sp. collagenous inclusions. ..... 135
Figure 3.15: L. pacificodae n. sp. binary fission within RER. ..... 135
Figure 3.16: L. pacificodae n. sp. highly vacuolated merogonial plasmodia. ..... 135
Figure 3.17: L. pacificodae n. sp. oblong sporogonial plasmodium ..... 136
Figure 3.18: L. pacificodae n. sp. parasitophorous vacuole formation. ..... 136
Figure 3.19: L. pacificodae n. sp. parasitophorous vacuole formation ..... 136
Figure 3.20: L. pacificodae n. sp. small light vesicles. ..... 136
Figure 3.21: L. pacificodae n. sp. tubule-filled vesicles. ..... 137
Figure 3.22: L. pacificodae n. sp. tubules in parasitophorous vacuoles. ..... 137
Figure 3.23: L. pacificodae n. sp. 2 spores per parasitophorous vacuole. ..... 137
Figure 3.24: L. pacificodae n . sp. host response (light micrograph) ..... 137
Figure 3.25: L. pacificodae n. sp. host response (transmission electron micrograph). ..... 138
Figure 3.26: L. pacificodae n. sp. host response in gonads. ..... 138
Figure 3.27: L. pacificodae n. sp. mitochondrion-like structures. ..... 138
Figure 3.28: L. wallae n . sp. xenoma in central venous sinus. ..... 138
Figure 3.29: L. wallae n. sp. secondary xenomas ..... 139
Figure 3.30: $L$. wallae n . sp . xenoma wall. ..... 139
Figure 3.31: L. wallae n . sp. binary fission without RER covering. ..... 139
Figure 3.32: L. wallae n. sp. merogonial plasmodium. ..... 139
Figure 3.33: L. wallae n. sp. products of cylindrical plasmodium (sporoblasts). ..... 140
Figure 3.34: L. wallae n. sp. early sporoblasts in chains of 4. ..... 140
Figure 3.35: L. wallae n. sp. dark material-filled vesicles. ..... 140
Figure 3.36: $L$. wallae n. sp. tubule-filled vesicles and 2 spores per PV. ..... 140
Figure 3.37: L. kenti n. sp. large xenoma in secondary lamella. ..... 141
Figure 3.38: L. kenti n. sp. xenoma wall. ..... 141
Figure 3.39: L. kenti n. sp. xenoma wall. ..... 141
Figure 3.40: L. kenti n. sp. multinucleate merogonial plasmodium. ..... 141
Figure 3.41: L. kenti n . sp. meront or merogonial plasmodia ..... 142
Figure 3.42: L. kenti n. sp. early parasitophorous vacuole formation. ..... 142
Figure 3.43: L. kenti n. sp. parasitophorous vacuole and dark material-filled vesicles. ..... 142
Figure 3.44: L. kentin. sp. binary fission within parasitophorous vacuole. ..... 142
Figure 3.45: L. kenti n. sp. binary fission of merogonial plasmodium. ..... 143
Figure 3.46: L. kenti n. sp. sporogonial plasmodium division "clover-leaf", ..... 143
Figure 3.47: L. kenti n . sp. products of sporogonial plasmodia division. ..... 143
Figure 3.48: L. kenti n. sp. tubule-like structures within dark material-filled vesicles. ..... 143
Figure 3.49: L. kenti n . sp. one spore per parasitophorous vacuole. ..... 144
Figure 3.50: L. lingcodae n. sp. xenomas in secondary lamellae. ..... 144
Figure 3.51: L. lingcodae n. sp. xenoma wall. ..... 144
Figure 3.52: L. lingcodae n. sp. xenoma wall details. ..... 144
Figure 3.53: L. lingcodae n. sp. merogonial plasmodium. ..... 145
Figure 3.54: L. lingcodae n. sp. meront or merogonial plasmodium and long tubule ..... 145
Figure 3.55: L. lingcodae n. sp. spindle plaque. ..... 145
Figure 3.56: L. lingcodae n. sp. dark granular material in vesicles. ..... 145
Figure 3.57: L. lingcodae n. sp. 4 spores per parasitophorous vacuole. ..... 146
Figure 3.58: L. lingcodae n. sp. mitochondrion-like structure. ..... 146
Figure 3.59: L. richardi n . sp. small xenoma in secondary lamella. ..... 146
Figure 3.60: L. richardi n . sp. xenoma packed with spores. ..... 146
Figure 3.61: L. richardi n . sp . xenoma wall. ..... 147
Figure 3.62: L. richardi n . sp. xenoma wall details. ..... 147
Figure 3.63: L. richardi n . sp. double-layered xenoma wall. ..... 147
Figure 3.64: L. richardin. sp. merogonial plasmodium within RER. ..... 147
Figure 3.65: L. richardi n. sp. ER cisternae in sporont. ..... 148
Figure 3.66: L. richardin. sp. binary fission of sporonts. ..... 148
Figure 3.67: L. richardi n. sp. details of xenoma cytoplasm. ..... 148
Figure 3.68: L. richardi n . sp. different stages within a parasitophorous vacuole. ..... 148
Figure 3.69: L. richardi n. sp. details of parasitophorous vacuole formation. ..... 149
Figure 3.70: L. richardin. sp. small light vesicles. ..... 149
Figure 3.71 : L. richardi n . sp. small light vesicles around meront. ..... 149
Figure 3.72: L. richardi n. sp. 2 spores per parasitophorous vacuole. ..... 149
Figure 3.73: L. richardi n. sp. tubule-filled vesicles and parasitophorous vacuoles. ..... 150
Figure 3.74: L. pacificodae n. sp. tubules oriented at poles. ..... 150
Figure 3.75: L. lingcodae n. sp. polyribosome structure. ..... 150
Figure 4.1: Regions of rDNA used in phylogenetic analyses. ..... 208
Figure 4.2: Consensus tree from rDNA Region 1 with bootstrap values. ..... 209
Figure 4.3: Consensus tree from rDNA Region 2 with bootstrap values. ..... 210
Figure 4.4: Consensus tree from rDNA Region 3 with bootstrap values. ..... 211
Figure 4.5: Consensus tree from rDNA Region 3 from gap matrix data alone. ..... 212
Figure 4.6: Consensus tree from rDNA Region 3 from nucleotide and gap matrix data. ..... 213
Figure 4:7: Details from Fig. 3.6 showing paralogous relationships of rDNA copies. ..... 214
Figure 4.8: Trees for partial EF-1 $\alpha$ gene obtained using all three optimality criteria ..... 215
Figure 4.9: Consensus tree from partial EF-1 $\alpha$ gene with bootstrap values ..... 216
Figure 4.10: Alternative trees for monophyly tests for rDNA Region 1 ..... 217
Figure 4.11: Alternative trees for monophyly tests for rDNA Region 2. ..... 218
Figure 4.12: Alternative trees for monophyly tests for rDNA Region 3 ..... 228
Figure 4.13: Alternative trees for monophyly tests for the partial EF-1 $\alpha$ gene. ..... 242

Figure 4.14: Relationships among species and genera from fishes by 3 analysis
methods. ..... 246
Figure 4.15: Consensus tree from SSU rDNA for genera from fishes with bootstrap values. ..... 248
Figure 4.16: Relationships among Loma and other species by 3 methods for partial RPB1 DNA. ..... 249
Figure 4.17: Consensus tree from partial RPB1 with bootstrap values. ..... 250
Figure 4.18: Alternative trees for monophyly tests for SSU rDNA for Loma and other genera. ..... 251
Figure 4.19: Alternative trees for monophyly tests for the partial RPB1 gene. ..... 254
Figure 4.20: Comparison of parasite and host trees ..... 256
Figure 5.1: Light micrograph of spores, 8 per sporophorous vesicle. ..... 280
Figure 5.2: Developmental stages undergoing binary fission. ..... 281
Figure 5.3: Early developmental stages of Thelohania butleri. ..... 282
Figure 5.4: Accumulation of dark material between stages of T. butleri. ..... 282
Figure 5.5: Sporoblasts and sporophorous vesicle inclusions in $T$. butleri. ..... 282
Figure 5.6: Spores and developmental stages of T. butleri. ..... 282
Figure 5.7: Spores within a matrix of fibrous material with round, dark inclusions. ..... 283
Figure 5.8: Spore showing 13 turns of the polar filament. ..... 283
Figure 5.9: Maximum parsimony, distance, and maximum likelihood trees. ..... 284
Figure 5.10: Distance consensus tree with bootstrap values. ..... 287
Figure 5.11: Maximum parsimony consensus tree with bootstrap values. ..... 288
Figure 5.12: Maximum likelihood consensus tree with bootstrap values. ..... 289
Figure 5.13: Alternate tree topologies that were compared in CONSEL. ..... 291
Figure 6.1: Spores distributed in clusters and freely in copepod haemocoel ..... 321
Figure 6.2: Light micrograph showing pyriform and elongate spores. ..... 321
Figure 6.3: Light micrograph showing elongate and ovoid spores. ..... 321
Figure 6.4: Light micrograph showing more pyriform and elongate spores. ..... 321
Figure 6.5: Light micrograph showing pyriform spores embedded in mucous. ..... 321
Figure 6.6: Early developmental stage of Vossbrinckus richardi n. gen., n. sp. ..... 322
Figure 6.7: Thick walled sporont of Vossbrinckus richardi n. gen., n. sp. ..... 322
Figure 6.8: Sporont of Vossbrinckus richardi n. gen., n. sp ..... 322
Figure 6.9: Early sporoblast of Vossbrinckus richardi n. gen., n. sp. ..... 322
Figure 6.10: Spore and early sporoblast. ..... 323
Figure 6.11: Sporoblast with thick, multi-layered coat in a stream of mucous. ..... 323
Figure 6.12: Sporoblast of elongate form with thick, multi-layered coat. ..... 323
Figure 6.13: Sporoblast of pyriform spore with crescentric nucleus. ..... 323
Figure 6.14: Mature pyriform spore. ..... 324
Figure 6.15: Mature elongate spore. ..... 324
Figure 6.16: Details of pyriform spore. ..... 324
Figure 6.17: Details of isofilar polar filament and spore cytoplasm. ..... 324
Figure 6.18: Several spores in streams of mucous. ..... 325
Figure 6.19: Anterior end of elongate spore. ..... 325
Figure 6.20: Ovoid spore in transverse section. ..... 325
Figure 6.21: Details of binucleate ovoid spore. ..... 325
Figure 6.22: Malformed or abortive spore with some double coils. ..... 326
Figure 6.23: Aberrant spore with polar filaments at an unusually steep angle. ..... 326
Figure 6.24: Maximum parsimony tree showing placement of V. richardi n. gen., n. sp.. ..... 327
Figure 6.25: Distance tree showing placement of $V$. richardi $n$. gen., n. sp. ..... 328
Figure 6.26: Maximum likelihood tree showing placement of $V$. richardi n . gen., n. sp ..... 329
Figure 6.27: Alternate trees used in AU and other tests. ..... 331
Figure 7.1: Phylogeny of Loma species showing corresponding morphological character states ..... 363
Figure A1.1: Example of a double signal sequence ..... 382
Figure A1.2: Flip analyzer first menu choices. ..... 383
Figure A1.3: Flip analyzer choosing a file ..... 383
Figure A1.4: Flip analyzer sample input file. ..... 384
Figure A1.5: Flip analyzer "Analyze" menu option ..... 384
Figure A1.6: Flip analyzer menu for shift interval. ..... 385
Figure A1.7: Flip analyzer results for indels of -1 and 1 ..... 385
Figure A1.8: Flip analyzer results for indels of -10 to 10 ..... 386
Figure A1.9: Flip analyzer results showing relative errors. ..... 386
Figure A1.10: Gap matrix first menu choice. ..... 387
Figure A1.11: Gap matrix choosing a file. ..... 387
Figure A1.12: Gap matrix sample input file. ..... 388
Figure A1.13: Gap matrix analysis menu options. ..... 388
Figure A1.14: Gap matrix results of first step of program. ..... 389
Figure A1.15: Gap matrix second menu option. ..... 389
Figure A1.16: Gap matrix results of second step of program ..... 390
Figure A1.17: Gap matrix third menu option. ..... 390
Figure A1.18: Gap matrix final matrix results. ..... 391
Figure A1.19: Gap matrix PAUP formatted output. ..... 391
Figure A1.20: Model of rRNA stem state changes. ..... 392
Figure A1.21: Weight matrix for rRNA stem state changes ..... 393
Figure A1.22: Stem state program sample input file. ..... 394
Figure A1.23: Stem state analysis menu choices. ..... 394
Figure A1.24: Stem state results of first step and further menu choices ..... 395
Figure A1.25: Stem state results of second step. ..... 395
Figure A1.26: Stem state third menu choice. ..... 396
Figure A1.27: Stem state results of third step ..... 396
Figure A1.28: Stem state menu for input of transition and transversion values. ..... 397
Figure A1.29: Stem state menu options for resolution ..... 397
Figure A1.30: Stem state final results. ..... 398
Figure A1.31: Stem state menu option for PAUP formatting. ..... 398
Figure A1.32: REALEM package help files ..... 399
Figure A1.33: REALEM package version and authorship details. ..... 399

## Contributions by others

## Chapter 2

Dr. Ross W. Shaw and Sheila C. Dawe maintained laboratory salmon. Dr. Michael L. Kent and Ross Shaw collected, examined, and contributed infected salmon from laboratory, farms and hatcheries. Sheila C. Dawe collected material for the wild salmon survey. Dr. J. Genaro Sanchez-Martinez performed experimental transmission with brook trout and provided gill material for this study. Dr. Michael Kent provided the initial idea for this work and supplied several ribosomal DNA primers.

## Chapter 3

Sheila C. Dawe helped to collect, prepare and screen histological sections for species of Loma. Dr. Ross W. Shaw collected shiner perch and salmon. Susan Shinn performed most of the ultra-thin sectioning for transmission electron microscopy. Dr. Michael L. Kent provided the initial idea of this work and several PCR primers.

## Chapter 4

Dr. J. Genaro Sánchez -Martinez supplied gill material from brook trout. Dr. Robert Adlard provided gill material from Australian surf bream. Sheila C. Dawe helped to collect many Pacific and Atlantic Loma spp. isolates and maintained laboratory isolates of Loma salmonae. Dr. Michael L. Kent provided the initial idea for this study and supplied some ribosomal DNA primers and material from a number of L. salmonae isolates, outgroups Glugea sp . and Microsporidium prosopium.

## Chapter 5

Dr. Jim Boutillier suggested the initial idea for this study and provided isolates of smooth pink shrimp.

## Chapter 6

Dr. Charles R. Vossbrinck provided the initial idea for this study. Dr. Patrick J. Keeling amplified and sequenced $\alpha$ and $\beta$ - tubulin genes from this species (mentioned in the discussion).

## Dedication

In memory of my father
for his music, joy and inspiration,
and for sternly asking me to finish my thesis
while he was finishing his life's project.

## General Acknowledgements

I am indebted to: Dr. Shannon Bennett for last-minute PAUP-command counseling and comments on thesis drafts; my sister Emily for emergency ML-crunching and personal advise; my supervisor Dr. Martin Adamson for the fun, a free-handed approach, and for untangling my thoughts and words; my committee for their help when I asked for it; my parents for their everpresent financial generosity and patience; Graham Blair for much needed philosophical pep talks re Thomas Kuhn; Dr. Lance Barrett-Lennard for his sagely guidance at the beginning of my laboratory training; my other lab mates Allyson Miscampbell, Tammy Laberge, Steve Connor and Ross Shaw for their tolerance and support in the rain; and many others at UBC, PBS, on the W. E. Ricker, in Halifax, P.E.I., Connecticut, Brazil, and other meeting places, who shared their scientific spaces, music and wisdom. I am grateful to the developers of the bicycle - a mechanism that can promote imaginative thought, inspire innovation and got me up the hill. Finally, I must thank the butchered, but un-surrendering unknowns of the biological world from which all this wonder comes.

## Chapter 1: General Introduction

## Outline and structure of the thesis

This thesis examines relationships of microsporidians at various levels using genetic and morphological data. Chapters are organized to correspond with taxonomic groups and levels of relationships examined: Chapter 2 presents variation within a single species (Loma salmonae) from salmon, Chapters 3 and 4 explore relationships among species and genera related to genus Loma from fishes, while Chapters 5 and 6 look at species from crustaceans and examine deeper relationships in phylum Microsporidia.

Molecular data frequently expose problems in older classifications of the microsporidia, therefore all these chapters re-visit historical taxonomic disputes in light of the new information. Included in the thesis are descriptions of new species (Chapter 3), a new genus (Chapter 6), and an examination of species boundaries (Chapter 4). Microsporidian ribosomal DNA is characterized by frequent indel differences between species that presented difficulty during this study; therefore, several programs were designed (Appendix 1) and used in other chapters of the thesis. Research objectives specific to particular taxonomic groups are summarized at the end of this introductory chapter and are explained in detail in subsequent chapters.

This chapter introduces microsporidian biology and critical problems in epidemiology, evolution, taxonomy and systematics of the microsporidia to provide a background for issues addressed in Chapters 2 through 6. The tools and methods used are also introduced here.
Common themes that appear in several chapters are introduced here and will be expanded upon in the final chapter (Chapter 7).

## General features: unity and diversity of the microsporidia

Microsporidians are small, unicellular, spore-forming organisms, as the name suggests. They have been recognized as a distinct group since Balbiani and others, in the 1880s (see Wittner, 1999), first examined their unique apparatus for parasitic transmission: the polar filament, a long, coiled tube in spores that everts, upon appropriate stimulus, to pierce host cells and cause infection. The parasitic lifestyle of these organisms may be responsible for another remarkable feature of the microsporidia after the polar filament: the high diversity in developmental pattern and life-cycles (reviewed in Cali \& Takvorian, 1999). Although obligate
microparasites like these might be expected to develop diverse life cycles and developmental patterns under strong selection to develop within and transmit between their hosts, it is impressive that microsporidia appear to have accomplished such adaptations with remarkably little machinery. For example, these tiny (often $\sim 1-4 \mu \mathrm{~m}$ ), single-celled eukaryotes have the most extraordinarily reduced cellular, genetic, and biochemical complexity of all eukaryotes. They lack centrioles, $9+2$ microtubule structures, and peroxisomes, have simplified Golgi and highly reduced mitochondria (Vávra \& Larsson, 1999; Katinka et al., 2001; Williams et al., 2002). They have the smallest genomes of all eukaryotes (as little as 2.3 Mbp ), with as few as 1,997 coding genes (Biderre et al., 1994; Biderre et al., 1995; Peyretaillade et al., 1998; Katinka et al., 2001; Méténier \& Vivarès, 2001), severely reduced intron and extra-coding DNA (Katinka et al., 2001). They also have small, 70S bacteria-like ribosomes with a 5.8 S ribosomal RNA gene (rDNA) that is fused to the large subunit (LSU) without a second internal transcribed spacer (ITS-2) (Vossbrinck \& Woese, 1986), and they lack the tricarboxylic acid cycle and other typical eukaryotic biochemical pathways (Katinka et al., 2001).

The most commonly observed stages of microsporidia are the spores. These are small, but highly refractile (they appear to glow under a light microscope) and often form massive infections visible to the naked eye. For example, they can replace virtually all of a host's abdominal muscle (e.g. Thelohania butleri of Chapter 5) or fill large cyst-like structures (explained later) on external organs (e.g. Loma salmonae of Chapters 2, 3, 4). Spores are the only stage of the life cycle capable of surviving outside the intracellular environment of the host, and are protected by a thick, layered, chitinous spore wall, which gives them resistance to damage in the environment, sometimes for more than a year (e.g. in L. salmonae Shaw et al., 2000c, and Marssoniella elegans Komárek \& Vávra, 1968). Under light microscopy alone, spores can sometimes be used to distinguish species, as they tend to vary in shape and size ( $\sim 2$ 6 and rarely up to $30 \mu \mathrm{~m}$ ) or surface structure. Transmission electron microscopy (TEM) reveals internal features of spores (Fig. 1.1), which are used widely to distinguish species. However, light microscopic and TEM features of spores may overlap extensively among species such that they cannot be differentiated (Shaw et al., 1997; Weiss \& Vossbrinck, 1999), and so cryptic species are not uncommon.

Other developmental stages occur only within host cells and detection and diagnosis usually requires TEM or careful light microscopy with appropriate stains (Weber et al., 1999). These more elusive life-cycle stages inside host cells can proliferate by complex and variable
vegetative divisions (Fig. 1.2). For example, in a single microsporidian species the two major developmental phases, merogony and sporogony, may undergo different division sequences in one cell, in different cells of the same host, or in different hosts where several hosts are required to complete the life cycle (e.g. species of Amblyospora) (Becnel \& Andreadis, 1999). The basic life cycle, shown in Fig. 1.2, begins with the action of the polar filament (a coiled tube within spores) and associated extrusion structures (anchoring disc and polaroplast membranes) upon an appropriate signal from a nearby host cell (e.g. pH or salinity) (Keohane \& Weiss, 1999). The filament extrudes and forcibly pierces the host cell, and injects its nucleus and some cytoplasm, enclosed in a membrane, into the host cell cytoplasm (or nucleoplasm in Nucleospora species) through the hollow tube. This first stage, called the sporoplasm, grows and becomes known as a meront, which develops in intimate association with the host mitochondria, deriving ATP and other needs (see Weidner et al., 1999; Katinka et al., 2001) from the host. Meronts divide repeatedly in the new cell (merogony) often forming multinucleate plasmodia, and eventually become known as sporonts (sometimes sporogonial plasmodia if multinucleate) as they begin a phase destined to form spores (sporogony). Sporonts divide to produce sporoblasts, which develop without further division into spores. Details of these stages, such as number of nuclei per plasmodium, shape of plasmodium, and form of plasmodial division during both merogony and sporogony are species-specific. Thus, the developmental sequence can be useful in describing and differentiating species (see Chapters 3, 5 and 6).

Spores resulting from such divisions will either infect nearby host cells of the same individual or will pass to the environment through excretions or upon death of the host. In some species, multiple forms of spore are produced by one species (polyspory or heterospory) in one host (see Chapter 6), or in separate hosts. In such cases each spore usually specializes in infecting a particular host or cell (see Becnel \& Andreadis, 1999), so monospory or polyspory can be indicative of host-use. However, the presence of only one spore form in a host is not sufficient evidence that there is not a second host in which other spore forms are produced. Similarly, the presence of several spore forms in a single host may not indicate the requirement of a second host, as species may have spore forms that are remnants of a complex life cycle, which are now destined for autoinfection or horizontal transmission to the same host, or may even regularly produce abortive (non-viable) spores (e.g. Edhazardia aedis).

Spores and earlier dividing stages may be enclosed in an envelope of host origin (called a parasitophorous vacuole) or parasite origin (called a sporophorous vesicle). These membranous sacs may form during merogony (merogonic vesicle) or sporogony (sporogonic
vesicle), or may be absent in some species. Some species have single, isolated nuclei in spores, meronts or other stages, while other species may have two nuclei that are either isolated (binucleate) or are closely pressed together (a diplokaryon). Diplokaryons divide synchronously, and in some species these two closely associated nuclei fuse together (karyogamy) and undergo meiosis or begin a cycle with uninucleate division. The sexual cycle has been observed in several species, involving karyogamy (nuclear fusion), synaptonemal complex formation, mingling of chromosomes, and reductive nuclear divisions (meiosis) (Loubès et al., 1976; Loubès, 1979; Hazard \& Brookbank, 1984; Larsson, 1986; Chen \& Barr, 1995). Species that appear to have always isolated nuclei (like Loma spp. of Chapters 2 to 4) may be haploid and entirely asexual (Lom \& Nilsen, 2003). This is still speculative, as chromosome and DNA-quantity studies to assess ploidy and detailed study of the nuclear and chromosomal behaviour in uninucleate species have not been performed. Where known, sex, chromosomal behaviour, and ploidy in microsporidia are unusual, and are still poorly understood (Hazard \& Brookbank, 1984; Hazard et al., 1985; Canning, 1988; Haig, 1993; Chen \& Barr, 1995; Vávra \& Larsson, 1999). However, there is one study showing a microsporidian with always-isolated nuclei that forms synaptonemal complexes (see Vávra \& Larsson, 1999).

Microsporidia are also diverse in host and geographic distribution. They occur in both single-celled and multicellular organisms, most phyla of invertebrates and all five classes of vertebrates (Sprague, 1977; Canning \& Lom, 1986; Wittner, 1999), but they are not known to occur in photosynthetic organisms. The majority of the more than 1,200 known species are from arthropod or fish hosts, perhaps because infections are more readily observed in these hosts, or perhaps because these are naturally better hosts. Another possibility is that microsporidia diversified earliest in arthropods and fishes. The basal position of species having aquatic-crustacean or aquatic-larval insect hosts in some studies (Baker et al., 1997; Nilsen, 2000; Bell et al., 2001) suggested the earliest ancestors of microsporidia may have required an aquatic transmission route for the spores. This aquatic-origin hypothesis seems less likely in light of newer studies showing more basal taxa (e.g. Antonospora scoticae) from terrestrial arthropod hosts. These results differ depending on the method of analysis used (Weiss \& Vossbrinck, 1999), and need to be confirmed by more genetic loci and by comparison with more species. Despite differences in phylogenetic results among microsporidia, all molecular data and morphological data suggest microsporidia form a monophyletic group, characterized by having a polar filament, unusually reduced cellular complexity, and considerable developmental flexibility.

## Historical problems in biology and systematics of the microsporidia

## Ubiquitous pests, pest controls and curiosities

Microsporidia are ubiquitous pathogens, recognized as the cause of economic losses and pathology in fishes and insects long before the tools were available to examine details of their cellular structures (see Wittner, 1999). For example, in the mid-1800s the European silk industry was devastated by a microsporidian causing pébrine disease in silkworms - the first species to be named - Nosema bombycis Nägeli, 1857 (see Canning, 1990). Louis Pasteur later studied this species and in 1870 published the first study showing control of a protozoan disease (Canning, 1990; Wittner, 1999). Recently there has been renewed economic interest in microsporidia because several species occur in humans with AIDS or other immune compromised conditions (Desportes et al., 1985; Baker et al., 1995; Wittner, 1999; Kotler \& Orenstein, 1999; Franzen \& Müller, 1999), while others are major pathogens in fisheries or aquaculture (Shaw \& Kent, 1999). Some species are of interest for their use in biological control of pest insects (Becnel \& Andreadis, 1999), and others have been used as models in studies of the evolution of virulence and transmission (Ebert, 1994; Agnew \& Koella, 1997; Koella \& Agnew, 1997).

Beyond pathogenesis and commercial or human impact, microsporidia have long been considered as curiosities by systematists for their resemblance to primitive protists, fungi (e.g. chitinous spore walls), even myxosporeans (based on the coiled filament), and were difficult to place as a group (Sprague, 1977a). Since the polar filament was recognized in the 1880s, separating microsporidia from other organisms, they have been placed together and variously ranked as a subgroup of yeasts and bacteria (placed in Schizomycetes by Nägeli in 1857), a sister-group to or subgroup of the myxosporeans, and finally raised to phylum status by Sprague (1977a) and almost simultaneously by Weiser (1977). Sprague \& Becnel (1998) proposed the correct name for the phylum as Microsporidia Balbiani, 1882, instead of names previously used (phylum Microspora Sprague, 1977, or phylum Microsporida Weiser, 1977). Sorting out species into a suitable system of classification within the group has been even more difficult because of their reduced cellular features and size, and the expression. of a wide array of developmental and nuclear conditions (Sprague, 1977a; b) that are difficult to observe without (and sometimes even with) fine-grained tools (e.g. TEM and DNA studies).

## Microsporidia as Fungi: the problem of parasitic reduction

Prior to the late 1990 s, evidence suggested microsporidia were primitive, amitochondrial eukaryotic protists that diverged early, before the endosymbiotic acquisition of the mitochondrion (Cavalier-Smith, 1987). Several studies supported this theory by showing microsporidia at the base of the eukaryotes, e.g. using ribosomal DNA (rDNA) (Vossbrinck et al., 1987), translation elongation factors 1 and 2 (EF-1 $\alpha \& E F-2$ ) (Kamaishi et al., 1996) and isoleucyl-tRNA synthetase gene analyses (Brown \& Doolittle, 1996). Many recent studies (see Mathis, 2000) show microsporidia to be a later-branching group, located in the crown of the tree close to or within the Fungi. Evidence of this late-branching includes the presence of an organelle which now appears to be a true mitochondrial homologue (Williams et al., 2002) with corresponding mitochondrial genes (Hirt et al., 1997; Germot et al., 1997; Peyretaillade et al., 1998; Katinka et al., 2001; Fast \& Keeling, 2001). Phylogenetic analyses of several independent genes also consistently show microsporidia as a later-diverged eukaryote (Hirt et al., 1999; Keeling et al., 2000; Van de Peer et al., 2000; Katinka et al., 2001), and suggest previous results in which the group was shown to be early branching were erroneous, perhaps resulting from long-branch attraction. If these results are correct, and microsporidia are members of the Fungi, their highly simplified cellular and genetic structure compared to their fungal ancestors must be an extreme example of parasitic reduction.

## Towards a testable theory of parasitic reduction in the microsporidia

Parasitic reduction of the microsporidial genome and cellular structures may have arisen through random mutation and loss of genes (substitutions and deletions) that were not needed in the intracellular host environment. Alternatively, parasitic reduction may have arisen under strong selection for the parasite to become smaller, biochemically or morphologically hidden, or to replicate DNA or cellular components more quickly. The first of these ideas was proposed by Vivarès \& Méténier (2000) and Katinka et al. (2001). These authors suggested the loss of primary biochemical pathways in microsporidia that were no longer needed in the host environment during the early phases of microsporidial parasitism led to reduced protein-protein interactions, which led to a loss of inter-genic DNA regions that normally encode regulatory signals, and ultimately led to the dramatic compaction of microsporidial genomes.

The second explanation for parasitic reduction arises from studies of transmission modes in the microsporidia reviewed in Dunn et al. (2001) and references therein. These authors reviewed studies that examined, in detail, how microsporidia hide undetected in host cells in
vertically transmitted species by interacting intimately with the host cell cytoskeleton, spindle poles, microtubules and mitochondria, and divide in close synchrony with the mitotic chromosomes (see Terry et al., 1998). Dunn et al. (2001) suggested parasitic reduction might have occurred as a result of selection imposed by this close association in vertical transmission, for which theory predicts the parasite will usually be under greater selection to remain hidden (so as not harm its host). The closeness of the association and the strength of selection it imposes should be correlated with the relative importance of vertical transmission in species that transmit both vertically and horizontally. Furthermore, experimental studies by Sweeney et al. (1989), Iwano \& Kurtti (1995) and Koella et al. (1998) suggest that microsporidians can adapt under selective pressures to become more efficient at vertical transmission when horizontal transmission is restricted. They can condense the developmental sequence, or lose production of an alternate spore type. Although Dunn et al. (2001) did not elaborate on how such selection could cause a reduction in genome size, they suggest vertical transmission is ancestral in the microsporidia and may be the primary cause of parasitic reduction in the phylum Microsporidia.

These two hypotheses explaining parasitic loss in the microsporidia can be used to formulate alternate predictions. The first hypothesis, in which loss was due to a lack of need for genes during the early stages of evolution to parasitism, one might predict loss of genome size and associated characters (e.g. biochemical or structural) to be randomly distributed in a "ratchet-like" pattern of progressive loss along evolutionary lineages. Relatives in a lineage would therefore have similar genomic and cellular losses, regardless of ecological differences. In contrast, the second hypothesis, in which loss was due to the need to be small and hide (especially with respect to mode of transmission), one might predict loss of genome size and other characters should correlate with the dominant mode of transmission in a lineage or sublineage. Sister-groups with different dependence on vertical transmission (i.e. different selection to become reduced) could be compared in a phylogeny and the correlation among these parasitic constraints and differences in genome size and cellular, developmental, or biochemical complexity could be tested.

Several other environmental selective pressures could be important in parasitic reduction of the microsporidia. For example, differences in transmission opportunity or other ecological factors may cause some species to require production of larger numbers of propagules per host cell (to increase net reproduction, $\mathrm{R}_{0}$ ) without disruption of the cell until spores are mature. This constraint could impose selection for reduced size or developmental complexity and, as with
vertical transmission, the correlation with parasitic loss could be tested by a comparative method using phylogenetic analyses.

There are some data that suggest the first hypothesis - that loss is random and ratchetlike due to lack of need - is not supported. For example, among 13 species of microsporidia there is an 8.5 -fold difference in genome size and a wide range of chromosome numbers, but these follow no pattern with respect to host group, lineage, or morphological characters (Méténier \& Vivarès, 2001). This lack of pattern in genome size or karyotype does not support the first hypothesis, because under this hypothesis, sister-groups would have been expected to lose similar overall amounts of DNA. This pattern instead presents an argument for selection on genome size imposed by environmental or other factors that differ among related sister-groups. Glugea atherinae, with the largest known genome size ( 19.5 Mbp ), can be compared to Spraguea lophii, a species in the same larger fish-parasitic clade of the microsporidian tree, which has one-third the genome size ( 6.2 Mbp ). One could speculate first that the large genome of $G$. atherinae may be due to the lower selection to be small or hidden compared to other microsporidians, and propose as a reason, that instead of hiding physically (by being small and transmitting vertically) it enlists the host cell to form xenomas which are greatly enlarged hostparasite complexes able to contain and disseminate enormous numbers of spores for horizontal transmission -- so the large size of the xenoma may free up constraints on parasite size. Another possibility is that $G$. atherinae may not need to reduce genome and cell size because it forms sacs around sporogonal stages that protect it from being detected this way, rather than by reduced size. By contrast, S. lophii, which has a smaller genome, also forms xenomas, but has all developing stages in direct contact with the host cytoplasm. Perhaps this direct contact increases the constraints on size (to avoid detection by the host cell), and therefore correlates with three-fold reduction in genome size compared to G. atherinae. Other possibilities are that S. lophii may transmit vertically, or may have transmitted vertically in its recent ancestry. Evidence for the latter is suggested by the production of two spore types in $S$. lophii, a trait often associated with past or present use of a second host and possibly vertical transmission in one of the hosts. In contrast, G. atherinae and other relatives produce just one spore type.

It is also possible that DNA has been gained, rather than lost, in some lineages, perhaps where selection for reduced size has been "relaxed" due to some feature of the host environment. There is some evidence that DNA gain may be important. For example, ribosomal DNA (rDNA) differences between closely related species (e.g. Edhazardia aedis and Amblyospora spp.) sometimes appear non-homologous, suggesting DNA sequence can be
gained by insertions in some species. Furthermore, the appearance of high evolutionary rates in small subunit (SSU) rDNA in species with larger genome sizes (e.g. N. bombycis) might be explained by accumulation of non-homologous inserts among relatives.

These alternatives need to be systematically examined by a comparative (Poulin, 1998) or experimental approach to test the alternate explanations for the dramatic genomic and cellular losses that characterize the microsporidia.

## Higher classification in flux: the problem of rapidly evolving morphology

Higher classification within phylum Microsporidia has been in flux as alternate authors have emphasized different single characters as important indicators of major groups (Sprague, 1977a; Weiss \& Vossbrinck, 1999 and references therein). For example, the classification systems of Tuzet et al. (1971), Sprague (1977a; b), Issi (1986), Larsson (1986), and Sprague et al. (1992) placed different emphasis on the membranes around spores, the form of sporogony, the nuclear condition, characters including spore shape and origin of the membrane around developing stages (host vs. parasite), and the chromosome cycle. Sprague (1977b) wisely suggested that the use of characters of certain classes for particular ranks of taxonomy is problematic, so organisms should be differentiated and ranked by as many characters as possible, even if this means each individual category of characters (e.g. membranes around spores or divisions during development) will not form a monophyletic group.

Even the best morphology-based classification systems that incorporate several characters do not closely reflect evolutionary relationships of the microsporidia suggested by molecular data (Baker et al., 1995; Weiss \& Vossbrinck, 1999). Baker et al. (1995) presented an SSU rDNA tree of 16 microsporidians on which they plotted morphological characters and showed that this tree was inconsistent with previous morphology-based classification systems, suggesting commonly used characters for higher-level classification were polyphyletic, and so inappropriate, at least when used alone. Updated rDNA analyses (Weiss \& Vossbrinck, 1999; Lom \& Nilsen, 2003) with more species and gamma-correction for long-branch attraction differ slightly, but confirm the overall conclusion of Baker et al. (1995). Results from another gene, RNA polymerase largest subunit II (RPB1) are consistent with the overall rDNA results (Cheney et al., 2001), confirming again that previous morphology-based classifications of the microsporidia require revision.

Lom \& Nilsen (2003) suggested that a new classification system based on DNA sequence data would be necessary. However, data are still lacking for some important groups (e.g. Metchnikovellidae), suggesting it is too early to devise a completely new system. The
gradual accumulation of molecular data has, instead, allowed for the re-classification of microsporidia in small steps, with single species being transferred or placed in new genera based on rDNA sequence. Lom \& Nilsen (2003) proposed some larger re-groupings based on rDNA, but did not have a solution for large, morphologically heterogeneous groups like "Group 3" (see Chapter 4), which contains Loma species, the subjects of Chapters 2 to 4 of this thesis. A good system of classification for the microsporidia is not only important for systematists, evolutionary biologists and taxonomists, but could help researchers in epidemiology, biochemistry and immunology to look for species with comparable biology to serve as models.

## Systematics of closely related species

For some 90 years, from the 1880s when the first microsporidian species were differentiated until the 1970s when the first TEM studies of microsporidia appeared, researchers were limited to light microscopy. This made differentiation of fine characters difficult; for example, Lemmermann (1900) viewed the bright-blue-green colour of refractile spores of Marssoniella elegans as an indication that these specimens (actually microsporidia) were bluegreen algae (see Chapter 6). Nevertheless, slight differences between specimens in a single character were commonly found and used to describe new species or even genera (see Sprague, 1977b). Canning (1990) suggested the practice of discriminating species and genera based on single characters (light microscopy or TEM) has led to overestimates of diversity for the microsporidia, i.e. excessive splitting. Molecular data that has accumulated over the past 10 years for more than 70 species of microsporidia suggest the opposite situation may have occurred, wherein, using only light microscopy, many taxa were excessively "lumped."

Microsporidia pose significant problems for species-identification even when using finegrained tools such as TEM and DNA, because of the high cost, time, and technical expertise to gather and interpret, and characterize such data across a species' range. Typically such data are reported for only a single infected host or a single tissue sample in species descriptions. Larsson (1999) discussed these problems and emphasized the need for researchers to present detailed TEM or DNA features for several individuals or isolates, and more importantly to give diagnostic characters that can be used in the field (i.e. light microscopic characters), without the need for laborious and expensive TEM or DNA confirmation of characters. While such thorough data cannot always be obtained due to time, financial, technical, and geographic sampling limitations, it might be argued that, short of this, species descriptions are virtually useless. Therefore, in Chapters 2 to 4 of this thesis, individuals were sampled from multiple localities for TEM, DNA and LM study, where possible.

Studies of closely related microsporidians rarely address what is meant by the word "species" (i.e. the operational species concept) probably because other questions need to be answered first, in particular, the question of whether microsporidia are sexual diploids or asexual haploids. Taxonomists of microsporidia are at an early stage in their understanding and characterization of the basic biology, reproduction, and distribution of variation in characters compared to taxonomists of some other groups (e.g. mammals, fishes), so detailed discussion of suitable species concepts may be premature. However, it can be argued that a useful species concept need not depend on the reproductive behaviour of the organism, and with the increasing use of DNA sequence in microsporidian species descriptions, lineage-based species concepts may be applicable and useful (see Sites \& Crandall, 1997). In many cases, DNA sequence may be the most practical or the only available character (Weiss \& Vossbrinck, 1999) for distinguishing close species of microsporidia. Researchers studying closely related microsporidians should attempt to sample DNA from across the range of a species, and to consider more carefully the use of lineage-based (cladistic) methods for defining species, e.g. those reviewed in Quicke (1993) and applied in numerous studies (Sites \& Crandall, 1997; Van Oppen et al., 2000; Rodriquez-Robles \& de Jesus-Escobar, 2000; Diekmann et al., 2001; Puorto et al., 2001; Bradley \& Baker, 2001; Lopez et al., 2002; Van Oppen et al., 2002; Chen et al., 2002; Bernardi et al., 2002). Chapters 2 to 4 and 7 develop and discuss operational species concepts appropriate to the data in this study.

However, phylogenetic analyses of DNA data cannot address most of the important biological and ecological questions we have about closely related species (e.g. host breadth, virulence, etc.); so, other data are essential. Transmission to reciprocal hosts is still an excellent test of reproductive boundaries (if species are sexual), particularly for microsporidia, which often undergo several generations within a single host. Nevertheless, DNA studies could be useful for assessing the size, distribution and duration of the "spore bank", or the supply of viable, resistant spores in the environment (see Didier \& Bessinger, 1999; Snowden \& Shadduck, 1999; Becnel \& Andreadis, 1999), and the reservoir of infection in alternate hosts. This period of spore viability in the environment or the use of reservoir hosts are essential epidemiological parameters and are important for understanding evolution and speciesboundaries formation. Molecular data combined with ecological data can help identify aspects of life cycles, such as use of alternate hosts or transmission routes (Vossbrinck et al., 1998).

## Introduction to Loma spp. and microsporidia from fishes

Microsporidians from fishes show some incongruence between morphology-based and molecular-based classification (Lom, 2002; Lom \& Nilsen, 2003) and thereby serve as a model group in which to study morphological plasticity and parasitic loss, processes that appear to characterize the phylum Microsporidia. For example, there have been several surprising results from rDNA sequencing suggesting some species with very different morphology belong as sister species (e.g. Heterosporis anguillarum and Glugea anomala in Nilsen et al., 1998; Ichthyosporidium and Loma in Nilsen, 2000; and Kabatana takedai, Spraguea and Microgemma in Lom et al., 2000); or suggesting species be renamed (e.g. Pleistophora sp. PA moved to Tuzetia weidneri and Pleistophora sp. LS to Perezia nelsoni in Canning et al., 2002; Glugea americanus moved to Spraguea in Pomport-Castillon et al., 2000; Glugea acerinae moved to Loma in Lom \& Pekkarinen, 1999; and then removed from Loma in Lom \& Nilsen, 2003; and that Ichthyosporidium sp. may not be congeneric with Ichthyosporidium giganteum in Karlsbakk et al., 2001).

There have been several reviews of the 14 genera and 156 species of microsporidians so far described from fishes (Canning \& Lom, 1986; Shaw \& Kent, 1999; Lom, 2002; Lom \& Nilsen, 2003). All microsporidians from fishes group together in molecular studies with just one exception so far (Nucleospora spp.) (Docker et al., 1997b; Nilsen et al., 1998; Nilsen, 1999; Weiss \& Vossbrinck, 1999; Nilsen, 2000; Bell et al., 2001; Matthews et al., 2001; Nilsen \& Chen, 2001, Lom \& Nilsen, 2003). Most species from fishes have isolated nuclei throughout development and form only one type of spore, although several exceptions exist. Such species are therefore thought to be directly transmitted; however, direct transmission has only been experimentally demonstrated in a few species so far (McVicar, 1975; Baxa-Antonio et al., 1992; Shaw et al., 1998; Kent \& Bishop-Stewart, 2003; Lom \& Nilsen, 2003; Lee et al., 2004). Several characters including the xenoma, sporophorous vesicle, and the possession of a diplokaryon appear to have arisen or been lost more than once in fish-parasitic microsporidia (Nilsen et al., 1998; Lom \& Nilsen, 2003), and are therefore no longer considered good characters for classification at the family-and genus-level. Although all species from fishes except Nucleospora spp. group together, a handful of species from non-fish hosts also fall within this clade. Thus, as with most groups of microsporidia, the fish-parasitic species have no single defining character that distinguishes them as a group (Lom \& Nilsen, 2003) and older classification appears to be based on polyphyletic characters.

Not surprisingly, the larger genera Glugea Thélohan, 1891, with 38 species, and Pleistophora Gurley, 1893, with 33 species, have proven to be polyphyletic groups (Nilsen et al., 1998; Nilsen, 2000; Pomport-Castillon et al., 2000; Matthews et al., 2001; Cheney et al., 2001; Lom \& Nilsen, 2003). In light of recent molecular data, it seems more species should have been placed in the collective group Microsporidium Balbiani, 1884, that was designed as a temporary holding place for species of uncertain placement, instead of using Pleistophora and Glugea, which appear too broadly defined at this time. Although the holding group Microsporidium already includes 47 species from fishes, it may still be preferential to other designations until DNA data are available. Given the difficulty with morphological classification, it may be necessary to define genera and species (or higher levels of classification) using DNA sequence in some part of the definition. In such cases DNA data must be available from the type species. Type species of Glugea, Pleistophora, and six of the other 12 genera from fishes have now been represented with SSU rDNA sequence. For the remaining six genera, comprising 33 species, DNA sequence from the type species has not been published.

In Chapter 4, DNA sequence is presented from Loma Morrison \& Sprague, 1981, the largest genus from fishes for which sequence was lacking from the type species. Genus Loma was created for species that were formerly lumped with Glugea, Pleistophora or Nosema and have xenomas usually in the gills, with a thin, amorphous xenoma wall and developmental stages intermixed in the xenoma. Genus Loma was erected by Morrison \& Sprague (1981a) for a specimen described from Atlantic cod (Gadus morhua), L. morhua Morrison \& Sprague, 1981, and designated this as the type species. In the same publication, specimens from haddock (Melanogrammus aeglefinus) first described by Nemeczek (1911) were transferred to genus Loma by Morrison \& Sprague (1981a), to make the new combination, Loma branchialis. (Nemeczek, 1911) Morrison \& Sprague, 1981. Loma branchialis had previously been placed in Nosema, then Glugea. Various authors who later examined material from haddock, the type host of $L$. branchialis, were not able to conclude with confidence that it matched Nemeczek's (1911) description for L. branchialis (originally Nosema branchiale Nemeczek, 1911) (Morrison \& Sprague, 1981a; b; Morrison \& Marryatt, 1986; Morrison, 1983). Therefore, these authors proposed L. morhua as the name for specimens in Atlantic cod and sometimes also haddock, but they recognized the possibility that other material from haddock or at other localities may be $L$. branchialis rather than L. morhua. Several authors (see Kabata, 1959; and see Morrison \& Sprague, 1981a) who had examined similar microsporidia from Atlantic cod, haddock, and other
gadids expressed similar uncertainty about the identity of specimens from different hosts and localities. There appeared to be at least two distinct spore and xenoma sizes in specimens from gadid hosts (see Kabata, 1959). The solution proposed by Morrison \& Sprague (1981a; b) was to consider $L$. morhua as the type species and $L$. branchialis as a separate species, although the differences were slight. Canning \& Lom (1986) found the latter solution ambiguous, and considered $L$. morhua as a junior synonym of $L$. branchialis, thus making the latter species the type species. Chapters 3 and 4 re-consider this issue in greater detail, presenting DNA and spore data.

Since genus Loma was created, it has expanded to include 14 named species distributed broadly, in Africa, South America, Russia, Europe, Australia, and the northern Atlantic and Pacific oceans, in an array of hosts (11 families of fishes, in 4 orders) (see Table 1.1). These species are typically described using characters of the xenoma, site of infection (gills or gut), and features of the sac formed around dividing stages. Several experimental studies on a single species, L. salmonae (Putz, Hoffmann \& Dunbar, 1965), from rainbow trout and Pacific salmon (Oncorhynchus spp.), have provided some understanding of the host specificity (Speare et al., 1998b; Shaw \& Kent, 1999; Shaw et al., 2000a; b; Ramsay et al., 2002), transmission (Kent et al., 1995; Shaw et al., 1998), life cycle (Sánchez et al., 1999; Sánchez et al., 2001c), development (Beaman et al., 1999; Sánchez et al., 2000), host-response (Speare et al., 1989; Speare et al., 1998c; Kent et al., 1999; Beaman et al., 1999; Mustafa, 2000; Shaw et al., 2001), and treatment (Higgins et al., 1998) of infection. To build biological hypotheses about species of Loma other than L. salmonae, considering these studies' results, it would be useful to know the relationships among Loma species. Ribosomal DNA studies suggest morphological relationships among members of genus Loma may be wrong (Lom \& Nilsen, 2003). For example, L. acerinae, which resembles "true" Loma species according to Lom \& Pekkarinen (1999), turns out to be an outlier that is not closely related to other Loma species. Furthermore, Lom \& Pekkarinen (1999) suggested that based on differences in the origins of the sac around developing stages, several other species of Loma appear to be quite unlike the type species. Thus, there are both rDNA and morphological arguments that the genus Loma is polyphyletic. Prior to the present study, DNA sequence was lacking for all but four described species ( $L$. salmonae, L. embiotocia, L. acerinae, and Loma sp. of Nilsen, 2000); therefore Chapter 4 examines DNA sequence from other Loma species (see Table 1.1), including the type species, to address the integrity of genus Loma.

Variation in DNA sequence, morphology and other biological properties (e.g. host preference, virulence) have rarely been characterized for individuals across the range of a microsporidian species; however, several studies show this variation exists and can be useful in understanding other observations (Didier et al., 1995; Hollister et al., 1995; Didier et al., 1996a; Didier et al., 1996b; Breitenmoser et al., 1999; Biderre et al., 1999; Gresoviac et al., 2000). In particular, these studies of intraspecific DNA variation have revealed microsporidian strains separated by geographic locality or host. Chapters 2 to 4 examine intraspecific variation among isolates of Loma species towards addressing such subdivision, and also examine species boundaries and the possibility of reservoir hosts. Chapter 2 begins by attempting to characterize sequence difference (at two genetic loci) across many isolates L. salmonae, a widespread, serious pathogen in commercial salmon that has been well-studied experimentally (Magor, 1987; Speare et al., 1989; Kent et al., 1990; Markey et al., 1994; Bruno et al., 1995; Bader et al., 1998; Kent et al., 1998; Kent, 2000). This study of variation in L. salmonae will provide a backdrop against which variation in less well-known species of Loma, examined in Chapters 3 and 4 , may be compared.

## Introduction to microsporidia from crustaceans

Microsporidians from crustaceans are ideal models for studying the evolution of microsporidia, particularly life-cycle or morphological adaptations in response to the host. For example, while many crustacean-parasitic species possess complex life cycles (e.g. Amblyospora-like species) in which the crustacean is the intermediate host, other species in crustaceans are directly transmitted (e.g. Thelohania contejeani). Some groups appear to have co-evolved with their hosts (e.g. species from copepods) while others have not (e.g. species from cladocerans) (Baker et al., 1998; Refardt et al., 2002; Moodie et al., 2003). Furthermore, many potentially important microsporidians from crustaceans have yet to be discovered, as evidenced by the frequency of discovery of new species (Larsson, 1996; 1999; Vossbrinck et al., 2004) in commercial crustaceans and copepod intermediate hosts of pest mosquitoes. Another important crustacean-parasitic microsporidian has recently been found in the sea louse, Lepeophtheirus salmonis, a salmon-parasitic copepod that itself is a significant pathogen (Freeman et al., 2003). Chapters 5 and 6 examine microsporidia from crustaceans (shrimp and copepod) in the context of host-parasite co-evolution and morphological and life-cycle adaptations, while examining taxonomic and classification questions.

Unfortunately no single work has reviewed all species of microsporidia found in crustacean hosts since Sprague (1977c). Sprague's (1977c) list of 18 genera and 88 species shows taxonomic diversity to be high for microsporidia from Copepoda ( 11 genera, 24 species) (Sprague, 1977b). For microsporidia from Decapoda, genus diversity is lower whereas species diversity is higher ( 8 genera, 32 species), but sampling effort is uneven and may explain such differences. Two other crustacean groups Amphipoda and Cladocera, known to act as hosts for more than a handful of microsporidians, have slightly fewer species, at 7 genera, 17 species, and 7 genera, 14 species, respectively (Sprague, 1977b; c). Sprague's (1977b) list is long outdated, and Larsson (1996) suggested there might be more than 130 species described from crustaceans to date. Larsson's $(1996 ; 1999)$ reviews of the literature to the genus level, suggest about $1 / 4$ of microsporidian genera ( 25 of 110) have type species in crustacean hosts, 16 of these being in aquatic microcrustaceans (copepods and cladocerans), although the importance of groups in these hosts has not been reflected in DNA studies or taxonomic reviews. Presumably, these groups have not been studied as well because of their distance from human or commercially important species; however, Refardt et al. (2002) showed species in aquatic crustaceans might sometimes be closely related to species with more economic or human interest.

Molecular data suggest microsporidians in crustaceans are a polyphyletic group (Refardt et al., 2002, Moodie et al., 2003) with species dispersed throughout the tree, which is in contrast to previous studies, suggesting microsporidians fall into clades corresponding to major host groups (e.g. fishes, aquatic insects, terrestrial insects) with only a few exceptions (Baker et al., 1995; Baker et al., 1997; Weiss \& Vossbrinck, 1999; Nilsen \& Chen, 2001; Lom \& Nilsen, 2003). While this pattern is unusual in the microsporidia, it is not unusual in parasites (Poulin, 1998). Incongruence between host and parasite phylogenies can occur for ecological and phylogeographic reasons. For example, when hosts colonize new localities their parasites may go extinct due to differences in transmission opportunity or sampling effects due to the patchiness of parasite populations (i.e. parasites tend to be even more aggregated than their hosts), thus, leaving host niches open to new parasite invasions (i.e. enabling host switching or host capture). The tendency for some species to be highly specific to host species while others remain more general is also expected to influence the pattern of host-parasite co-evolution, and while one could speculate that microsporidians in aquatic crustaceans may have different hostspecificity, there is no data for most species.

Past or present use of crustaceans as intermediate hosts in two-host life cycles may explain the phylogenetic dispersion of crustacean-parasitic microsporidia. To date, 13 species of
microsporidia have been shown to require a crustacean as an intermediate host (Becnel \& Andreadis, 1999, Micieli et al., 2000, and references therein; Andreadis \& Vossbrinck, 2002, Vossbrinck et al., 2004) and, as direct transmission has rarely been demonstrated, many other species (particularly polysporous species) are suspected to require two hosts. The only two-host life-cycles thus far demonstrated involve a freshwater cyclopoid copepod and a mosquito final host. The presence of more than one spore form (polyspory) in other species is ambiguous; it may be evidence of spores specialized for a second host, for environment vs. autoinfection, for infection of alternate tissues, or for vertical vs. horizontal transmission. Similarly, the presence of a single form of spores (monospory) is not clear evidence of requirement for a single host, because a second, specialized spore form may have been missed or not produced in the host under study. Specialized spores may also look identical. Chapter 6 examines a new species from a copepod in light of these issues.

Chapter 5 examines a species resembling Thelohania butleri from its type host (shrimp) and locality, to address evolutionary and taxonomic questions. Thelohania species are heterogeneous with respect to morphology and life-cycle. Some species may be directly transmitted with only horizontal transmission, while others have vertical and horizontal transmission (e.g. T. solenopsae). Thelohania species generally begin development (merogony) with diplokaryotic nuclei and then produce uninucleate sporoblasts and spores, but this may not be universal, as the nuclear condition has not been carefully studied in most species. Sporogony varies widely (Sprague, 1977b) among species. Molecular data (Moodie et al., 2003) from several Thelohania species and the very wide host distribution of the more than 75 species in this genus (Sprague, 1977b; Larsson, 1999) also strongly suggest the genus is polyphyletic, confirming indications by others (Hazard \& Oldacre, 1975; Sprague et al., 1992; Larsson, 1999) that the genus is too broadly defined and needs revision.

## Introduction to tools of study

## Genes chosen

Ribosomal RNA genes (rDNA) are widely used and highly informative markers of both higher- and lower-level relationships in microsporidia. The multiple copy number of these genes improves the chances of amplification from small amounts of tissue, lightly infected tissue, poorly preserved material, or mechanically damaged DNA (e.g. from bead-beaten spores). Because of this ease of amplification and the interspersion of highly conserved and highly variable sequence regions, rDNA genes have proven highly informative for both
epidemiological studies and deep-phylogeny of the microsporidia (Vossbrinck et al., 1993; Zhu et al., 1994; Baker et al., 1994; Weiss et al., 1994; Baker et al., 1995; Barlough et al., 1995; Baker et al., 1997; Docker et al., 1997a; Keeling \& McFadden, 1998; Vossbrinck et al., 1998; Fries et al., 1999; Adler et al., 2000; Müller et al., 2000; Gresoviac et al., 2000; Khattra et al., 2000; Brown \& Kent, 2002).

Ribosomal DNA has several well-known difficulties that are particularly problematic in microsporidia. The two most important of these are: difficulty with alignment in non-conserved regions due to high sequence variation and large numbers of indels in or near loop-regions (Brown \& Adamson, 2000); and multiple gene copies (paralogs) in a single genome (Bell et al., 2001; Cheney et al., 2001). Erroneous phylogenies may result if gene copies are not compared to those that are homologous by descent (orthologs), i.e. copies that are homologous by duplication in ancestors (paralogous copies) may not correctly resolve nodes after the duplication event. Another complicating feature of rDNA in phylogenetic analysis is that the multiple copies in a genome can evolve in concert or may diverge from one another in different species-lineages to varying degrees, depending on the processes of mutation, recombination and homogenization among copies (i.e. replication slippage and unequal crossing over, or gene conversion). For some species, rDNA copies are located in tandem on a chromosome and mutations will be homogenized by unequal crossing over (Dover, 1982), whereas in other species, copies are dispersed on different chromosomes and are homogenized less efficiently, or not at all. In microsporidia, rDNA is dispersed differently in closely related species (Peyretaillade et al., 1998; Vivarès \& Méténier, 2000; Méténier \& Vivarès, 2001). Perhaps because of the differential dispersion, rDNA evolves at different rates in different lineages of microsporidia; for example, it appears to evolve more slowly among fish-parasitic microsporidians compared to other microsporidians (Nilsen et al., 1998; Cheney et al., 2000; Nilsen \& Chen, 2001). Despite these problems, the small subunit (SSU), internal transcribed spacer (ITS), and large subunit (LSU) have been found to be sufficiently informative to reveal important features of microsporidian evolution and show consistency with other data (e.g. RPB1, a single-copy gene).

The potential difficulties with rDNA (or any multi-copy gene) make it desirable to have data from an independent, single-copy nuclear gene. Single-copy genes may be, in some ways, even more problematic, particularly because they can be hard to amplify from small amounts or low quality material. It can also be difficult to design PCR primers for single-copy genes because they are not as well represented in Genbank for microsporidia or close outgroups to the
microsporidia. When this work began, there were relatively few microsporidian gene sequences in Genbank, and these were mostly from human- or mammal-parasitic species that are only distant relatives of groups to be examined in this study (e.g. fish-parasitic or crustacean-parasitic species). Recently, the complete genome sequence for the microsporidian E. cuniculi has become available (Katinka et al., 2001), making it easier to design primers to target a great number of single-copy genes; however, E. cuniculi is also a distant relative of taxa examined in this study. Therefore, two genes that had been characterized for microsporidia in the fishparasitic clade were chosen: translation elongation factor-1 alpha (EF-1 $\alpha$ ) and RNA polymerase largest subunit II (RPB1). EF-1 $\alpha$ has been characterized in a wide range of phylogenetic studies (Hashimoto et al., 1994; Cho et al., 1995; Kamaishi, et al. 1996; Yamamoto, et al., 1997; Hashimoto et al., 1997), and has been sequenced in the fish-parasitic microsporidian, Glugea plecoglossi. Recent analysis in greater detail (Roger et al., 1999; Moreira et al., 1999; Inagaki et al., 2003), suggests that while this gene may be useful, it interacts with the cytoskeleton and other translation machinery, most importantly rRNA, and may occur in multiple copies in some eukaryotes. It also may evolve at rates that are inconsistent among lineages, though such inconsistencies have not yet been characterized within the microsporidia. The gene was used in Chapters 2 and 4 to complement rDNA results, with the understanding that rate differences or multiple copies (paralogs) may be found. Very few EF-1 $\alpha$ sequences are available from microsporidia; therefore this study helps expand the database for this gene. Studies suggest RPB1 is more likely to be a strictly single-copy gene than is EF-1 $\alpha$ (Allison et al. 1985; Klenk et al. 1995; Cermakian et al. 1996; Morse et al. 1996; Croan \& Ellis, 1996; Stiller \& Hall, 1997; and Reiger \& Shultz, 1997). This gene has recently been shown to give results congruent to those for gamma-corrected rDNA phylogenies (Hirt et al., 1999; Cheney et al., 2001), and so RPB1 was chosen (Chapter 4) to provide evidence independent of rDNA data.

## Phylogenetic analysis methods and Modeltest (Posada \& Crandall, 1998)

In order to examine the robustness of phylogenetic results in light of the underlying models of evolution, phylogeny reconstruction was performed using all three major kinds of optimality criteria: maximum parsimony (MP), minimum evolution distance (ME), and maximum likelihood (ML) in PAUP* Phylogenetic Analysis Using Parsimony (Swofford, 2000). Because each data set may be unique in how closely it conforms to or violates assumptions of each of these methods, application of all three methods can be informative, and used as a means of data-exploration, as well as an attempt to estimate the true phylogeny. For
example, where topologies differ among methods, a low estimated shape parameter (calculated in ML) would suggest strong rate heterogeneity and favour ML trees over MP and ME trees. An unusual MP tree topology, compared to ME or ML trees, would suggest there might be longbranch attraction, and favour results from the other methods. Differences in ME trees compared to other methods may suggest base-frequency differences among species are significant (Swofford, 2000).

To further examine each data set and improve ML analyses by specifying a more realistic model of DNA substitution, Modeltest (Posada \& Crandall, 1998) was used. This widely used program evaluates the fit of the data to a nested set of 64 models with and without assumptions such as equal base frequencies, equal transition or transversion rates, rates equal among sites, no invariable sites, and includes all of the simple, commonly used models (e.g. JC $=$ Jukes and Cantor; K80 = Kimura; HKY; GTR = general time reversible, etc.) that have been proposed over the years, using these parameters. The resulting best-fit model and calculated parameters estimated from the data were used as the basis for calculating starting trees for ML analyses in PAUP*, and after that, parameters suggested to be important in the best model were re-estimated from the data for final ML tree searches (Appendix 2 shows a sample input file for Modeltest; Appendix 3 shows a sample output file from Modeltest; and Appendix 4 shows a sample PAUP* input command file for incorporating Modeltest results into ML analyses).

Monophyly constraints and hypothesis testing with AU and other tests (Shimodaira \& Hasegawa, 2001)

A set of phylogenetic trees can be compared by several likelihood-based methods to evaluate the relative statistical confidence in each tree, using the Kishino-Hasegawa (KH) test, the Shimodaira-Hasegawa (SH) test, and the Approximately Unbiased (AU) test (Kishino \& Hasegawa, 1989; Goldman et al., 2000; Shimodaira \& Hasegawa, 2001; Shimodaira, 2002). These methods resolve more finely the statistical confidence in nodes (topologies) of interest than the familiar percent bootstrap support statistics routinely reported on branches. These methods also allow one to compare a set of hypothetical topologies to assess which topologies are statistically most likely. While the KH, SH, and weighted versions of these tests (WKH and WSH) have been demonstrated to have significant problems with biases (Shimodaira, 2002), the AU test was shown to be nearly unbiased (Shimodaira \& Hasegawa, 2001). This test involves a multi-scale bootstrap procedure devised by Efron et al. (1996) and expanded on by Shimodaira (2000), which re-samples the data (bootstrap re-sampling with replacement), but changes sequence lengths in each replicate. The procedure was shown to correct bias that occurs in other
statistical bootstrap-based methods (see details in Shimodaira, 2000, and Chapter 4).
Shimodaira (2002) also demonstrated that the AU test was less likely to reject correct trees than the KH test (i.e. type 1 error), and is not over-conservative like the SH test (i.e. accepting too many trees).

The AU and other such tests in the CONSEL software package of Shimodaira \& Hasegawa (2001) were applied in this study to test a variety of hypothetical trees based on phylogenetic predictions relating to species boundaries or corresponding to various host groups (e.g. under a host-parasite co-evolution hypothesis) or taxonomic groups of the parasite. Appendix 5 gives further details on the creation of hypothetical trees and construction of formal hypotheses. Appendix 7 shows a sample file in which monophyletic clades were forced in PAUP* for later analysis in CONSEL. Appendix 8 shows the resulting output file that is used as input in CONSEL and Appendix 9 shows a sample CONSEL output. Goldman et al. (2000) argued that the widespread use of the KH and SH tests to select the best set of trees from a set of equally most likely trees found during a search (e.g. heuristic ML search) violates an important assumption of these tests - that the trees under consideration must be chosen a priori, separate from any information in the data (see Appendix 6, and Goldman et al., 2000). Shimodaira (2002) did not mention that any such a priori rule applies in use of the AU test, but the rule of $a$ priori may still be important for this test, based on the basic principles discussed in Goldman et al. (2000). Therefore, in Chapters 4, $5 \& 6$, where these tests were used, alternative trees were constructed by constraining groups according to a priori hypotheses, and, although the initial topologies were selected using heuristic ML searches, the set of trees being compared differed only at nodes determined by the a priori hypotheses. The merits of this approach and the degree to which it conforms to or violates assumptions of these tests are discussed in Appendix 6.

## Software design for analysis of gaps (indels), double signal, and rDNA stems

Ideas for three software programs emerged during analysis of rDNA sequences for this study of microsporidia. I designed these programs (described in detail in Appendix 1), and code was written by Michael Coury for Java 2 (JDK 1.2 or J2SE). The first version of the software package, called REALEM, was completed in May 2000. The simplest program, "Flip Analyzer," was devised to extract sequences from DNA traces that have long stretches of double peaks that have been caused by an insertion or deletion (indel) difference between two amplified copies of an rDNA or other gene. The second program, "Gap Matrix," can be used to create a data matrix that carries phylogenetically informative information on the location of indels or
alignment gaps (actually gap end positions) that can then be analyzed by maximum parsimony in PAUP*.

These two programs may be useful in studies of microsporidian rDNA, which has many indels (alignment gaps) that would normally have to be deleted, along with sometimesinformative sequences. Flip Analyzer was used in Chapters 2 to 6 to retrieve sequences for analysis and Gap Matrix was used in Chapters $4,5 \& 6$ for data-exploration, particularly to examine paralogs in Chapter 4.

The third program, "Stem States," was designed for analysis of rDNA stem regions according to a substitution model explained in Appendix 1. Its use on microsporidian sequences has been demonstrated elsewhere (Brown \& Adamson, 2000), and is presented only in Appendix 1 for interest as a tool that is part of the REALEM package.

## Light and transmission electron microscopy (TEM)

Light microscopy of fresh or fixed material, sectioned histological preparations, or resinembedded sections is often sufficiently informative for diagnosing infection, particularly with appropriate stains (see Weber et al., 1999). It can be sufficiently informative for distinguishing species, and can be faster and cheaper than TEM. Light microscopy was always used as the first step in examining specimens for the presence of xenomas or spores (Chapters 2 to 6 ); however, for characterizing new species (Chapters $3 \& 6$ ) and confirming identity with previously described species (Chapter 5), TEM provides far superior detail. Meronts, internal features of spores, and host and parasite cytoplasmic features are virtually impossible to detect and differentiate without TEM. The most significant weakness of TEM, compared to light microscopy as a tool for study of microsporidia, besides the cost, time, and appropriate facilities, is the inability to represent three-dimensionality. Many sequential ultrathin sections (serial sectioning) must be made and examined to obtain an understanding, or measurements of the dimensions of structures. For example, cylindrical merogonial plasmodia in Loma species will be difficult to assess (will appear as uninucleate round cells) in some sectioning planes. Fortunately, the plane-of-section effect can be diminished, along with the number of sections needed to understand three-dimensional structures by looking at heavily infected tissues, which have a greater probability of revealing replicate structures cut randomly in different planes. TEM poses greater problems for rare stages and smaller structures (for which serial sections may go right through the object of interest), or for lighter infections when there is a lower likelihood of having structures cut in enough different planes to present a clear picture of the structures. These problems will be discussed in subsequent chapters.

## Objectives

Using genetic variation within Loma salmonae (a serious pathogen of Pacific salmon, Oncorhynchus spp.) and 10 other Loma species from fishes, and reconstructing and statistically evaluating phylogenetic relationships within and among Loma species, and also among relatives of Thelohania butleri (a pathogen of shrimp), and among relatives of a new microsporidian found in a cyclopoid copepod, this thesis attempts to (1) test the integrity of strains, species and genera with genetic and morphological data, and use results to distinguish and describe new species; (2) characterize diagnostically useful developmental, spore, and DNA features and identify morphological characters that are most and least problematic in classification of particular groups of taxa; (3) examine ecological or epidemiological features (e.g. prevalence, intensity, signs of host response) that may contribute to understanding each disease; (4) estimate evolutionary relationships, and use this information to develop biogeographic and coevolutionary hypotheses. Furthermore, Chapter 7 (General Conclusions) will re-visit themes introduced in this chapter, illuminating how the results of this thesis help us to (5) form better questions about how these unusual organisms maintain plasticity while having such a reduced morphology and genome.

## LITERATURE CITED

Adler, P. H., Becnel, J. J. and Moser, B. 2000. Molecular characterization and taxonomy of a new species of Caudosporidae (Microsporidia) from black flies (Diptera: Simuliidae) with host-derived relationships of the North American caudosporids. Journal of Invertebrate Pathology 75:133-143.

Agnew, P. and Koella, J. C. 1997. Virulence, parasite mode of transmission, and host fluctuating asymmetry. Proceedings of the Royal Society of London Series B 264:9-15.

Allison, L. A., Moyle, M., Shales, M. and Ingles, C. J. 1985. Extensive homology among the largest subunits of eukaryotic and prokaryotic RNA polymerases. Cell 42:599-610.

Andreadis, T. G. and Vossbrinck C. R. 2002. Life cycle, ultrastructure and molecular phylogeny of Hyalinocysta chapmani (Microsporidia: Thelohaniidae), a parasite of Culiseta melanura (Diptera: Culicidae) and Orthocyclops modestus (Copepoda: Cyclopidae). Journal of Eukaryotic Microbiology 49(4):350-64.

Azevedo, C. and Matos, E. 2002. Fine structure of a new species, Loma myrophis (Phylum Microsporidia), parasite of the Amazonian fish Myrophis playrhynchus (Teleostei, Ophichthidae). European Journal of Protistology 37:445-452.

Bader, J. A., Shotts, E. B., Steffens, W. L. and Lom, J. 1998. Occurrence of Loma cf. salmonae in brook, brown and rainbow trout from Buford Trout Hatchery, Georgia, USA. Diseases of Aquatic Organisms 34:211-216.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Andreadis, T. G. 1998. Phylogeny of Amblyospora (Microsporida: Amblyosporidae) and related genera based on small subunit ribosomal DNA data: a possible example of host parasite cospeciation. Journal of Invertebrate Pathology 71:199-206.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Maddox, J. V. 1997. Phylogenetic position of Amblyospora Hazard \& Oldacre (Microspora: Amblyosporidae) based on small subunit rRNA data and its implication for the evolution of the microsporidia. Journal of Eukaryotic Microbiology 44(3):220-225.

Baker, M. D., Vossbrinck, C. R., Didier, E. S., Maddox, J. V. and Shadduck, J. A. 1995. Small subunit ribosomal DNA phylogeny of various microsporidia with emphasis on AIDS related forms. Journal of Eukaryotic Microbiology 42(5):564-570.

Baker, M. D., Vossbrinck, C. R., Maddox, J. V. and Undeen, A. H. 1994. Phylogenetic relationships among Vairimorpha and Nosema species (Microspora) based on ribosomal RNA sequence data. Journal of Invertebrate Pathology 64:100-106.

Barlough, J. E., McDowell, T. S., Milani, A., Bigornia, L., Slemenda, S. B., Pieniazek, N. J. and Hedrick, R. P. 1995. Nested polymerase chain reaction for detection of Enterocytozoon salmonis genomic DNA in chinook salmon Oncorhynchus tshawytscha. Diseases of Aquatic Organisms 23:17-23.

Baxa-Antonio, D., Groff, J. M. and Hedrick, R. P. 1992. Experimental horizontal transmission of Enterocytozoon salmonis to chinook salmon, Oncorhynchus tshawytscha. Journal of Protozoology 39:699-702.

Beaman, H. J., Speare, D. J., Brimacombe, M. and Daley, J. 1999. Evaluating protection against Loma salmonae generated from primary exposure of rainbow trout, Oncorhynchus mykiss (Walbaum), outside of the xenoma-expression temperature boundaries. Journal of Fish Diseases 22:445-450.

Becnel, J. J. and Andreadis, T. G. 1999. Microsporidia in insects. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 447501.

Bekhti, M. and Bouix, G. 1985. Loma salmonae (Putz, Hoffman et Dunbar, 1965) et Loma diplodae n. sp., microsporidies parasites de branchies de poissons téléosteens: implantation et données ultrastructurales. Protistologica 21(1):47-59.

Bell, A. S., Aoki, T. and Yokoyama, H. 2001. Phylogenetic relationships among microsporidia based on rDNA sequence data, with particular reference to fish-infecting Microsporidium Balbiani 1884 species. Journal of Eukaryotic Microbiology 48(3):258-265.

Bernardi, G., Holbrook, S. J., Schmitt, R. J., Crane, N. L. and DeMartini, E. 2002. Species boundaries, populations and colour morphs in the coral reef three-spot damselfish (Dasycyllus trimaculatus) species complex. Proceedings of the Royal Society of London Series B 269:599-605.

Biderre, C., Mathis, A., Deplazes, P., Weber, R., Méténier, G. and Vivarès, C. P. 1999. Molecular karyotype diversity in the microsporidian Encephalitozoon cuniculi. Parasitology 118:439-445.

Biderre, C., Pages, M., Méténier, G., Canning, E. U. and Vivarès, C. P. 1995. Evidence for the smallest nuclear genome ( 2.9 Mb ) in the microsporidium Encephalitozoon cuniculi. Molecular and Biochemical Parasitology 74:229-231.

Biderre, C., Pages, M., Méténier, G., David, D., Bata, J., Prensier, G. and Vivarès, C. P. 1994. On small genomes in eukaryotic organisms: molecular karyotypes of two microsporidian species (Protozoa) parasites of vertebrates. Comptes Rendues de l'Academie des Sciences. Paris Life Science 317:300-304.

Bradley, R. D. and Baker, R. J. 2001. A test of the genetic species concept: Cytochrome-b sequences and mammals. Journal of Mammalogy 82(4):960-973.

Breitenmoser, A. C., Mathis, A., Bürgi, E., Weber, R. and Deplazes, P. 1999. High prevalence of Enterocytozoon bieneusi in swine with four genotypes that differ from those identified in humans. Parasitology 118:447-453.

Brown, A. M. V. and Adamson, M. L. 2000. Microsporidian evolution doing cart-wheels: new rDNA phylogenies using gap matrices and stem models. Program guide and abstracts - 75th annual meeting of the American Society of Parasitologists. June 24-28, 2000, San Juan, Puerto Rico. p. 61.

Brown, J. R. and Doolittle, W. F. 1996. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proceedings of the National Academy of Sciences USA 92:2441-2445.

Brown, A. M. V. and Kent M. L. 2002. Molecular diagnostics for Loma salmonae and Nucleospora salmonis (microsporidia) In Molecular diagnostics of salmonid diseases. Cunningham, C. O. (ed.). Kluwer Academic Publishers, Dordrecht p. 267-283.

Bruno, D. W., Collins, R. O. and Morrison, C. M. 1995. The occurrence of Loma salmonae (Protozoa: Microspora) in farmed rainbow trout, Oncorhynchus mykiss Walbaum, in Scotland. Aquaculture 133:341-344.

Cali, A. and Takvorian, P. M. 1999. Developmental morphology and life cycles of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 85-128.

Canning, E. U. 1988. Nuclear division and chromosome cycle in microsporidia. BioSystems 21:333-340.
Canning, E. U. 1990. Microspora. in Margulis, L., J. O. Corliss, M. Melkonian, D. J. Chapman, ed., Handbook of Protoctista. Bartlett, Boston. p.53-72.

Canning, E. U. and Lom, J. 1986. The microsporidia of vertebrates. Academic Press, London. 289 pp.
Canning, E. U., Curry, A. and Overstreet, R. M. 2002. Ultrastructure of Tuzetia weidneri sp. n. (Microsporidia: Tuzetiidae) in skeletal muscle of Litopenaeus setiferus and Farfantepenaeus aztecus (Crustacea: Decapoda) and new data on Perezia nelsoni (Microsporidia: Pereziidae) in L. setiferus. Acta Protozoologica 41:63-77.

Cavalier-Smith, T. 1987. The origin of eukaryote and archaebacterial cells. Annals of the New York Academy of Sciences 503:17-54.

Cermakian, N., Ikeda, T. M., Cedergren, R. and Gray, M. W. 1996. Sequences homologous to yeast mitochondrial and bacteriophage T3 and T7 RNA polymerases are widespread throughout the eukaryotic lineage. Nucleic - Acids Research 24(4):648-654.

Chen, C. A., Chen, C.-P., Fan, T.-Y., Yu, T.-K. and Hseih, H.-L. 2002. Nucleotide sequences of ribosomal internal transcribed spacers and their utility in distinguishing closely related Perinereis polychaetes (Annelida; Polychaeta; Nereidae). Marine Biotechnology 4:17-29.

Chen, W. and Barr, A. R. 1995. Chromosomal evidence on the sporogony of Amblyospora californica (Microspora: Amblyosporidae) in Culex tarsalis (Diptera: Culicidae). Journal of Eukaryotic Microbiology 42(2):103-108.

Cheney, S. A., Lafranchi-Tristem, N. J., Bourges, D. and Canning, E. U. 2001. Relationships of microsporidian genera, with emphasis on the polysporous genera, revealed by sequences of the largest subunit of RNA polymerase II (RPB1). Journal of Eukaryotic Microbiology 48(1):111-117.

Cheney, S. A., Lafranchi-Tristem, N. J. and Canning, E. U. 2000. Phylogenetic relationships of Pleistophora-like microsporidia based on small subunit ribosomal DNA sequences and implications for the source of Trachipleistophora hominis infections. Journal of Eukaryotic Microbiology 47:280-287.

Cho, S., Mitchell, A., Regier, J. C., Mitter, C., Poole, R. W., Friedlander, T. P. and Zhao, S. 1995. A highly conserved nuclear gene for low-level phylogenetics: elongation factor- $1 \alpha$ recovers morphology-based tree for heliothine moths. Molecular Biology and Evolution 12(4):650-656.

Croan, D. and Ellis, J. 1996. Phylogenetic relationships between Leishmania, Viannia and Sauroleishmania inferred from comparison of a variable domain within the RNA polymerase II largest subunit gene. Molecular and Biochemical Parasitology 79:97-102.

Desportes, I., Le Charpentier, Y., Calian, A., Bernard, F., Cochand-Priollet, B., Lavergne, A., Ravisse, P. and Modigliani, R. 1985. Occurrence of a new microsporidian: Enterocytozoon bieneusi n. g., n. sp., in the enterocytes of a human patient with AIDS. Journal of Protozoology 26:179-187.

Didier, E. S. and Bessinger, G. T. 1999. Host-parasite relationships in microsporidiosis: animal models and immunology. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 225-257.

Didier, E. S., Vossbrinck, C. R., Baker, M. D., Rogers, L. B., Bertucci, D. C. and Shadduck, J. A. 1995. Identification and characterization of three Encephalitozoon cuniculi strains. Parasitology 111:411-421.

Didier, E. S., Rogers, L. B., Orenstein, J. M., Baker, M. D., Vossbrinck, C. R., Van Gool, T., Hartskeerl, R., Soave, R. and Beaudet, L. M. 1996a. Characterization of Encephalitozoon (Septata) intestinalis isolates cultured from nasal mucosa and bronchoalveolar lavage fluids of two AIDS patients. Journal of Eukaryotic Microbiology 43:34-43.

Didier, E. S., Visvesvara, G. S., Baker, M. D., Rogers, L. B., Bertucci, D. C., DeGroote, M. A. and Vossbrinck, C. R. 1996b. A microsporidian isolated from an AIDS patient corresponds to Encephalitozoon cuniculi III, originally isolated from domestic dogs. Journal of Clinical Microbiology 34:2835-2837.

Diekmann, O. E., Bak, R. P. M., Stam, W. T. and Olsen, J. L. 2001. Molecular genetic evidence for probable reticulate speciation in the coral genus Madracis from a Caribbean fringing reef slope. Marine Biology 139:221-233.

Docker, M. F., Devlin, R. H., Richard, J., Khattra, J. and Kent, M. L. 1997a. Sensitive and specific polymerase chain reaction assay for detection of Loma salmonae (Microsporea). Diseases of Aquatic Organisms 29:4148.

Docker, M. F., Kent, M. L., Hervio, D. M. L., Khattra, J. S., Weiss, L. M., Cali, A. and Devlin, R. H. 1997 b. Ribosomal DNA sequence of Nucleospora salmonis Hedrick, Groff and Baxa, 1991 (Microsporea:

Enterocytozoonidae): implications for phylogeny and nomenclature. Journal of Eukaryotic Microbiology 44(1):55-60.

Dover, G. A. 1982. Molecular drive: a cohesive mode of species evolution. Nature 299:111-117.
Dunn, A. M., Terry, R. S. and Smith, J. E. 2001. Transovarial transmission in the microsporidia. Advances in Parasitology 48:57-100.

Ebert, D. 1994. Virulence and local adaptation of a horizontally transmitted parasite. Science 265:1084-1086.
Efron, B., Halloran, E. and Holmes, S. 1996. Bootstrap confidence levels for phylogenetic trees. Proceedings of the National Academy of Sciences USA 93:13 429-13 434.

Fast, N. M. and Keeling, P. J. 2001. $\alpha$ and $\beta$ subunits of pyruvate dehydrogenate E1 from the microsporidium Nosema locustae: mitochondrion-derived carbon metabolism in microsporidia. Molecular and Biochemical Parasitology 117:201-209.

Faye, N., Toguebaye, B. S. and Bouix, G. 1995. On the cytology and development of Loma boopsin. sp. (Microspora, Glugeidae), parasite of Boops boops (Pisces, Teleostei, Sparidae) from the coasts of Senegal. Archiv Fur Protistenkunde 146:85-93.

Fomena, A., Coste, F. and Bouix, G. 1992. Loma camerounensis new species (Protozoa: Microsporida) a parasite of Oreochromis niloticus Linnaeus 1757 Teleost Cichlidae in fish-rearing ponds in Melen Yaounde Cameroon. Parasitology Research 78(3):201-208.

Franzen, C. and Müller, A. 1999. Molecular techniques for detection, species differentiation, and phylogenetic analysis of microsporidia. Clinical Microbiology Reviews 12(2):243-285.

Freeman, M. A., Bell, A. S. and Sommerville, C. 2003. A hyperparasitic microsporidian infecting the salmon louse, Lepeophtheirus salmonis: an rDNA-based molecular phylogenetic study. Journal of Fish Diseases 26:667676.

Fries, I., Paxton, R. J., Tengö, J., Slemenda, S. B., da Silva, A. J. and Pieniazek, N. J. 1999. Morphological and molecular characterization of Antonospora scoticae n. gen., n. sp. (Protozoa, Microsporidia) a parasite of the communal bee, Andrena scotica Perkins, 1916 (Hymenoptera, Andrenidae). European Journal of Protozoology 35:183-193.

Germot, A., Philippe, H. and Le Guyader, H. 1997. Evidence for loss of mitochondria in microsporidia from a mitochondrial-type HSP70 in Nosema locustae. Molecular and Biochemical Parasitology 87:159-68.

Goldman, N., Anderson, J. P. and Rodrigo, A. G. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49(4):652-670.

Gresoviac, S. J., Khattra, J. S., Nadler, S. A., Kent, M. L., Devlin, R. H., Vivarès, C. P., de la Fuente, E. and Hedrick, R. P. 2000. Comparison of small subunit ribosomal RNA gene and internal transcribed spacer sequences among isolates of the intranuclear microsporidian Nucleospora salmonis. Journal of Eukaryotic Microbiology 47(4):379-387.

Haig, D. 1993. Alternatives to meiosis: the unusual genetics of red algae, microsporidia, and others. Journal of Theoretical Biology 163:15-31.

Hashimoto, T., Nakamura, T., Nakamura, F., Shirakura, T., Adachi, J., Goto, N., Okamoto, K. and Hasegawa, M. 1994. Protein phylogeny gives a robust estimation for early divergences of eukaryotes: phylogenetic place of a mitochondria-lacking protozoan, Giardia lamblia. Molecular Biology and Evolution 11(1):65-71.

Hashimoto, T., Nakamura, Y., Kamaishi, T. and Hasegawa, M. 1997. Early evolution of eukaryotes inferred from protein phylogenies of translation elongation factors $1 \alpha$ and 2 . Archiv fuer Protistenkunde 148:287-295.

Hazard, E. I. and Brookbank, J. W. 1984. Karyogamy and meiosis in an Amblyospora sp. (Microspora) in the mosquito Culex salinarius. Journal of Invertebrate Pathology 44:3-11.

Hazard, E. I. and Oldacre, S. W. 1975. Revision of microsporida (Protozoa) close to Thelohania, with descriptions of one new family, eight new genera, and thirteen new species. U. S. Department of Agriculture Technical Bulletin 1530. U. S. Department of Agriculture, Washington, D. C.

Hazard, E. I., Fukuka, T. and Becnel, J. J. 1985. Gametogenesis and plasmogamy in certain species of Microspora. Journal of Invertebrate Pathology 46:63-69.

Higgins, M. J., Kent, M. L., Moran, J. D. W., Weiss, L. M. and Dawe, S. C. 1998. Efficacy of the fumagillin analog TNP-470 for Nucleospora salmonis and Loma salmonae infections in chinook salmon Oncorhynchus tshawytscha. Diseases of Aquatic Organisms 34:45-49.

Hirt, R. P., Healy, B., Vossbrinck, C. R., Canning, E. U. and Embley, T. M. 1997. Identification of a mitochondrial Hsp70 orthologue in Vairimorpha necatrix: molecular evidence that microsporidia once contained mitochondria. Current Biology 7:995-998.

Hirt, R. P., Logsdon, J. M., Healy, B., Dorey, M. W., Doolittle, W. F. and Embley, T. M. 1999. Microsporidia are related to fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proceedings of the National Academy of Sciences USA 96:580-585.

Hollister, W. S., Canning, E. U., Colbourn, N. I. and Aarons, E. J. 1995. Encephalitozoon cuniculi isolate from the urine of an AIDS patient, which differs from canine and murine isolates. Journal of Eukaryotic Microbiology 42:367-372.

Inagaki, Y., Blouin, C., Susko, E. and Roger, A. J. 2003. Assessing functional divergence in $\mathrm{EF}-1 \alpha$ and its paralogs in eukaryotes and archaebacteria. Nucleic Acids Research 31(14):4227-4237.

Issi, I. V. 1986. Microsporidia as a phylum of parasitic protozoa, p. 6-136 In Beyer, T. V. and Issi, I. V. (eds) Protozoology, vol. 10. Nauka, Leningrad, USSR. [in Russian with English summary]

Iwano, H. and Kurtti, T. J. 1995. Identification and isolation of dimorphic spores from Nosema furnacalis (Microspora: Nosematidae). Journal of Invertebrate Pathology 65:230-236.

Kabata, Z. 1959. On two little-known microsporidia of marine fishes. Parasitology 49:309-315.
Kamaishi, T., Hashimoto, T., Nadamura, Y., Nakamura, F., Murata, S., Okada, N., Okamoto, K., Shimizu, M. and Hasegawa, M. 1996. Protein phylogeny of translation elongation factor EF-1 $\alpha$ suggests microsporidians are extremely ancient eukaryotes. Journal of Molecular Evolution 42:257-263.

Karlsbakk, E., Askeland, J. and Plarre, H. 2001. Microsporidian infections in two Norwegian wrasse species (Labridae), Symphodus melops (L.) and Labrus bergylta Ascanus. 10th International Conference of the EAFP, Trinity College, Dublin, 9-14 September.

Katinka, M. D., Duprat, S., Cronillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J. and Vivarès, C. P. 2001. Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. Nature 414:450-453.

Keeling, P. J., Luker, M. A. and Palmer, J. D. 2000. Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. Molecular Biology and Evolution 17(1):23-31.

Keeling, P. J. and McFadden, G. I. 1998. Origins of microsporidia. Trends in Microbiology 6:19-23.
Kent, M. L. 2000. Marine netpen farming leads to infections with some unusual parasites. International Journal for Parasitology 30:321-326.

Kent, M. L. and Bishop-Stewart, J. K. 2003. Transmission and tissue distribution of Pseudoloma neurophilia (Microsporidia) of zebrafish, Danio rerio (Hamilton). Journal of Fish Diseases 26:423-426.

Kent, M. L., Dawe, S. C. and Speare, D. J. 1995. Transmission of Loma salmonae (Microsporea) to chinook salmon in sea water. Canadian Veterinary Journal 36:98-101.

Kent, M. L., Docker, M., Khattra, J., Vossbrinck, C.R., Speare, D. J. and Devlin, R. H. 1999. A new Microsporidium sp. (Microsporidia) from the musculature of the Mountain Whitefish Prosopium williamsoni from British Columbia: morphology and phylogeny. Journal of Parasitology 85(6):1114-1119.

Kent, M. L., Groff, J. M., Traxler, G. S., Zinkl, J. G. and Bagshaw, J. W. 1990. Plasmacytoid leukemia in seawater reared chinook salmon Oncorhynchus tshawytscha. Diseases of Aquatic Organisms 8:199-209.

Kent, M. L., Traxler, G. S., Kieser, D., Richard, J., Dawe, S. C., Shaw, R. W., Prosperi-Porta, G., Ketcheson, J. and Evelyn, T. P. T. 1998. Survey of salmonid pathogens in ocean-caught fishes in British Columbia, Canada. Journal of Aquatic Animal Health 10:211-219.

Keohane, E. M. and Weiss, L. M. 1999. The structure, function, and composition of the microsporidian polar tube. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 196-224.

Khattra, J. S., Gresoviac, S. J., Kent, M. L., Myers, M. S., Hedrick, R. P. and Devlin, R. H. 2000. Molecular detection and phylogenetic placement of a microsporidian from English sole (Pleuronectes vetulus) affected by X-cell pseudotumors. Journal of Parasitology 86(4):867-871.

Kishino, H. and Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29:170-179.

Klenk, H.-P., Zillig, W., Lanzendorfer, M., Grampp, B. and Palm, P. 1995. Location of protist lineages in a phylogenetic tree inferred from sequences of DNA-dependent RNA polymerases. Archiv fur Protistenkunde. 145:221-230.

Koella, J. C. and Agnew, P. 1997. Blood-feeding success of the mosquito Aedes aegypti depends on the transmission route of its parasite Edhazardia aedis. Oikos 78:311-316.

Koella, J. C., Agnew, P. and Michalakis, Y. 1998. Coevolutionary interactions between host life histories and parasite life cycles. Parasitology 116, S47-S55.

Komárek, J. and Vávra, J. 1968. In memoriam of Marssoniella Lemm, 1900. Archiv fuer Protistenkunde 111(1):1217.

Kotler, D. P. and Orenstein, J. M. 1999. Clinical syndromes associated with microsporidiosis. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 258-292.

Larsson, J. I. R. 1986. Ultrastructure, function and classification of microsporidia. p. 325-390. In. J. O. Corliss and D. J. Patterson (eds.) Progress in Protistology, vol. 1. Biopress, Ltd., Bristol, United Kingdom.

Larsson, J. I. R. 1996. Microsporidia in microcrustaceans. Scandinavian Section Society of Protozoologists 18th Annual Meeting. 202.

Larsson, J. I. R. 1999. Identification of microsporidia. Acta Protozoologica 38:161-197.
Lee, S.- J., Yokoyama, H. and Ogawa, K. 2004. Modes of transmission of Glugea plecoglossi (Microspora) via the skin and digestive tract in an experimental infection model using rainbow trout, Oncorhynchus mykiss (Walbaum). Journal of Fish Diseases 27:435-444.

Lemmermann, E. 1900. Beiträge zur Kenntnis der Planktonalgen. Berichte der Deutschen Botanischen Gesellschaft 18:272-275.

Lom, J. 2002. A catalogue of described genera and species of microsporidians parasitic in fish. Systematic Parasitology 53:81-99.

Lom, J. and Nilsen, F. 2003. Fish microsporidia: fine structural diversity and phylogeny. International Journal for Parasitology 33:107-127.

Lom, J., Dyková, I., Wang, C. H., Lo, C. F. and Kou, G. H. 2000. Ultrastructural justification of the transfer of Pleistophora anguillarum Hoshina, 1959 to the genus Heterosporis Schubert, 1969. Diseases of Aquatic Organisms 43:255-231.

Lom, J. and Pekkarinen, M. 1999. Ultrastructural observations on Loma acerinae (Jírovec, 1930) comb. nov. (Phylum Microsporidia). Acta Protozoologica 38:61-74.

Lopez, J. V., Peterson, C. L., Willoughby, R., Wright, A. E., Enright, E., Zoladz, S., Reed, J. K. and Pomponi, S. A. 2002. Characterization of genetic markers for in vitro cell line identification of the marine sponge Axinella corrugata. The Journal of Heredity 93(1):27-36.

Loubès, C. 1979. Recherchés sur la meiose chez les Microsporidies: Consequences sur les cycles biologiques. Journal of Protozoology 26:200-208.

Loubès, C., Maurand, J., Gasc, C., de Buron, I. and Barral, J. 1984. Étude ultrastructurale de Loma dimorphan. sp., microsporidie parasite de poissons gobiidae languedociens. Protistologica 14(4):579-589.

Loubès, C., Maurand, J. and Rousset-Galangau, V. 1976. Présence de complexes synaptonématiques dans le cycle biologique de Gurleya chironimi Loubès et Maurand, 1975: un argument en faveur d'une sexualité chez les Microsporidies? Comptes Rendus de l'Academie des Sciences 282:1025-1027.

Magor, B. G. 1987. First report of Loma sp. (Microsporida) in juvenile coho salmon (Oncorhynchus kisutch) from Vancouver Island, British Columbia. Canadian Journal of Zoology 65:751-752.

Markey, P. T., Blazer, V. S., Ewing, M. S. and Kocan, K. M. 1994. Loma sp. in salmonids from the eastern United States: associated lesions in rainbow trout. Journal of Aquatic Animal Health 6:318-328.

Mathis, A. 2000. Microsporidia: emerging advances in understanding the basic biology of these unique organisms. International Journal for Parasitology 30:795-804.

Matthews, J. L., Brown, A. M. V., Larison, K., Bishop-Stewart, J. K. and Kent, M. L. 2001. Pseudoloma neurophilia, n. gen., n. sp., a new microsporidium from the central nervous system of the zebrafish. Journal of Eukaryotic Microbiology 48:227-233.

McVicar, A. H. 1975. Infection of plaice Pleuronectes platessa L. with Glugea (Nosema) stephani (Hagenmüller 1899) (Protozoa: Microsporidia) in a fish farm and under experimental conditions. Journal of Fish Biology 7:611-619.

Méténier, G. and Vivarès, C. P. 2001. Molecular characteristics and physiology of microsporidia. Microbes and Infection 3:407-415.

Micieli, M. V., García, J. J. and Becnel, J. J. 2000. Horizontal transmission of Amblyospora albiafasciati Garcia and Becnel, 1994 (Microsporidia: Amblyosporidae), to a copepod intermediate host and the neotropical mosquito, Aedes albifasciatus (Macquart, 1837). Journal of Invertebrate Pathology 75:76-83.

Moodie, E. G., Le Jambre, L. F. and Katz, M. E. 2003. Thelohania montirivulorum sp. nov. (Microspora: Thelohaniidae), a parasite of the Australian freshwater crayfish, Cherax destructor (Decapoda: Parastacidae): fine ultrastructure, molecular characteristics and phylogenetic relationships. Parasitology Research 91:215-228.

Morrison, C. M. 1983. The distribution of the microsporidian Loma morhua in tissues of the cod Gadus morhua L. Canadian Journal of Zoology 61:2155-2161.

Morrison, C. M. and Marryatt, V. 1986. Further observations on Loma morhua Morrison \& Sprague, 1981. Journal of Fish Diseases 9:63-67.

Morrison, C. M. and Sprague, V. 1981a. Electron microscopical study of a new genus and new species of microsporida in the gills of Atlantic cod Gadus morhua L. Journal of Fish Diseases 4:15-32.

Morrison, C. M. and Sprague, V. 1981 b . Light and electron microscope study of microsporida in the gill of haddock, Melanogrammus aeglefinus (L.). Journal of Fish Diseases 4:179-184.

Morrison, C. M. and Sprague, V. 1983. Loma salmonae (Putz, Hoffman and Dunbar, 1965) in the rainbow trout, Salmo gairdneri Richardson, and L. fontinalis sp. nov. (Microsporidia) in the brook trout, Salvelinus fontinalis (Mitchill). Journal of Fish Diseases 6:345-353.

Moreira, D., Le Guyader, H. and Philippe, H. 1999. Unusually high evolutionary rate of the elongation factor $1 \alpha$ genes from the ciliophora and its impact on the phylogeny of eukaryotes. Molecular Biology and Evolution 16(2):234-245.

Morse, R., Collins, M., D. O'Hanlon, K., Wallbanks, S. and Richardson, P. T. 1996. Analysis of the $\beta^{\prime}$ subunit of DNA-dependent RNA polymerase does not support the hypothesis inferred from 16 S rRNA analysis that Oenococcus oeni (Formerly Lueconostocoenos) is a tachytelic (fast-evolving) bacterium. International Journal of Systematic Bacteriology 46(4): 1004-1009.

Müller, A., Trammer, T., Chioralia, G., Seitz, H. M., Diehl, V. and Fanzen, C. 2000. Ribosomal RNA of Nosema algerae and phylogenetic relationship to other microsporidia. Parasitology Research 86:18-23.

Mustafa, A., Speare, D. J., Daley, J., Conboy, G. A. and Burka, J. F. 2000. Enhanced susceptibility of seawater cultured rainbow trout, Oncorhynchus mykiss (Walbaum), to the microsporidian Loma salmonae during a primary infection with the sea louse, Lepeophtheirus salmonis. Journal of Fish Diseases 23:337-341.

Nemeczek, A. 1911. Beiträge zur Kenntnis der Myxo- und Microsporidien der Fishce. Archiv für Protistenkunde 22:143-169.

Nilsen, F. 1999. Small subunit rDNA phylogeny of Bacillidium sp. (Microspora, Mrazekiidae) infecting oligochaetes. Parasitology 118:553-558.

Nilsen, F. 2000. Small subunit ribosomal DNA phylogeny of microsporidia with particular reference to genera that infect fish. Journal of Parasitology 86(1):128-133.

Nilsen, F. and Chen, W. J. 2001 . rDNA phylogeny of Intrapredatorus barri (Microsporida: Amblyosporidae) parasitic to Culex fuscanus Wiedemann (Diptera: Culicidae). Parasitology 122:617-623.

Nilsen, F., Endresen, C. and Hordvick, I. 1998. Molecular phylogeny of microsporidians with particular reference to species that infect the muscles of fish. Journal of Eukaryotic Microbiology 45:535-543.

Ovcharenko, N. O., Sarabeev, V. L., Wita, I. and Czaplińska, U. 2000. Loma mugili sp. n., a new microsporidium from the gills of grey mullet (Mugil soiuy). Vestnik zoologii 34(4-5):9-15.

Peyretaillade, E., Biderre, C., Peyret, P., Duffieux, F., Méténier, G., Gouy, M., Michot, B. and Vivarès, C. P. 1998. Microsporidian Encephalitozoon cuniculi, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. Nucleic Acids Research 26(15):35133520.

Pomport-Castillon, C., De Jonkheere, J. F. and Romestand, B. 2000. Ribosomal DNA sequences of Glugea anomala, G. stephani, G. americanus and Spraguea lophii (Microsporidia): phylogenetic reconstruction. Diseases of Aquatic Organisms 40:125-129.

Posada, D. and Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14(9):817-818.

Poulin, R. 1998. Evolutionary ecology of parasites. Chapman \& Hall, London, UK. 212 pp.
Puorto, G., da Graça Salomão, M., Theakston, R. D. G., Thorpe, R. S., Warrell, D. A. and Wüster, W. 2001. Combining mitochondrial DNA sequences and morphological data to infer species boundaries: phylogeography of lanceheaded pitvipers in the Brazilian Atlantic forest, and the status of Bothrops pradoi (Squamata: Serpentes: Viperidea). Journal of Evolutionary Biology 14:527-538.

Putz, R. E., Hoffman, G. L. and Dunbar, C. E. 1965. Two new species of Pleistophora (Microsporidia) from North American fish with a synopsis of Microsporidia of freshwater and euryhaline fishes. Journal of Protozoology 12(2): 228-236.

Quicke, D. L. J. 1993. Principles and techniques of contemporary taxonomy. Blackie Academic \& Professional. London. 311 pp .

Ramsay, J. M., Speare, D. J., Dawe, S. C. and Kent, M. L. 2002. Xenoma formation during microsporidial gill disease of salmonids caused by Loma salmonae is affected by host species (Oncorhynchus tshawytscha, $O$. kisutch, $O$. mykiss) but not by salinity. Diseases of Aquatic Organisms 48:125-131.

Reiger, J. C. and Shultz, J. W. 1997. Molecular phylogeny of the major arthropod groups indicates polyphyly of crustaceans and a new hypothesis for the origin of hexapods. Molecular Biology and Evolution 14(9):902913.

Refardt, D., Canning, E. U., Mathis, A., Cheney, S. A., Lafranchi-Tristem, N. J. and Ebert, D. 2002. Small subunit ribosomal DNA phylogeny of microsporidia that infect Daphnia (Crustacea: Cladocera). Parasitology 124:381-389.

Rodriquez-Robles, J. A. and de Jesus-Escobar, J. M. 2000. Molecular systematics of new world gopher, bull, and pinesnakes (Pituophis: Colubridae), a transcontinental species complex. Molecular Phylogenetics \& Evolution 14(1):35-50.

Roger, A. J., Sandblom, O., Doolittle, W. F. and Philippe, H. 1999. An evaluation of elongation factor $1 \alpha$ as a phylogenetic marker for eukaryotes. Molecular Biology \& Evolution 16(2):218-233.

Sánchez, J. G., Speare, D. J. and Markham, R. J. F. 1999. Nonisotopic detection of Loma salmonae (Microspora) in rainbow trout (Oncorhynchus mykiss) gills by in situ hybridization. Veterinary Pathology 36:610-612.

Sánchez, J. G., Speare, D. J. and Markham, R. J. F. 2000. Normal and aberrant tissue distribution of Loma salmonae (Microspora) within rainbow trout, Oncorhynchus mykiss (Walbaum), following experimental infection at water temperatures within and outside of the xenoma-expression temperature boundaries. Journal of Fish Diseases 23:235-242.

Sánchez, J. G., Speare, D. J., Markham, R. J. F., Wright, G. M. and Kibenge, F. S. B. 2001c. Localisation of the initial developmental stages of Loma salmonae in rainbow trout (Oncorhynchus mykiss). Veterinary Pathology 38(5):540-546.

Sandeep, B. V. and Kalvati, C. 1985. A new microsporidian, Loma trichiuri n. sp., from the gill of a marine fish, Trichiurus savala Cuv. (Trichiuridae). Indian Journal of Parasitology 9(2):257-259.

Shaw, R. W. and Kent, M. L. 1999. Fish microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 418-446.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 1998. Modes of transmission of Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 33(2):151-156.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000c. Viability of Loma salmonae (Microsporidia) under laboratory conditions. Parasitology Research 86:978-981.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2001. Phagocytosis of Loma salmonae (Microsporidia) spores in Atlantic salmon (Salmo salar), a resistant host, and chinook salmon (Oncorhynchus tshowytscha), a susceptible host. Fish \& Shellfish Immunology 11:91-100.

Shaw, R. W., Kent, M. L., Docker, M. F., Brown, A. M. V., Devlin, R. H. and Adamson M. L. 1997. A new species of Loma (Microsporea) in shiner perch (Cymatogaster aggregata). Journal of Parasitology 83(2):296-301.

Shimodaira, H. 2000. Another calculation of the p-value for the problem of regions using the scaled bootstrap resamplings. Technical Report No. 2000-35. Stanford University.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology. 51(3):492-508.

Shimodaira, H. and Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17(12):1246-1247.

Sites, J. W. and Crandall, K. A. 1997. Testing species boundaries in biodiversity studies. Conservation Biology 11(6):1289-1297.

Snowden, K. F. and Shadduck, J. A. 1999. Microsporidia in higher vertebrates. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 393-417.

Speare, D. J., Beaman, H. J., Jones, S. R. M., Markham, R. J. F. and Arsenault, G. J. 1998b. Induced resistance in rainbow trout, Oncorhynchus mykiss (Walbaum), to gill disease associated with the microsporidian gill parasite Loma salmonae. Journal of Fish Diseases 21(2):93-100.

Speare, D. J., Brackett, J. and Ferguson, H. W. 1989. Sequential pathology of the gills of coho salmon with a combined diatom and microsporidian gill infection. Canadian Veterinary Journal 30:571-575.

Speare, D. J., Daley, J., Markham, R. J. F., Sheppard, J., Beaman, H. J. and Sánchez, G. J. 1998c. Loma salmonaeassociated growth rate suppression in rainbow trout, Oncorhynchus mykiss (Walbaum), occurs during early onset xenoma dissolution as determined by in situ hybridization and immunohistochemistry. Journal of Fish Diseases 21(5):345-354.

Sprague, V. 1977a. Classification and phylogeny of the microsporidia. In Bulla, L. A. and Cheng, T. C. (eds) Comparative Pathology. 2. Systematics of the Microsporidia. New York and London, Plenum Press 333 pp .

Sprague, V. 1977b. Annotated list of species of microsporidia. In Bulla, L. A. and Cheng, T. C. (eds) Comparative Pathology. 2. Systematics of the Microsporidia. New York and London, Plenum Press 333 pp.

Sprague, V. 1977c. The zoological distribution of microsporidia. In Bulla, L. A. and Cheng, T. C. (eds) Comparative Pathology. 2. Systematics of the Microsporidia. New York and London, Plenum Press 333 pp.

Sprague, V. and Becnel, J. J. 1998. Note on the name-author-date combination for the taxon microsporidies Balbiani, 1882, when ranked as a phylum. Journal of Invertebrate Pathology 71:91-94.

Sprague, V., Becnel, J. J. and Hazard, E. I. 1992. Taxonomy of phylum Microspora. Critical Reviews in Microbiology 18(5/6):285-395.

Stiller, J. W. and Hall, B. D. 1997. The origin of red algae: implications for plastid evolution. Proceedings of the National Academy of Sciences USA. 94:4520-4525.

Sweeney, A. W., Doggett, S. L. and Gullick, G. 1989. Laboratory experiments on infection rates of Amblyospora dyxenoides (Microsporida, Amblyosporidae) in the mosquito Culex annulirostris. Journal of Invertebrate Pathology 53:85-92.

Swofford, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Terry, R. S., Smith, J. E. and Dunn, A. M. 1998. Impact of a novel feminizing microsporidian parasite on its crustacean host. Journal of Eukaryotic Microbiology 45:497-501.

Tuzet, O., Maurand, J., Fize, A., Michel, R. and Fenwich, B. 1971. Proposition d'un nouveau cadre systematique por les genres de Microsporidies. Comptes Rendus de L'Academie des Sciences 272:1268-1271.

Van de Peer, Y., Ben Ali, A. and Meyer, A. 2000. Microsporidia: accumulating molecular evidence that a group of amitochondriate and suspectedly primitive eukaryotes are just curious fungi. Gene 246:1-8.

Van Oppen, M. J. H., Willis, B. L., Van Rheede, T. and Miller, D. J. 2002. Spawning times, reproductive compatibilities and genetic structuring in the Acropora aspera group: evidence for natural hybridization and semi-permeable species boundaries in corals. Molecular Ecology 11:1363-1376.

Van Oppen, M. J. H., Willis, B. L., van Vugt, H. W. J. A. and Miller, D. J. 2000. Examination of species boundaries in the Acropa cervicornis group (Scleractinia, Cnidaria) using nuclear DNA sequence analyses. Molecular Ecology 9:1363-1373.

Vávra, J. and Larsson, J. I. R. 1999. Structure of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 7-84.

Vivarès, C. P. and Méténier, G. 2000. Towards the minimal eukaryotic parasitic genome. Current Opinion in Microbiology 3:463-467.

Vossbrinck, C. R., Andreadis, T. G. and Debrunner-Vossbrinck, B. A. 1998. Verification of intermediate hosts in the life cycles of microsporidia by small subunit rDNA sequencing. Journal of Eukaryotic Microbiology 45:290-292.

Vossbrinck, C. R., Andreadis, T. G., Vávra, J. and Becnel, J. J. 2004. Molecular phylogeny and evolution of mosquito parasitic microsporidia (Microsporidia: Amblyosporidae). Journal of Eukaryotic Microbiology. 51(1):88-95.

Vossbrinck, C. R., Baker, M. D., Didier, E. S., Debrunner-Vossbrinck, B. A. and Shadduck, J. A. 1993. Ribosomal DNA sequences of Encephalitozoon hellem and Encephalitozoon cuniculi: species identification and phylogenetic reconstruction. Journal of Eukaryotic Microbiology 40(3):354-362.

Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. and Woese, C. R. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326(6111):411-414.

Vossbrinck, C. R. and Woese, C. R. 1986. Eukaryotic ribosomes that lack a 5.8 RNA. Nature 320(20):287-288.
Weber, R., Schwartz, D. A. and Deplazes, P. 1999. Laboratory diagnosis of microsporidiosis. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 315-362.

Weidner, E., Findley, A. M., Dolgikh, V. and Sokolova J. 1999. Microsporidian biochemistry and physiology. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 172-195.

Weiser, J. 1977. Contribution to the classification of microsporidia. Vestnik Ceskoslovenske Spolecnosti Zoologicke 41:308-320.

Weiss, L. M. and Vossbrinck, C. R. 1999. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 129-171.

Weiss, L. M., Zhu, X., Cali, A., Tanowitz, H. B. and Wittner, M. 1994. Utility of microsporidian rRNA in diagnosis and phylogeny: a review. Folia Parasitologica 11:81-90.

Williams, B. A. P., Hirt, R. P., Lucocq, J. M. and Embley, M. T. 2002. A mitochondrial remnant in the microsporidian Trachipleistophora hominis. Nature 418:865-869.

Wittner, M. 1999. Historic perspective on the microsporidia: expanding horizons. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 1-6.

Yamamoto, A., Hashimoto, T., Asago, E., Hasegawa, M. and Goto, N. 1997. Phylogenetic position of the mitochondrion-lacking protozoan Trichomonas tenax, based on amino acid sequences of elongation factors $1 \alpha$ and 2. Journal of Molecular Evolution 44:98-105.

Zhu, X., Wittner, M., Tanowitz, H. B., Cali, A. and Weiss, L. M. 1994. Ribosomal RNA sequences of Enterocytozoon bieneusi, Septata intestinalis and Ameson michaelis: phylogenetic construction and structural correspondence. Journal of Eukaryotic Microbiology 41(3):204-209.


Figure 1.1: Diagrammatic representation of a microsporidian spore, showing features commonly used in species diagnosis, for example, thickness of chitinous wall layers (endospore and exospore), features of anchoring disk, number of polar filament coils (often seen in cross section like this), condition of nucleus (uninucleate and isolated in this case), form(s) of polaroplast membranes, presence and size of posterior vacuole. Polyribosomes are not used for diagnosis but are commonly packed in the remaining cytoplasm of the spore.

Figure 1.2: Diagrammatic representation of several microsporidian life cycles (using symbols similar to those of Canning, 1990). Circle/oval shapes = membrane-bound microsporidial cells. Doubled-ovals $=$ spores covered with a spore wall. Dark circles in cells $=$ nuclei. Stipple-filled cells $=$ merogonic stages. Solid-filled cells $=$ sporogonic stages or spores. $\mathbf{A}=$ Some species have a diplokaryotic nucleus throughout the life cycle or for part of the life cycle. However, for the simplicity of this diagram, only uninucleate form is shown. In Thelohania species the diplokaryotic condition may persist until the onset of sporogony. $\mathbf{B}=$ Merogonic divisions (merogony) or the vegetative propagation can be numerous and are often by simple binary fission (mitosis). Sacs (sporophorous vesicle or parasitophorous vacuole) around these stages are present in some species, but are not shown here for the simplicity of this diagram. $\mathbf{C}=\mathrm{A}$ merogonic plasmodium which occurs in some species of microsporidia. $\mathbf{D}$ to $\mathbf{H}=$ At the onset of sporogony, the surface of stages thickens, cytoplasmic contents change, and sometimes a sac is formed if it does not already exist. Various different kinds of division are observed, shown for a case of octosporoblastic sporogony: $\mathbf{D}=$ multiple fissions of a sporogonial plasmodium, here the first 2 steps are plasmotomy producing smaller plasmodia, and the last step is binary fission; $\mathbf{E}=$ rosette like budding of sporogonial plasmodium to simultaneously produce products (called multiple division, not to be confused with multiple fission); $\mathbf{F}=$ binary fissions without the formation of a plasmodium, here a series of 3 binary fissions producing 2 -cell, 4 -cell and 8 cell stages as was seen in the species in Chapter 5 where the dotted line represents a sac that forms around these sporogonic stages omitted here for simplicity of this diagram; $\mathbf{G}$ and $\mathbf{H}=$ merogony/sporogony patterns in species of Loma (Chapter 3), although a rosette (E) was also seen. Note that a sac around the dividing stages can be formed at various points during division of cells. $\mathbf{G}=$ multiple fissions of a cylindrical plasmodium (the first is plasmotomy, the second is binary fission) as was seen in some species of Loma (Chapter 3). $\mathbf{H}=$ binary fission without the production of a plasmodium. A variety of modifications and combinations of these developmental patterns exits. Close relatives may have different developmental sequences. Some species involve use of multiple hosts.


Table 1.1: Species of Loma Morrison \& Sprague, 1981, showing host order, family, species and geographic locality. Dark shading = Species examined in this study. Light shading = DNA sequences compared (source: Genbank or this study). $\quad=L$. morhua was considered synonymous with $L$. branchialis by Canning \& Lom (1986), but Chapter 4 suggests these are separate. If $L$. morhua were a junior synonym of $L$. branchialis, then the host breadth $\left(^{*}\right)$ of $L$. branchialis would include other Atlantic gadids. $*=$ Chapter 2 shows that variant " $L$. salmonae SV" of Sánchez et al. (2001) is not conspecific with L. salmonae, but may be conspecific with L. fontinalis. $\Psi=$ See Appendix 12 for name equivalences for these five species examined in Chapters $3 \& 4 . \phi=$ also one Blenniidae (Lyphophyrs pholis), and Zosterisessor ophiocephalus. $\varphi=$ formerly Trichiurus savala. (N. Amer. = North America).

| Loma species | Host order | Host family | Host species | Locality |
| :---: | :---: | :---: | :---: | :---: |
| L. acerinae (Jirovec, 1930) Lom \& Pekkarinen, 1999 | Perciformes | Percidae | Gymnocephalus cernuus | Finland |
| L. boopsi Faye, Toguebaye \& Bouix, 1995 | Perciformes | Sparidae | Boops boopsi | Africa |
| L. branchialis (Nemeezek, 1911 ) Morison \& Sprague, 1981 (type) $\downarrow$ | Gadiformes | Gadidae sus | Melanogrammus aeglefinus* | Atlantic |
| L. camerounensis Fomena, Coste \& Bouix, 1992 | Perciformes | Cichlidae | Oreochromis niloticus | Africa |
| L. dimorpha Loubès, Maurand, Gasc, De Buron \& Barral, 1984 | Perciformes | Gobiidae ${ }^{\text {¢ }}$ | Gobius niger + others | France |
| L. diplodae Bekhti \& Bouix, 1985 | Perciformes | Sparidae | Diplodus sargus | France |
| L.embiotocia Shaw, Kent, Docker, Brown, Devlin \& Adammon, 199 | Perciformes | Embiotocidae | Cymatogaster aggregata | Pacific N - Amer |
| L. fontinalis Morrison \& Sprague, 1983 * | Salmoniformes | Salmonidae | Salvelinus fontinalis | Eastern N. Amer. |
| L. morhua Morrison \& Sprague, 1981 | Gadiformes | Gadidae - | Gadus morhua | Atlantic N. Amer. |
| L. mugili Ovcharenko, Sarabeev, Wita \& Czaplińska, 2000 | Perciformes | Mugilidae | Mugil soiuy | Russia |
| L. myrophis Azevedo \& Matos, 2002 | Anguilliformes | Ophichthidae | Myrophis platyrhynchus | South America |
| L. salmonae (Putz, Hoffman \& Dunbar, 1965) Morrison \& Sprague, 1981 'L. Salmonae SV' of Sánchezet al: 2001 * | Salmoniformes Salmoniformes | Salmonidae Salmonidae | Oncorhynchuss spp. T others Salvelinus fontinalis | Westem N. Amer. Eastern N. Amer. |
| L. trichiuri Sandeep \& Kalvati, 1985 | Perciformes | Trichiuridae | Lepturacanthus savala ${ }^{\text {¢ }}$ | India |
| Loma sp of Adlard, unpublished $=\\|=1 \mathrm{mam}$ | Perciformes | Sparidae te | Acanthopagrus latus | Australia $=0$ |
| Loma sp. of Bekhti, 1984 | Perciformes | Cichlidae | Tilapia melanopleura | France |
| Lomă sp:ofKentet al, $1998^{\text {T}}$ (described in Chapter 3) <br> Lomasp of K ent et al, 199 T $^{\Psi}$ (described in Chapter 3 ) <br> Loma sp of Kent et al., $1998{ }^{4}$ (described in Chapter33) <br> Loma Sp. of Kent et al., $1998{ }^{4}$ (described in Chapter 3) <br> Loma speofKentetal, $1998^{*}$ (described in Chapter 3) | Gadiformes <br> Gadiformes <br> Gadiformes <br> Scorpaeniformes <br> Scorpaeniformes | Gadidae <br> Gadidae: <br> Gadidae <br> Hexagrammidaè <br> Anoplopomatidae | Gadus macrocephalus <br> Theragra chalcogramma <br> Microgadus proximus <br> Ophiodonelongatus <br> Anoplopoma fimbria | Pacific N. Amer. <br> Pacificic: Amer. <br> Pacific N: Amer: <br> Pacific N Amer <br> Pacific N A Amer. |
| Loma sp. of Nilsen, 2000 | Gadiformes | Lotidae | Enchelyopus cimbrius | Atlantic |
| Loma sp. of Narasimhamurti et al., 1990 | Perciformes | Belontiidae | Trichogaster trichopterus. | India |
| Loma sp. of Narasimhamurti et al., 1989 | Perciformes | Sciaenidae | unidentified sciaenid | India |

# Chapter 2: DNA distinguishes populations of Loma salmonae (Microsporidia) in salmon and trout (Oncorhynchus spp.) and identifies a cryptic species in brook trout (Salvelinus fontinalis) 

## INTRODUCTION

Loma salmonae (formerly Pleistophora salmonae Putz, Hoffman \& Dunbar, 1965) is a pathogen in the gills and other tissues of farmed and wild salmonids, causing morbidity and mortality in farmed fish (Magor, 1987; Speare et al., 1989; Kent et al., 1990; Markey et al., 1994; Bruno et al., 1995; Bader et al., 1998; Kent et al., 1998; Kent, 2000). Several studies suggest that wild populations of L. salmonae have been prevalent and widespread historically. They have occurred in $75 \%$ of wild yearling rainbow trout in California (Wales \& Wolf, 1955), and regularly as an epizootic in northwestern and eastern North America, Japan and Europe (Hauck, 1984; Bekhti \& Bouix, 1985; Canning \& Lom, 1986; Kent et al., 1989; Bruno et al., 1995; Shaw \& Kent, 1999). Extremely low-level infection with L. salmonae can now be detected using a PCR-probe designed by Docker et al. (1997a). This probe, capable of detecting 0.01 spores per $50 \mu \mathrm{l}$ PCR reaction (Docker et al., 1997a), was recently shown to detect $L$. salmonae infection in wild salmon smolts at a level too low to be detected by light microscopy (Shaw et al., 2000c). This study employs Docker et al.'s (1997a) probe to detect infections in wild, farmed and laboratory salmonids, and then uses DNA sequencing to examine differences among isolates of $L$. salmonae.

Like other microsporidians, L. salmonae is a very small ( $\sim 2-6 \mu \mathrm{~m}$ ), spore-forming, obligate, intracellular parasite, and like many species from fishes, L. salmonae forms large ( $\sim$ 1 mm ) cyst-like nodules (xenomas), which release infective spores to the environment. Salmon ingest spores, which then extrude tiny tubes (polar filaments) through which the nucleus and cytoplasm (sporoplasm) travel into host gut epithelial cells. Eventually these early stages migrate to the lamina propria and enter the blood, perhaps transported by macrophages (Shaw et al., 1998) to the heart or other tissues where they undergo a series of division cycles (merogony)
(Sánchez et al., 2000; Sánchez et al., 2001c). Gills are infected by transport cells containing merogonic stages, which mature into spores in xenomas, and then spores can be released into the environment or autoinfect the fish (for details, see Rodriguez-Tovar et al., 2003). The parasite can potentially lie dormant and undetected in the heart or other body tissues. One study suggests wild salmon could carry light infections (as smolts) that are not detected microscopically (Shaw et al., 2000b), and other studies suggest salmon may have cryptic infections (without formation of xenomas) under some temperature conditions in which salmon can develop resistance to further infection (Speare et al., 1998b; Beaman et al., 1999a, b; Kent et al., 1999; Shaw et al., 2001). Such infections may endure for an unknown amount of time and later pass to other fish in fresh or seawater (Kent et al., 1995; Ramsay et al., 2002). Loma salmonae might be haploid (perhaps clonal) or diploid, or could have an unusual ploidy cycle like that seen in other distantly related microsporidians (Hazard \& Brookbank, 1984).

Sympatric species that are morphologically similar to L. salmonae exist in six nonsalmonid fishes in British Columbia (Shaw et al., 1997; Kent et al., 1998), and one salmonid in eastern North America (Morrison \& Sprague, 1983). Prior to the present study and those in the following chapters, few or no characters were available to distinguish these species, besides host, or in some cases transmissibility to alternate hosts. Thus, the present study examines DNA variation among morphologically homogeneous isolates of $L$. salmonae - a species with a wide geographic range and multiple hosts - and compares molecular data across the geographic- and host-range of $L$. salmonae to that of a questionable variant ("L. salmonae SV") in brook trout. Data will be interpreted by considering whether genetic discontinuities suggesting populationlevel or species-level differences exist across isolates of L. salmonae. The basis for this approach is the working species definition most appropriate for the data gathered here. The general working species definition to be used here (based on arguments in Wheeler \& Meier, 2000) has two parts. The first criterion for distinguishing one species from another when in sympatry is that each must possess separate discrete characters or overlapping characters with a statistically separate mean. The second criterion is that the suite of characters that provides evidence for species must agree across the sampled populations. To fix the interpretation around a valid member of the species, data will be compared to isolates from the type host (Oncorhynchus mykiss) and location (California) for L. salmonae.

Knowledge of host and geographic range of L. salmonae was important in selecting isolates for this study. Several transmission studies have suggested L. salmonae is distinct from Loma species found in non-salmonid hosts and is exclusively a parasite of salmonids. For
example, Shaw and Kent (1999) and Shaw et al. (2000b) demonstrated that L. salmonae is able to infect all seven Oncorhynchus species in the laboratory, but not Atlantic salmon Salmo salar, Arctic char Salvelinus alpinus, herring Clupea pallasi, prickly sculpin Cottus asper, shiner perch Cymatogaster aggregata, sticklebacks Gasterosteus aculeatus, goldfish Carassius auratus, and guppies Poecilia spp. However, Shaw et al. (2000b) were successful in infecting two other nonOncorhynchus spp. salmonids (brown trout Salmo trutta, and brook trout), at least under experimental conditions. Others (Speare et al., 1998a; Sánchez et al., 2001a; b) found brook trout to be resistant or sometimes susceptible to experimental infection. These authors hypothesized that resistant brook trout may have developed partial resistance to infection during an undetected prior exposure to L. salmonae. However, these infections in brook trout were not easily transmitted to other hosts of L. salmonae (Sánchez et al., 2001a; b), thus, the authors suggested there is a second strain of L. salmonae, "L. salmonae SV" which prefers brook trout and was present in the chinook salmon Oncorhynchus tshawytscha that were used to infect these brook trout (Sánchez et al., 2001a; b; Speare \& Daley, 2003). An alternative explanation of the results is that brook trout are not actually susceptible to $L$. salmonae, but may carry cryptic infections with a second species of Loma, L. fontinalis, which was originally described from brook trout at a nearby location (Morrison \& Sprague, 1983; and see Brown \& Kent, 2002). This alternative requires that L. fontinalis (morphologically indistinguishable from $L$. salmonae by light microscopy) occurred in the supposedly infection-free hatcheries from which experimental brook trout were supplied, perhaps causing only low-level or undetected pathogenesis (for example, causing no xenoma formation). Shaw et al.'s (2000c) PCR-based study showing that Loma species are able to exist at sub-clinical levels lends support to this theory. The present study attempts to characterize molecular markers that may help in future experimental transmission studies of this kind.

The question of whether there exist alternate strains of L. salmonae with different host preferences is even more intriguing in light of studies presenting differences in prevalence in different hosts in the wild (Shaw et al., 2000b) and different susceptibilities in the lab (Shaw et al., 2000a; b; Ramsay et al., 2002). These observations could be explained by differences in host response alone or differences in the parasite's host preference among several strains maintained in individuals or groups of fish in the laboratory or the wild, or a combination of both host- and parasite-factors. Wild $L$. salmonae could be made up of several related strains, each specialized to infect different salmon species, or characterized by different levels of virulence in different hosts. Teasing apart parasite-driven prevalence differences (e.g.
specificity for certain hosts or competing strains with different characteristic virulence) versus host-driven differences (i.e. host susceptibility due to prior exposure, life-history, etc.) would require experiments beyond the scope of the current study; however, a first step in such an investigation would be to look for genetically distinct strains of the parasite in these different hosts or host populations.

In summary, the first goal of this study was to examine the prevalence of L. salmonae in populations of five species of Pacific salmon: chinook $O$. tshawytscha, sockeye $O$. nerka, coho O. kisutch, pink O. gorbuscha, and chum $O$. keta, and in doing so, examine the utility of Docker et al.'s (1997a) sensitive and specific PCR probe. The second goal of this study was to look for genetic markers that distinguish populations of L. salmonae in different geographic locations in the wild, the laboratory, farms, and in different host species populations of L. salmonae, including the type locality and type host (Californian rainbow trout, O. mykiss) by sequencing a highly variable portion of the parasite's ribosomal DNA (rDNA) region from these isolates. The third goal of this study was to characterize sequence variation among isolates, including the type locality and type host for L. salmonae, for a second, independent nuclear gene, elongation factor-1alpha (EF-1 $\alpha$ ), expected to vary sufficiently to serve as a reasonably variable, independent intra- and inter-specific marker for Loma species (Cho et al., 1995; Moreira et al., 1999; and see Chapter 4). The fourth goal of this study was to compare sequence data for these two independent genes among isolates of $L$. salmonae, including the type host and type location, to verify that, as expected given morphological similarities and host and transmission data, $L$. salmonae fits the working definition for species (above). Given this definition, the fifth goal of this study was to examine the claim by Sánchez et al. (2001a; b) and Speare \& Daley (2003) that the species or variant from brook trout that was given the name " $L$. salmonae SV " is a strain of L. salmonae rather than a distinctly separate species. The last goal of this study was to characterize of the specificity of Docker et al.'s (1997a) PCR-probe as a "species-specific marker" for L. salmonae from many locations by gathering sequence data across the PCRprobe's priming sites from a wide range of $L$. salmonae isolates. A continuation of the latter investigation comparing this probe's priming sites for other species of Loma will be presented in Chapter 4.

## MATERIALS AND METHODS

## Specimen collection

Salmon species (Oncorhynchus spp.) from wild populations were caught by hook and line or seine from rivers, lakes and bays on Vancouver Island or surrounding areas in British Columbia, Canada, or were caught by trawling aboard the W. E. Ricker vessel from the east and west coasts of Vancouver Island during Fisheries and Oceans Canada research surveys. All wild salmon used in the PCR-test for Loma salmonae were collected during a single collection trip through the Georgia and Johnston Straits in Sep. 1997, whereas wild salmon used in DNA sequencing came from various localities between Nov. 1996 and Sep. 1997 (see Table 2.2). Fishes were killed with a blow to the head or an overdose of tricaine methanesulfonate (MS222) prior to examination. Fresh gills were examined at 400X and 1000X magnification for the presence of $L$. salmonae before placement in $95 \%$ molecular grade ethanol. Fishes used in the PCR-test were not examined for $L$. salmonae prior to fixation of gills in ethanol.

Salmon and rainbow trout (Oncorhynchus spp.) from farms and hatcheries (Table 2.2) were generally examined on-site by the naked eye or sometimes under a dissection microscope (100X - 250X) before they were kindly donated to the present authors. Gills were fixed in ethanol or frozen.

Salmon from the laboratory (Table 2.2) were raised at the Pacific Biological Station (PBS), Nanaimo, British Columbia, Canada and maintained as described in Kent et al. (1995), Shaw et al. (1998) and Shaw et al. (2000c). These fish were maintained as hosts for an ongoing population of $L$. salmonae to be used in a variety of experiments. Loma salmonae was initially introduced to this laboratory stock from a nearby commercial farm in British Columbia, which presumably contracted $L$. salmonae in the sea net pens. Gills from laboratory-reared salmon were examined microscopically for the presence of L. salmonae, and either fixed in ethanol or placed in Earl's Buffered Saline Solution (EBSS) and stored at $4^{\circ} \mathrm{C}$ to be used in spore isolation.

Dr. J. Genaro Sánchez -Martinez from the Atlantic Veterinary College, Prince Edward Island (P.E.I.), Canada provided gill material from laboratory-held brook trout (originally from a nearby hatchery). The gills showed xenomas after a 3rd trial following experimental transmission with spores originally from chinook salmon from the Pacific Biological Station, in

British Columbia, Canada (refer to Speare et al., 1998a; Shaw \& Kent, 1999; Sánchez et al., 2001a; b; Speare \& Daley, 2003).

## Spore isolation

Spores were isolated from gills of laboratory-raised salmon in batches of 6 to 17 fish per isolation, using Percoll (SIGMA Ronkonkoma, New York) gradients as described in Shaw et al. (1998). Gill tissue was allowed to autolyze for several days in cold EBSS before being ground using a Polytron (Luzer, Switzerland) tissue homogenizer. The resulting slurry was mixed with distilled, deionized $\mathrm{H}_{2} \mathrm{O}\left(\mathrm{ddH}_{2} \mathrm{O}\right)$, centrifuged at 2000 xg for 10 min , and the resulting spore pellet was washed, gently resuspended in $\mathrm{ddH}_{2} \mathrm{O}$ and centrifuged at 500 xg for 45 sec . The new pellet was resuspended, laid onto a $35 \%$ Percoll gradient and centrifuged to pellet spores. This was repeated twice for $50 \%$ and $65 \%$ Percoll gradients. Spores were stored at $-70^{\circ} \mathrm{C}$.

## DNA isolation

DNA isolation from purified spore concentrates required bead beating to first break open spores. Bead beating followed the procedure of Docker et al. (1997a), involving shaking isolated spores with 0.5 mm silica beads in a Mini-Beadbeater (Biospec Products, Bartlesville, OK, USA) with TE buffer ( 10 mM tris, 1 mM EDTA) for 1 to 3 min . Prior to DNA isolation from ethanol-fixed gills, ethanol was removed by soaking for 15 minutes in lysis buffer ( 10 mM Tris, 1 mM EDTA, $10 \mathrm{mM} \mathrm{NaCl}, 1 \%$ SDS). For the PCR-test, individual salmon from each species were grouped together in fives. Sixty to 65 mg of gill tissue ( 4 to 8 filaments, depending on the size of fish) was used for each fish. For all other ethanol-fixed samples, approximately 60 to 65 mg of ethanol-fixed gill tissue was taken from a single fish. Bead-beaten spores or ethanol-fixed gills were then digested in lysis buffer with $0.5 \mathrm{mg} / \mathrm{ml}$ proteinase K for $4-6$ hours at $37^{\circ} \mathrm{C}$ in a rotating incubator. DNA was phenol chloroform extracted (one phenol step, two phenol: chloroform: isoamyl alcohol 25:24:1 steps, and one chloroform: isoamyl alcohol 24:1 step), precipitated in cold $95 \%$ ethanol, washed twice with $70 \%$ ethanol, vacuum dried, resuspended in $40 \mu \mathrm{l}$ distilled water and stored for use at $-20^{\circ} \mathrm{C}$.

## Polymerase chain reaction test (PCR-test)

The polymerase chain reaction test (PCR-test) for L. salmonae was performed following the methods detailed in the sensitive and specific PCR assay of Docker et al. (1997a). Briefly, 25 pmol of each diagnostic primer LS-1 and LS-2 (in the ITS and LSU rDNA regions, respectively, see Docker et al. 1997a) were added to each $50 \mu \mathrm{l}$ reaction, along with $5 \mu \mathrm{l}$ 10X standard PCR buffer (Boehringer Mannheim, Germany), $1.5 \mathrm{mM} \mathrm{MgCl} 2,0.2 \mathrm{mM}$ dNTPs, 1.25
units Taq DNA polymerase, $0.6 \mu \mathrm{~g}$ template DNA, and molecular biology grade $\mathrm{H}_{2} \mathrm{O}$. Thermal cycling was performed in a Perkin Elmer Cetus DNA Thermal Cycler $480\left(94^{\circ} \mathrm{C}\right.$ for $3 \mathrm{~min}, 35$ cycles of $94^{\circ} \mathrm{C}$ for $1 \mathrm{~min}, 53^{\circ} \mathrm{C}$ for $1 \mathrm{~min}, 72^{\circ} \mathrm{C}$ for 1 min , followed by $72^{\circ} \mathrm{C}$ for 5 min ). Positive control L. salmonae DNA from laboratory-reared chinook and a negative control (water only) replicate were always run alongside test PCR reactions. The 272 base pair product was visualized on standard $1.5 \%$ agarose gels with ethidium bromide stain. The PCR test was repeated in case of false results. To test for false negatives due to inhibitors in DNA extractions, the DNA pools that resulted in negative tests for $L$. salmonae were run in a repeat PCR in which $0.6 \mu \mathrm{~g}$ of positive control L. salmonae DNA was added to the starting reaction in addition to the sample. Where there was no amplification in these positive "spiked" runs, DNA inhibitors could be responsible for negative results, so for these samples "spiked" PCR was repeated with 3 x more Taq polymerase, and if it still failed to amplify, inhibitors were cleaned from the original template DNA by repeating phenol:chloroform extraction, ethanol precipitation, and washes, or with QIAquick PCR Purification Kit (Qiagen, Santa Clarita, CA). The range in prevalence was estimated by multiplying the minimum number of fish that could be infected per positive pool (e.g. one out of five), or the maximum number (e.g. five out of five) by the number of positive pools and dividing this number by the total number of fish collected.

## PCR for sequencing and cloning

PCR for sequencing of ribosomal DNA (rDNA) and elongation factor- $1 \alpha$ (EF-1 $\alpha$ ) genes was performed in a thermal cycler (above) in $25 \mu \mathrm{l}$ reactions (generally 4 replicates pooled together) with standard PCR buffer $2.5 \mathrm{mM} \mathrm{MgCl}_{2}, 0.2 \mathrm{mM} \mathrm{dNTP}, 15 \mathrm{pmol}$ of each primer, and 1-3 units of Taq DNA polymerase (Invitrogen Corp., Carlsbad, CA). Conditions were $95^{\circ} \mathrm{C}$ for $1 \mathrm{~min}, 35$ cycles of $95^{\circ} \mathrm{C}$ for $50 \mathrm{sec}, 54^{\circ} \mathrm{C}$ for 30 sec (or as low as $50^{\circ} \mathrm{C}$ for difficult amplifications), $72^{\circ} \mathrm{C}$ for 90 sec , and a final extension of $72^{\circ} \mathrm{C}$ for 5 min . Primers for sequencing ribosomal DNA (rDNA) and amplifying various regions in the small subunit (SSU), internal transcribed spacer (ITS) and large subunit (LSU) were: forward M5P - CAC CAG GTT GAT TCT GCC pos. 1-18 at 5' end of the SSU (18eMIC in Docker et al., 1997b); Seqlf - CGT TGT AGT TCT AGC AGT pos. 719-736 in the SSU (provided by M. F. Docker); L7f - ATT AGT GAG ACC TCR GCC pos. 1001-1019 in the SSU (Loma f in Shaw et al. 1997); reverse SeqR - AAC AGG GAC KYA TTC ATC pos. 1218-1235 in the SSU (this study); 580R GGT CCG TGT TTC AAG ACG G pos. 1847-1865 in the LSU (Vossbrinck et al. 1987). Elongation factor-1 $\alpha(\mathrm{EF}-1 \alpha)$ primers were designed to match conserved regions of the gene using an
alignment of a wide range of eukaryote sequences including the microsporidian Glugea plecoglossi and several fungi. These primers amplified about 1120 base pairs from 43 amino acids downstream of the start codon. Primers and positions relative to G. plecoglossi were: forward EFZ - TTG CTT CAT TGG NCA CGT MGA pos. 32; EFX - AGA AAG AGG TAG AGG TAC TT pos. 143; EFV - GTA CAT ATC GTG GTA TTA pos. 198; reverse EFD - TGC ACC TGT ACT ACY CTN CCN GT pos. 806; EFW - AAG TCA CAT TTT CAC CTT T pos. 1203; EFY - CAA TTG CAC CGA TTC CGA C pos.1314. PCR products were visualized and excised from 1.5\% agarose TBE or TAE gels, and then freeze-thaw extracted or $\beta$-agarase digested to remove agarose. They were then sequenced directly or cloned.

## Cloning

PCR products were isolated in $0.8 \%$ agarose and cleaned for ligation using Ultraclean 15 MOBIO DNA Purification Kit (BIO/CAN Scientific Inc. Mississauga, ON) and cloned using the TOPO TA Cloning PCR Version 2.1 (Invitrogen Corp., Carlsbad, CA) using $1 / 2$ volume. Clones were screened for presence of the insert in $10 \mu \mathrm{l}$ PCR reactions using Taq DNA Polymerase (Invitrogen Corp., Carlsbad, CA) with standard reagents and screening primers M13-20 and M13 Rev (conditions: $94^{\circ} \mathrm{C}$ for 2 min , 34 cycles of $92{ }^{\circ} \mathrm{C}$ for $45 \mathrm{sec}, 55^{\circ} \mathrm{C}$ for 45 $\sec 72^{\circ} \mathrm{C}$ for 1 min 30 sec , followed by $72^{\circ} \mathrm{C}$ for 5 min ). Positive clones from master plates were grown in 3 ml of standard XY culture with 50 mM ampicillin by shaking at 220 rpm at 37 ${ }^{\circ} \mathrm{C}$ overnight. Plasmids were isolated for sequencing using the Rapid Plasmid Miniprep System (Gibco BRL, Gaithersburg, MD) following directions of the manufacturer.

## DNA sequencing

Sequencing was performed on the ABI PRISM 377 DNA automated sequencer using
BigDye Terminator Version 3.1 fluorescent dye-labelled terminators with forward and reverse primers and PCR conditions as recommended for the Taq terminators. Wherever possible, PCR products were sequenced in both directions. Multiple PCR products and multiple clones were sequenced to check for Taq or sequencer errors.

## Sequence analysis

Sequences were easily aligned by eye using the program ESEE Version 3.2s (Eyeball SEquence Editor, by Eric Cabot, 1998). Polymorphisms (two alleles of a gene) in the population produced polymorphic sites in directly sequenced PCR products (i.e. in which two or more nucleotide signals appeared at a site in a sequence). Because of uncertainty as to whether these double signals were erroneous (from poor quality or contaminated template, or from
sequencing artefacts) or reflect valid polymorphism in the data (such as two different copies of a multiple-copy gene or different alleles of a single-copy gene both amplified in PCR and sequencing reactions), the original sequence output data was carefully re-examined for evidence of sequencer software error or background signal. In cases where there was still uncertainty, sequencing was repeated. Valid polymorphism of this kind was indicated using the standard degenerate nucleotide code. Where a species had more than one nucleotide at a site (a polymorphism) and another species has only one nucleotide at this site, the difference was classified into one of two types. Where one nucleotide in the polymorphic species matched that of the second species, it was referred to as a polymorphic difference or 'half substitution', whereas if both nucleotides in the polymorphic species do not match the other species, this was counted as a full substitution. Percent sequence difference was calculated by adding the number of substitutions, the number of indels (regardless of size) and half of the number of polymorphic differences (i.e. giving these a half substitution value), divided by the total number of alignment positions for which both sequences are known. Where multiple replicates of PCR , sequencing, or different clones produced different sequences for one isolate all the differences from each isolate were added, and this number was divided by the total number of sequenced nucleotides for all sequences in that isolate.

## RESULTS

## PCR-test for Loma salmonae in wild salmon

The PCR-test for $L$. salmonae showed that all five species of wild-caught Pacific salmon (Fig. 2.1) from Georgia and Johnston Straits could harbour this parasite. Band brightness, reflecting quantity of amplified PCR product, differed between samples (pools) in different lanes (Fig. 2.1). These relative differences in quantity of amplified PCR product were consistent over repeated PCR runs. Figure 2.1 also shows little non-specific priming.

Number of fish sampled, PCR pools tested, L. salmonae-positive pools, and estimated prevalence for each salmon species are shown in Table 2.1. Salmon species differed in frequency of $L$. salmonae-positive PCR-test results (Fig. 1 and Table 2.1). Chinook salmon $O$. tshawytscha had the highest fraction of L. salmonae-positive pools, followed in descending order, by coho salmon $O$. kisutch, sockeye salmon $O$. nerka, pink salmon $O$. gorbuscha and chum salmon $O$. keta. Range in prevalence, calculated as described (see Materials and Methods) overlapped among all salmon species.

Many samples produced L. salmonae-negative PCR-test results; however, these samples were tested for inhibitors as described in the Materials and Methods, by spiking with $L$. salmonae DNA and repeating PCR amplification. Samples that did not amplify after being spiked were deemed to have inhibitors affecting the template DNA (Fig. 2.2). After inhibitorremoval, all spiked samples could be amplified. All corresponding un-spiked inhibitor-cleaned template DNA remained $L$. salmonae-negative. This suggested false negatives did not occur in this study. DNA sequencing was attempted for all $L$. salmonae-positive pools; however, many positive pools could not be PCR-amplified with sequencing primers. Several positive pools could be amplified but not sequenced. In some cases products could be sequenced but sequences bore no resemblance to Loma spp. or other microsporidian sequences in Genbank, suggesting non-specific amplification.

## Intraspecific variation in the rDNA genes

Intraspecific variation in rDNA sequences was examined among 26 isolates of $L$. salmonae including seven isolates from the laboratory, 10 isolates from farms, and nine isolates from wild populations. The number of nucleotide positions sequenced for each isolate and the number of rDNA sites where intraspecific variation was observed are shown in Table 2.2. This intraspecific variation is symbolized by a number for each nucleotide difference compared to a
reference $L$. salmonae sequence shown in Fig. 2.3. For simplicity, the most common sequence found in this study (shared by 18 of 26 isolates, all but one from British Columbia) was chosen as the reference sequence in this figure. I will hereafter refer to this more commonly observed rDNA sequence in Fig. 2.3 as the "BC-form".

Figure 2.3 shows differences between the BC-form and a published L. salmonae sequence (Genbank U78736) that was obtained from fishes at the same locality, reared in the same building (Pacific Biological Station, Nanaimo, Canada). The rDNA region differed between the BC-form and published L. salmonae by $0.22 \%$ in 908 bp of SSU, $0.77 \%$ in 452 bp of LSU and did not differ in the 37 bp ITS region. Figure 2.3 shows how variation was distributed across the ribosomal region for L. salmonae isolates. Among 1832 alignment positions ( 1829 bp ), including 1343 bp SSU, 37 bp ITS, and 452 bp LSU, most of the differences occurred in the LSU, some in the SSU, and none in the ITS region.

Unfortunately, much of the rDNA variation among isolates occurred in regions not sequenced for more than half of the isolates, replicate clones or PCR runs, thereby making phylogenetic analysis impractical or uninformative for this data. Instead, I will describe the diversity and distribution of sequence differences, as well as reporting \% differences from the BC-form, calculated as described (Materials and Methods).

The eight isolates of $L$. salmonae that differed in rDNA sequence from the BC-form (Table 2.2) were from the laboratory (BA4 and BA9) and farms in British Columbia (I-27), California (CAL and L6), Colorado (Coll and Col2) and Chile (Lsc2). The Colorado isolate (Coll) had the most differences (seven differences) while the Indian Bay British Columbia isolate (I-27) was nearly identical (one difference). Table 2.2 also shows that isolates from farms were more variable than wild and British Columbia isolates. Some sequence differences were shared among isolates (shown in bold in Table 2.2 and Fig. 2.3), while others were not. Wild-caught isolates were all of the BC-form (including four isolates corresponding to pools from the PCR-test above), in four species of salmon: chinook (S52), coho (S21), pink (S41) and sockeye (S71).

Differences within isolates, generally observed as differences between two or more clones or PCR-runs, were incorporated in the \% difference calculations shown in Table 2.3 (see explanation of calculations in Materials and Methods). Table 2.3 shows that in the SSU gene, \% difference from the BC-form $L$. salmonae isolate was extremely low, presumably comparable to that expected from Taq errors (i.e. 1 in 1000). Percent difference was higher in the LSU than the SSU gene for all isolates. Laboratory isolates, BA4 and BA9, were similar to the BC-form,
with a small difference between them in the LSU. Californian isolates (L6 and CAL) were less similar, and Colorado isolates (Coll and Col2) were most diverged from BC-form isolates in the LSU rDNA.

## Intraspecific variation in the EF-1 $\alpha$ gene

The EF-1 $\alpha$ gene was sequenced for 11 isolates of L. salmonae including four isolates from the laboratory, four isolates from farms, and three isolates from wild populations. The number of nucleotide positions sequenced for each isolate and sites with intraspecific variation are shown in Table 2.2, as explained for rDNA (above). For a reference sequence, I chose the most common sequence, which was observed from six BC isolates, one Californian and one Chilean isolate. This isolate, hereafter named "BC-EF-1 $\alpha$ " is shown in an alignment in Fig. 2.4, along with the hypothesized amino acid sequence.

Figure 2.4 shows differences between the BC-EF-1 $\alpha$ sequence and a published microsporidian EF-1 $\alpha$.sequence, Glugea plecoglossi Genbank D321239. Between these species there was $14.3 \%$ amino acid difference (over 377 amino acid alignment positions), plus one two-amino acid indel. Nucleotide difference was 28.1 \% plus one indel (over 1132 alignment positions).

Only three out of 11 isolates of $L$. salmonae differed in sequence from the BC-EF-1 $\alpha$ (Table 2.2). Two of these isolates were from the laboratory (BA4 and BA9) and one was from a wild-caught sockeye (SP-24) in British Columbia. Out of six substitutional differences from the $B C-E F-1 \alpha$ sequence, two were synonymous and four were non-synonymous. Overall the intraspecific difference from the reference sequence was low, at $0.03 \%$ per 27374 sequenced nucleotides.

## Variation between $L$. salmonae isolates and "L. salmonae SV" from brook trout

Loma salmonae and the brook trout derived "L. salmonae SV" variant possessed 19 differences across 1830 alignment positions ( $1.4 \%$ ) in the SSU, 10.5 differences over 452 alignment positions ( $2.3 \%$ ) in the LSU, and $2.3 \%$ difference over 1030 alignment positions of the partial EF-1 $\alpha$ (see Chapter 4 for further analyses).

## DISCUSSION

## Intraspecific rDNA sequence variation in $L$. salmonae

Intraspecific sequence variation was expected to be relatively low within a single species that is sexual (or clonally) reproducing. Sequence variation should not show distinct discontinuities among well-mixing (sympatric) conspecifics according to the working species definition mentioned in the introduction. This study found low sequence difference at two independent loci (rDNA and EF-1 $\alpha$ ) across isolates of $L$. salmonae from southern British Columbia, and several others from elsewhere (western USA and Chile), as would be expected if these isolates represent a single, valid species. This low level of genetic variation, at most 0.079 \% among isolates in the SSU rDNA, was below that found among divergent strains of another distantly related microsporidian, Nucleospora salmonis (formerly Enterocytozoon salmonis) from salmon, which has 0.24 \% SSU divergence within isolates (Docker et al., 1997b). By comparison, the minimum and maximum differences among isolates of $N$. salmonis were higher, at $0.1 \%$ and $1.66 \%$, respectively (Gresoviac et al., 2000). This higher inter-isolate variation in N. salmonis from a broad range of geographic localities (British Columbia, California, Idaho, Oregon, Montana, Colorado, Nebraska, Washington, Chile, Norway, France) and several salmonid hosts (chinook and Atlantic salmon, rainbow, brook, and lake trout) further emphasizes the low level of genetic variation in $L$. salmonae, supporting the suggestion that $L$. salmonae is a valid, relatively genetically homogeneous species. Gresoviac et al. (2000) also reported an even higher difference ( $12.72 \%$ ) among isolates, compared to one isolate from English sole; however, this level of difference was so high that the English sole isolate is almost certainly a separate species, rather than an isolate of $N$. salmonis. Similarly, inter-isolate divergence in the SSU is higher (up to $1.73 \%$ ) in Enterocytozoon bieneusi, a close relative of $N$. salmonis (Zhu et al., 1994; Hartskeerl et. al., 1995). Intraspecific SSU variation has also been observed to be higher, at 0.2 to $11.4 \%$ in other microsporidians in genera Encephalitozoon, Vairimorpha, and Nosema (Baker et al., 1995). However, all of these species are so distantly related to $L$. salmonae that the degree of rDNA copy dispersion and homogenization or the overall mutation rate could differ significantly. Small subunit rDNA variation, at least interspecifically, has recently been shown to be lower in the mainly fish-parasitic clade in which L. salmonae falls (Nilsen et al., 1998; Cheney et al., 2000; Nilsen \& Chen, 2001; Lom \& Nilsen,
2003), therefore, this study confirms this low SSU variation also occurs at an intraspecific level for at least one species in this clade.

Similarly, the ITS region was completely invariant among all $L$. salmonae isolates in this study, whereas in E. bieneusi the ITS can have up to 1.2 to $3.9 \%$ variation (Breitenmoser et al., 1999). The ITS regions in L. salmonae and E. bieneusi are dramatically different in length (compare 37 to 243 bp ). This suggests the ITS in L. salmonae has lost length, or not gained length that is not highly functionally constrained, such that $L$. salmonae's ITS may be too short to allow the accumulation of non-deleterious sequence variation. The ITS is even longer in $N$. salmonis ( 660 bp ), and also has higher variation ( 1.16 to $3.09 \%$ ) than L. salmonae, but again, these differences may only suggest differences in constraints upon the gene, rather than biological differences in the divergence or mixing among the isolates.

Despite the low genetic variation in L. salmonae suggested by these results, rDNA sequence variation suggested a pattern divided at least partially along geographic lines into one variant observed in most British Columbian (BC) isolates, and the other variants found in nonBC isolates. By comparison, for $N$. salmonis, a distant relative of $L$. salmonae, genetic variation was also along geographic lines (Gresoviac et al., 2000), although $N$. salmonis shows an eastwest division and no particular pattern with BC vs. non- BC isolates.

In this study, rDNA variation was lower among BC isolates than among western USA (Colorado, California and Idaho) isolates, suggesting western USA farm-isolates may be more genetically diverse or contain separate $L$. salmonae variants or strains. Some of this higher variation, for example in isolate "Coll" from Colorado, is represented by shared polymorphisms (bold numbers in Table 2.2), while other variation is in the form of unshared substitutions (differences in single PCR products or single clones, e.g. numbers not in bold in Table 2.2). Unlike shared substitutions, unshared substitutions are more suspect, as they are more likely to be the result of sequencing artifacts. For example, unshared substitutions could be the result of Taq error; however, Taq error should be distributed randomly throughout the isolates sequenced in this study. Instead, results showed that unshared substitutions were clustered among a few farm/hatchery isolates, suggesting these were not caused by random Taq error. Furthermore, two laboratory isolates with unshared rDNA substitutions were among the few that also possessed EF-1 $\alpha$ substitutions, suggesting unshared substitutions also reflect real variation in $L$. salmonae, rather than Taq error. Whereas for a single-copy gene the question of Taq error might be addressed by repeated sequencing of PCR products, for multi-copy rDNA genes this solution cannot be applied, as PCR may amplify several copies of the gene simultaneously, so
such substitutions might be true polymorphisms in diverged rDNA copies or diverged strains of the parasite in the host gill. Therefore, the greater genetic diversity in $L$. salmonae rDNA among farm/hatchery isolates from non- BC localities appears to be valid and not merely due to sequencing artifacts. Perhaps more importantly, the more diverse isolates are also those from rainbow trout ( $O$. mykiss), suggesting host species or fresh water host habitat may be the isolating factor as much, or more than geographic locality. This idea is consistent with genetic results from another distantly related microsporidian with similar geographic and hostdistribution, N. salmonis. Gresoviac et al. (2000) found N. salmonis to have greatest rDNA sequence diversity in rainbow trout, as seems to be the case for $L$. salmonae in this study. This higher genetic diversity could suggest that rainbow trout was the ancestral host for $L$. salmonae. Similarly, L. salmonae may have originated in fresh water from another fresh water salmonid host. Or, alternatively, these results could suggest isolates in fresh water rainbow trout are simply more geographically and genetically isolated than are isolates in anadromous salmon, in which there is more opportunity for genetic mixing. A further explanation for greater genetic diversity in L. salmonae from farmed rainbow trout is that farms may obtain their stock and infections from multiple sources, a possibility which may have important consequences potentially enhancing the potential for evolution of highly virulent strains (Bull et al., 1991).

This study also presents some rDNA polymorphic "markers" that could be used as tools for identifying the geographic or even host origin and accompanying differences in biology of $L$. salmonae variants. These markers consist of three shared polymorphisms, a transversion at position 1575 in Colorado isolates (\#12), a "T" insert at position 1668 in Colorado and Californian isolates (\#13), and a Y/C transversion polymorphism at position 1725 in four isolates (\#14). In order for such markers to be informative, one would need to look at many more isolates from each source population to determine how widely these polymorphisms are shared. The present data, although representing a small sample, also present the possibility of a division in L. salmonae along host lines, corresponding with rainbow trout $O$. mykiss versus non-rainbow trout hosts, as the shared polymorphisms \#12 and \#13 occur only in isolates from rainbow trout. By comparison, N. salmonis showed greater sequence diversity in rainbow trout (Gresoviac et al., 2000) but did not show any particular "signature" or shared differences in species from such hosts, nor any other pattern characterizing sequences from a particular host.

Lack of variation among wild BC isolates of $L$. salmonae was unexpected, particularly considering the population substructuring of the parasite that should occur, given the wide geographic range of the host species, the variety of spawning habitats (river systems), and
spawning times of these hosts. If $L$. salmonae transmits largely during the host's fresh water phase, one would expect some substructuring based on this. One possible explanation for the uniform rDNA sequence in BC isolates is that the parasite normally transmits in seawater or at least not exclusively during freshwater phases of the salmon's life. Several studies support this hypothesis (Shaw et al., 2000c; Ramsay et al., 2002; Kent et al., 1995) suggesting that $L$. salmonae is a ubiquitous saltwater pathogen that intermixes throughout BC salmon hosts. Prevalence estimates found in this study (discussed below) also tend to support this hypothesis. The lack of variation in rDNA among BC L. salmonae isolates also presents some intriguing genetic questions. For example, could this be caused by bottleneck, recent colonization (founder effects) of $L$. salmonae in BC, or a low mutation/immigration rate in the rDNA compared to the rate of homogenization or concerted evolution among rDNA copies? Such speculations have little corroborating evidence, and would require a significantly better understanding of the basic biology and population genetics of the parasite at the least.

## Intraspecific differences in the $\mathrm{EF}-1 \alpha$ gene

Elongation factor- $1 \alpha$ was intended to act in this study as a marker independent of rDNA, to test against results from rDNA, which may be misleading because of the potential rDNA paralogs. Although EF-1 $\alpha$ could only be sequenced from a subset of the L. salmonae isolates, this gene appeared to be insufficiently variable to be an informative marker. The $0.03 \%$ interisolate variation in this gene (including several non-synonymous substitutions) is low enough to suggest Taq error could be responsible for a significant portion of the variation. Recent studies suggest EF-1 $\alpha$ may not really evolve independently of rDNA genes, even though these genes may not be closely linked physically on chromosomes, they may be under similar selective constraints due to their roles in translation (Moeira et al., 1999; Roger et al., 1999). Therefore, the unexpectedly low variation in both rDNA and EF-1 $\alpha$ in L. salmonae may arise from some common constraint on these two genetic regions. However, mutations observed in this study were at the intraspecific level and so they may be more likely to have accumulated in a nearly neutral way, thereby making it plausible that rDNA and EF-1 $\alpha$ are reasonably independent markers in L. salmonae. In conclusion, this gene may be potentially informative at this level, but a greater sequence length may be necessary to show informative variation within $L$.
salmonae.

## Prevalence of $L$. salmonae in five wild salmon species

The PCR survey showed that $L$. salmonae may occur in all five wild Pacific salmon species in BC , a result that is consistent with both experimental transmission studies (Shaw \& Kent, 1999; Shaw et al., 2000b; Ramsay et al., 2002), and numerous studies of wild salmon (reviewed in Shaw \& Kent, 1999). The PCR survey also provided some evidence that the frequency of $L$. salmonae may be highest in chinook salmon, followed by coho salmon, and lowest in chum and pink salmon. Other studies have shown similar patterns in potential chinook and coho host preference (Shaw et al., 2000a; Ramsay et al., 2002) although more complete data are needed. These prevalence data are imprecise due to the pooling of samples, and so only provided broadly overlapping, general estimates; however, they suggest that prevalence of this parasite differs depending on host species. This could result from innate susceptibility differences. For example, salmon species (or populations) could differ in innate immunity to infection with $L$. salmonae and other parasites due to differences in allelic diversity in the major histocompatability complex (MHC), as was demonstrated experimentally for sticklebacks exposed to G. anomala (Kurtz et al., 2004). Alternately, salmon species may differ in the proportion of individuals with resistance developed due to prior exposure (Speare et al., 1998b; 1998c; Shaw et al., 2000a; b; Sánchez et al., 2001 a; Ramsay et al., 2002). Lack of variation in rDNA sequence in these isolates suggests overall genetic homogeneity of $L$. salmonae, particularly in the wild. This suggests prevalence differences were more likely to be the result of differences in the hosts' responses or histories of exposure, rather than due to differences among divergent strains.

Although the overall prevalence estimated using PCR in this study would appear to be only about 5-28\% of salmon infected, the true level of infection could be far higher, both because the PCR test has limits to its sensitivity and because $L$. salmonae may have occurred in other tissues, such as the heart, for some time (Sánchez et al., 2000; Sánchez et al., 2001c). In addition, some specimens with Loma-like xenomas in the gills or positive PCR-test results failed to amplify with sequencing primers, possibly indicating the presence of a different Loma species or perhaps a different strain of $L$. salmonae that differs at the rDNA sequencing primer sites. The sequencing primer sites have been shown to amplify a range of Loma species (Brown et al., 1998; and see Chapter 4). Problems with inhibitors would be a plausible explanation. Many of these samples could not be sequenced after inhibitors were removed, suggesting the parasite DNA was lost during inhibitor removal, or spore DNA may have become degraded or too sheared to amplify with sequencing primers due to handling at some stage or inadequate fixative penetration through thick spore walls. In this study, inhibitors were frequently observed in gills,
particularly those from seawater. Future researchers, especially those employing the PCR-test with this marker, should take care to examine false negatives by some procedure similar to the one used here, or develop a DNA isolation protocol that effectively removes these inhibitors. Gresoviac et al. (2000) reported similar difficulties in amplifying some N. salmonis isolates, suggesting there may also be variation at priming sites or there may be cryptic species in this microsporidian group.

## The question of strain "SV"

Sánchez et al. (2001a; 2001b) hypothesized a strain or variant of L. salmonae with a preference for brook trout, Salvelinus fontinalis, and a lower virulence and transmissibility to rainbow trout or chinook salmon (Sánchez et al., 2001a; b) occurred in brook trout gills after they were fed L. salmonae spores from rainbow trout at Atlantic Veterinary College, Prince Edward Island (P.E.I.), Canada (Sánchez et al., 2001a; b; Speare \& Daley, 2003). The present study investigated sequence from the resulting infections in brook trout and also many isolates of spores from the same laboratory and nearby wild locations from which the spores used in experimental infections in P.E.I. originated (namely, laboratory-reared chinook salmon from the Pacific Biological Station, in British Columbia, Canada). Sánchez et al. (2001a) observed that it was peculiar to find brook trout that were experimentally fed spores in the first two trials could not be infected (Speare et al., 1998a; Shaw \& Kent, 1999), whereas, on a 3rd trial became heavily infected. Still more surprising was the observation that the resulting spores were found to be only very weakly infective to rainbow trout (Sánchez et al., 2001a). The authors suggested that in initial trials the brook trout may have been resistant due to prior exposure to $L$. salmonae, but they could not explain the different biological features of the resulting spores, except to suggest that this was from a genetic variant, which they named "L. salmonae SV".

In this study, rDNA and EF-1 $\alpha$ sequences were obtained from both the starting laboratory populations of L. salmonae that were brought to the P.E.I. laboratory for exposure to rainbow trout and brook trout, and from the hypothesized strain "L. salmonae SV" in brook trout in the PEI laboratory. Results show that both rDNA and EF sequences from the "SV" strain differed from those of BC L. salmonae, at a level greater than that found between sibling species of Loma (Shaw et al., 1997; Brown et al., 1998; Brown \& Kent, 2002; Matthews et al., 2001; and see Chapter 4). While a thorough analysis of these results will be presented in depth in Chapter 4, here I note that $L$. salmonae and the hypothesized "SV" strain possessed 19 differences ( $1.4 \%$ ) in the SSU and 10.5 differences ( $2.3 \%$ ) in the LSU, levels that are higher
than those found between any $L$. salmonae isolates in BC , California, Idaho, Colorado, or Chile (refer to Table 2.2), and greater than the level observed between L. salmonae and two sister species (L. embiotocia and an undescribed Loma sp. from lingcod Ophiodon elongatus from Kent et al., 1998) that were not transmissible (Shaw \& Kent, 1999; Shaw et al., 2000a; R. W. Shaw, personal communication) to reciprocal hosts (compare L. salmonae vs. L. embiotocia $0.75 \%$ and $1.2 \%$ and $L$. salmonae vs. Loma sp. from lingcod $1.0 \%$ and $1.0 \%$, for SSU and LSU, respectively). The partial EF-1 $\alpha$ sequence difference between these species ( $2.3 \%$ ) also differed beyond that found in L. salmonae isolates in this study, or other species pairs in genus Loma (see Chapter 4), which implies the "SV" strain could be a distinct species of Loma, rather than a strain.

Percent sequence difference aside, the "SV" strain hypothesis is unlikely based on the observation that DNA sequence from the same laboratory source populations over a range of years before, during, and after "SV" turned up did not show any sequence variation of this level in any clone or PCR product, despite some 40,000 rDNA nucleotides sequenced. In comparison, gills from five individual brook trout with the "SV" infection were analyzed, and all found to have only very small sequence divergence ( $0.09 \%$ in SSU and $0.1 \%$ in LSU), which implies "SV" did not come from the BC laboratory but instead raises the possibility that this "SV" variant is a separate species of Loma from those in the Pacific northwest and western USA. It may be native either to brook trout alone, rainbow trout and brook trout, or some other host. This theory as to the nature of the "SV" variant requires only that there was a cryptic infection in brook trout in the P.E.I. laboratory, a possibility worth considering in light of the ability of other Loma species to reside undetected in the heart or other internal organs (Speare et al., 1998b; Beaman et al., 1999; Sánchez et al., 1999; Sánchez et al., 2000; Sánchez et al., 2001c). In future studies, laboratory fish considered "naive" to Loma species infections or from hatcheries with no history of a reportable disease should be carefully examined for these species, perhaps using molecular probes.

A further question that remains to be addressed is whether this infection, named " $L$. salmonae SV" by Sánchez et al. (2001a; b), is conspecific with L. fontinalis Morrison \& Sprague, 1983 - a species with a preference for brook trout and a different morphology only under transmission electron microscopy. Attempts were made to amplify and sequence tissue containing positively identified $L$. fontinalis from brook trout close to the type locality in Halifax, Nova Scotia, Canada. While this tissue was weakly amplifiable using the L. salmonaespecific LS1 and LS2 PCR-primers of Docker et al. (1997a), the tissue had been fixed in
formalin and embedded in paraffin, and so insufficient product could be obtained for sequencing (Brown and Kent, 2002). The weak success of PCR-amplification suggests that L. fontinalis may be a close relative of $L$. salmonae. However, sequencing did show that the LS1 priming site of Docker et al.'s (1997a) probe is identical in L. salmonae, the "SV" variant or species from brook trout, and several other undescribed Loma species from British Columbia (Brown \& Kent, 2002, and Chapter 4). In particular, this similarity in one priming site of the probe suggests it might not be so $L$. salmonae-specific as was originally thought. The most compelling example is the identity of the LS1 priming site and close similarity of the LS2 priming site in L. salmonae and a Loma species from lingcod (described in Chapter 3), which was not found to be not transmissible experimentally to salmon (Shaw \& Kent, 1999), suggesting the probe may not be ideal for separating L. salmonae from relatives with welldemonstrated species-level morphological and biological differences.

Bader et al. (1998) and Markey et al. (1994) identified L. salmonae-like infections in salmonids from the eastern USA in brook, brown and rainbow trout. Spore and xenoma features from those studies suggest they may be slightly different from $L$. salmonae from BC or the type locality (California), such that the authors speculated they might be strains, variants, or species with an affinity for these fresh water hosts. Unfortunately, DNA sequence and detailed transmission electron-microscopic features were not obtained from those specimens, so the possibility remains that an eastern USA variant of L. salmonae or possibly a new sister species exists. One possible explanation for morphological differences between L. salmonae (type specimen) and Loma sp. from eastern USA (refer to Bader et al., 1998) is that $L$. salmonae diverged along an east-west gradient or divide across North America. A similar genetic divide exists between western North American/Pacific vs. eastern North American/Atlantic isolates of the salmonid pathogen $N$. salmonis (Gresoviac, et al., 2000). This suggests similar processes may be involved for L. salmonae; however, $N$. salmonis is only distantly related and has a higher rate of rDNA evolution, so such comparisons should be treated cautiously.

Even though genetic variation in L. salmonae's eastern range (eastern North America) is not yet known, the large sequence difference ( $1.4 \%$ in the SSU, $2.3 \%$ in the LSU, and $2.3 \%$ in the partial EF-1 $\alpha$ ) between L. salmonae and "SV" from brook trout compared to the low sequence difference (at most $0.079 \%-0.659 \%$ in the SSU and LSU respectively) among all isolates of $L$. salmonae from a wide geographic and host range presented here supports the idea that "L. salmonae SV" from brook trout is a separate species that does not occur in salmon or rainbow trout in western North America (and see Chapter 4 for further evidence). Presumably
the infection in the laboratory arose by means other than transmission from infected chinook salmon, such as from an underlying cryptic infection of laboratory brook trout.

## CONCLUSIONS

This study demonstrated the utility of Docker et al.'s (1997a) PCR-probe as a marker capable of detecting infections with L. salmonae, in that it is consistent with other studies suggesting chinook and coho salmon may be preferred hosts for L. salmonae. Each of the five species of wild pacific salmon carried $L$. salmonae at $2.3 \%$ prevalence (at least) at the time and location surveyed, suggesting $L$. salmonae may be widespread in the wild, perhaps often present as a low-level infection. This study also demonstrated the utility of rDNA and EF-1 $\alpha$ sequence as genetic markers of population-level differences in L. salmonae, although both gene regions varied at a low level compared to that in another microsporidian parasite of salmon ( $N$. salmonis). Intraspecific sequence variation supported the idea that L. salmonae is genetically homogeneous and a valid species with sequences from a wide range of localities and hosts nearly identical to those from the type host and type locality of $L$. salmonae, rainbow trout ( $O$. mykiss) from California. Ribosomal DNA, and to some extent also EF-1 $\alpha$ were able to show some population-level difference. For example, sequence variation was greater among farm/hatchery isolates in Colorado, California and in the freshwater host rainbow trout ( $O$. mykiss), showing that some shared polymorphic differences exist between geographic isolates. This may be useful in future for developing population-specific probes. Isolates from wild salmon in British Columbia were highly genetically homogeneous. For example, the SSU rDNA sequence was 0 to $0.079 \%$ different across isolates of $L$. salmonae-like infection from western North America and was invariant in L. salmonae from the wild in BC. The latter result could indicate that $L$. salmonae isolates exchange mutations through sexual recombination, or an alternative explanation for the lack of sequence variation in a parasitic microorganism is that a highly effective clone (or strain) has out-competed all others. The genetic similarity as well as laboratory experiments (Shaw et al., 2000b) suggest western L. salmonae may naturally infect both salmon and rainbow trout, even though the hosts are not presently sympatric through much of their range, and interestingly, spores from chinook salmon were shown, in some cases, to be able to infect brook trout under laboratory conditions (Shaw et al. 2000b). In contrast, the "SV" variant from brook trout from a laboratory in P.E.I., which was genetically different from all
isolates of $L$. salmonae, was also shown to be not very infective to either rainbow trout or salmon (Sánchez et al., 2001a; b; Speare \& Daley, 2003), preferring Salvelinus species as hosts. This suggests "SV" represents a distinct species under the working definition used here; however, it is not clear whether these species are naturally sympatric. A study of DNA variation across eastern North American L. salmonae-like infections would help address such species boundary questions.

## FURTHER INVESTIGATION

The results of this study suggest that rDNA, and particularly the partial LSU region can be informative for distinguishing $L$. salmonae isolates; however, a more variable marker is needed to provide sufficient variation for an informative statistical analysis of population substructuring among L. salmonae isolates from different hosts and localities. Preliminary results from dot-blot assays with a probe for GT repeat microsatellites (see Appendix 10) suggest that $L$. salmonae may have microsatellite-like repeats, but that these are probably rare, as hybridization was extremely weak. Preliminary attempts to develop an enriched microsatellite genomic library for $L$. salmonae were unsuccessful as there was evidence that DNA isolated from purified spores was far too sheared to obtain good growth of colonies with large-fragment inserts. Future studies could continue to better characterize rDNA or other gene differences in L. salmonae and to distinguish them from other Loma sp. infections in the eastern USA, Europe, Japan, and elsewhere. Future studies should examine genetic sequence data from L. fontinalis isolates from brook trout to determine whether the "SV" sequence is identical with L. fontinalis.

## ACKNOWLEDGEMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada strategic grant 582073 to M. L. Adamson. I wish to thank the crew of the research vessel W. E. Ricker for assistance in collection of salmon. I also gratefully acknowledge the assistance of fish farms for supplying access and material for this study. I am grateful to Dr. P. J. Keeling and his lab for assistance with cloning.

## LITERATURE CITED

Bader, J. A., Shotts, E. B., Steffens, W. L. and Lom, J. 1998. Occurrence of Loma cf. salmonae in brook, brown and rainbow trout from Buford Trout Hatchery, Georgia, USA. Diseases of Aquatic Organisms 34:211-216.

Baker, M. D., Vossbrinck, C. R., Didier, E. S., Maddox, J. V. and Shadduck, J. A. 1995. Small subunit ribosomal DNA phylogeny of various microsporidia with emphasis on AIDS related forms. Journal of Eukaryotic Microbiology 42(5):564-570.

Beaman, H. J., Speare, D. J., Brimacombe, M. and Daley, J. 1999. Evaluating protection against Loma salmonae generated from primary exposure of rainbow trout, Oncorhynchus mykiss (Walbaum), outside of the xenoma-expression temperature boundaries. Journal of Fish Diseases 22:445-450.

Bekhti, M. and Bouix, G. 1985. Loma salmonae (Putz, Hoffman et Dunbar, 1965) et Loma diplodae n. sp., microsporidies parasites de branchies de poissons teleosteens: implantation et donnees ultrastructurales. Protistologica 21(1):47-59.

Breitenmoser, A. C., Mathis, A., Bürgi, E., Weber, R. and Deplazes, P. 1999. High prevalence of Enterocytozoon bieneusi in swine with four genotypes that differ from those identified in humans. Parasitology 118:447-453.

Brown, A. M. V. and Kent M. L. 2002. Molecular diagnostics for Loma salmonae and Nucleospora salmonis (microsporidia) In Molecular diagnostics of salmonid diseases. Cunningham, C. O. (ed.). Kluwer Academic Publishers, Dordrecht p. 267-283.

Brown, A. M. V., Kent, M. L. and Adamson, M. L. 1998. Phylogeny of microsporidian parasites of fishes reveals dispersed ribosomal RNA genes in relatives of Loma salmonae. In: Program guide and abstracts - 73rd Annual Meeting of the American Society of Parasitologists 16-20 August 1998. Kona, Hawaii: American Society of Parasitologists, 1998. p. 80.

Bruno, D. W., Collins, R. O. and Morrison, C. M. 1995. The occurrence of Loma salmonae (Protozoa: Microspora) in farmed rainbow trout, Oncorhynchus mykiss Walbaum, in Scotland. Aquaculture 133:341-344.

Bull, J. J., Molineux, I. J. and Rice, W. R. 1991. Selection of benevolence in a host-parasite system. Evolution. 45(4):875-882.

Canning, E. U. and Lom, J. 1986. The Microsporidia of Vertebrates. Academic Press, Toronto, Canada, 289 pp.
Cheney, S. A., Lafranchi-Tristem, N. J. and Canning, E. U. 2000. Phylogenetic relationships of Pleistophora-like microsporidia based on small subunit ribosomal DNA sequences and implications for the source of Trachipleistophora hominis infections. Journal of Eukaryotic Microbiology 47:280-287.

Cho, S., Mitchell, A., Regier, J. C., Mitter, C., Poole, R. W., Friedlander, T. P. and Zhao, S. 1995. A highly conserved nuclear gene for low-level phylogenetics: elongation factor-1 $\alpha$ recovers morphology-based tree for heliothine moths. Molecular Biology and Evolution 12(4):650-656.

Docker, M. F., Devlin, R. H., Richard, J., Khattra, J. and Kent, M. L. 1997a. Sensitive and specific polymerase chain reaction assay for detection of Loma salmonae (Microsporea). Diseases of Aquatic Organisms 29:4148.

Docker, M. F., Kent, M. L., Hervio, D. M. L., Khattra, J. S., Weiss, L. M., Cali, A. and Devlin, R. H. 1997b. Ribosomal DNA sequence of Nucleospora salmonis Hedrick, Groff and Baxa, 1991 (Microsporea: Enterocytozoonidae): implications for phylogeny and nomenclature. Journal of Eukaryotic Microbiology 44(1):55-60.

Gresoviac, S. J., Khattra, J. S., Nadler, S. A., Kent, M. L., Devlin, R. H., Vivarès, C. P., de la Fuente, E. and Hedrick, R. P. 2000. Comparison of small subunit ribosomal RNA gene and internal transcribed spacer sequences among isolates of the intranuclear microsporidian Nucleospora salmonis. Journal of Eukaryotic Microbiology 47(4):379-387.

Hartskeerl, R. A., Van Gool, T., Schuitema, R. J., Didier, E. S. and Terpstra, W. J. 1995. Genetic and immunological characterization of the microsporidian Septata intestinalis Cali, Kotler and Orenstein, 1993: reclassification to Encephalitozoon intestinalis. Parasitology 110:277-285.

Hauck, A. K. 1984. Mortality and associated tissue reactions of chinook salmon, Oncorhynchus tshawytscha (Walbaum), caused by the microsporidian Loma sp. Journal of Fish Diseases 7:217-229.

Hazard, E. I. and Brookbank, J. W. 1984. Karyogamy and meiosis in an Amblyospora sp. (Microspora) in the mosquito Culex salinarius. Journal of Invertebrate Pathology 44:3-11.

Kent, M. L., Elliot, D. G., Groff, J., M. and Hedrick, R. P. 1989. Loma salmonae (Protozoa: Microspora) infections in seawater reared coho salmon Oncorhynchus tshawytscha. Diseases of Aquatic Organisms 20:231-233.

Kent, M. L. 2000. Marine netpen farming leads to infections with some unusual parasites. International Journal for Parasitology 30:321-326.

Kent, M. L., Dawe, S. C. and Speare, D. J. 1995. Transmission of Loma salmonae (Microsporea) to chinook salmon in sea water. Canadian Veterinary Journal 36:98-101.

Kent, M. L., Docker, M., Khattra, J., Vossbrinck, C.R., Speare, D. J. and Devlin, R. H. 1999. A new Microsporidium sp . (Microsporidia) from the musculature of the Mountain Whitefish Prosopium williamsoni from British Columbia: morphology and phylogeny. Journal of Parasitology 85(6):1114-1119.

Kent, M. L., Groff, J. M., Traxler, G. S., Zinkl, J. G. and Bagshaw, J. W. 1990. Plasmacytoid leukemia in seawater reared chinook salmon Oncorhynchus tshawytscha. Diseases of Aquatic Organisms 8:199-209.

Kent, M. L., Traxler, G. S., Kieser, D., Richard, J., Dawe, S. C., Shaw, R. W., Prosperi-Porta, G., Ketcheson, J. and Evelyn, T. P. T. 1998. Survey of salmonid pathogens in ocean-caught fishes in British Columbia, Canada. Journal of Aquatic Animal Health 10:211-219.

Kurtz, J., Kalbe, M., Aeschlimann, P. B., Haberli, M. A., Wegner, K. M., Reusch, T. B. and Milinski, M. 2004. Major histocompatability complex diversity influences parasite resistance and innate immunity in sticklebacks. Proceedings of the Royal Society of London Series B 271(1535):197-204.

Lom, J. and Nilsen, F. 2003. Fish microsporidia: fine structural diversity and phylogeny. International Journal for Parasitology 33:107-127.

Magor, B. G. 1987. First report of Loma sp. (Microsporida) in juvenile coho salmon (Oncorhynchus kisutch) from Vancouver Island, British Columbia. Canadian Journal of Zoology 65:751-752.

Mallet, J. 1995. A species definition for the modern synthesis. Trends in Ecology and Evolution 10:294-299.
Markey, P. T., Blazer, V. S., Ewing, M. S. and Kocan, K. M. 1994. Loma sp. in salmonids from the eastern United States: associated lesions in rainbow trout. Journal of Aquatic Animal Health 6:318-328.

Matthews, J. L., Brown, A. M. V., Larison, K., Bishop-Stewart, J. K. and Kent, M. L. 2001. Pseudoloma neurophilia, n. gen., n. sp., a new microsporidium from the central nervous system of the zebrafish. Journal of Eukaryotic Microbiology 48:227-233.

Moreira, D., Le Guyader, H. and Philippe, H. 1999. Unusually high evolutionary rate of the elongation factor $1 \alpha$ gene from the ciliophora and its impact on the phylogeny of eukaryotes. Molecular Biology and Evolution 16(2):234-245.

Morrison, C. M. and Sprague, V. 1983. Loma salmonae (Putz, Hoffman and Dunbar, 1965) in the rainbow trout, Salmo gairdneri Richardson, and L. fontinalis sp. nov. (Microsporidia) in the brook trout, Salvelinus fontinalis (Mitchill). Journal of Fish Diseases 6:345-353.

Nilsen, F. and Chen, W. J. 2001. rDNA phylogeny of Intrapredatorus barri (Microsporida: Amblyosporidae) parasitic to Culex fuscanus Wiedemann (Diptera: Culicidae). Parasitology 122:617-623.

Nilsen, F., Endresen, C. and Hordvick, I. 1998. Molecular phylogeny of microsporidians with particular reference to species that infect the muscles of fish. Journal of Eukaryotic Microbiology 45:535-543.

Putz, R. E., Hoffman, G. L. and Dunbar, C. E. 1965. Two new species of Pleistophora (Microsporidia) from North American fish with a synopsis of Microsporidia of freshwater and euryhaline fishes. Journal of Protozoology 12(2):228-236.

Ramsay, J. M., Speare, D. J., Dawe, S. C. and Kent, M. L. 2002. Xenoma formation during microsporidial gill disease of salmonids caused by Loma salmonae is affected by host species (Oncorhynchus tshawytscha, $O$. kisutch, O. mykiss) but not by salinity. Diseases of Aquatic Organisms 48:125-131.

Rodriguez-Tovar, L. E., Wright, G. M., Wadowska, D. W., Speare, D. J. and Markham, R. J. F. 2003. Ultrastructural study of the late stages of Loma salmonae development in the gills of experimentally infected rainbow trout. Journal of Parasitology 89(3):464-474.

Roger, A. J., Sandblom, O., Doolittle, W. F. and Philippe, H. 1999. An evaluation of elongation factor $1 \alpha$ as a phylogenetic marker for eukaryotes. Molecular Biology \& Evolution 16(2):218-233.

Sánchez, J. G., Speare, D. J. and Markham, R. J. F. 1999. Nonisotopic detection of Loma salmonae (Microspora) in rainbow trout (Oncorhynchus mykiss) gills by in situ hybridization. Veterinary Pathology 36: 610-612.

Sánchez, J. G., Speare, D. J. and Markham, R. J. F. 2000. Normal and aberrant tissue distribution of Loma salmonae (Microspora) within rainbow trout, Oncorhynchus mykiss (Walbaum), following experimental infection at water temperatures within and outside of the xenoma-expression temperature boundaries. Journal of Fish Diseases 23:235-242.

Sánchez, J. G., Speare, D. J., Markham, R. J. F. and Jones, S. R. M. 2001a. Experimental vaccination of rainbow trout against Loma salmonae using a live low-virulence variant of L. salmonae. Journal of Fish Biology 59:442-448.

Sánchez, J. G., Speare, D. J., Markham, R. J. F. and Jones, S. R. M. 2001b. Isolation of a Loma salmonae variant: biological characteristics and host range. Journal of Fish Biology 59:427-441.

Shaw, R. W. and Kent, M. L. 1999. Fish microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 418-446.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 1998. Modes of transmission of Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 33(2):151-156.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000a. Innate susceptibility differences in chinook salmon Oncorhynchus tshawytscha to Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 43:49-53.

Shaw, R. W., Kent, M. L., Brown, A. M. V., Whipps, C. M. and Adamson, M. L. 2000b. Experimental and natural host specificity of Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 40:131-136.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000c. Viability of Loma salmonae (Microsporidia) under laboratory conditions. Parasitology Research 86:978-981.

Shaw, R. W., Kent, M. L., Docker, M. F., Brown, A. M. V., Devlin, R. H. and Adamson M. L. 1997. A new species of Loma (Microsporea) in shiner perch (Cymatogaster aggregata). Journal of Parasitology 83(2):296-301.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2001 . Phagocytosis of Loma salmonae (Microsporidia) spores in Atlantic salmon (Salmo salar), a resistant host, and chinook salmon (Oncorhynchus tshawytscha), a susceptible host. Fish \& Shellfish Immunology 11:91-100.

Speare, D. J., Arsenault, G. J. and Buote, M. A. 1998a. Evaluation of rainbow trout as a model for use in studies on pathogenesis of the branchial microsporidian Loma salmonae. Contemporary Topics in Laboratory Animal Science 37:55-58.

Speare, D. J., Beaman, H. J., Jones, S. R. M., Markham, R. J. F. and Arsenault, G. J. 1998b. Induced resistance in rainbow trout, Oncorhynchus mykiss (Walbaum), to gill disease associated with the microsporidian gill parasite Loma salmonae. Journal of Fish Diseases 21(2):93-100.

Speare, D. J., Brackett, J. and Ferguson, H. W. 1989. Sequential pathology of the gills of coho salmon with a combined diatom and microsporidian gill infection. Canadian Veterinary Journal 30:571-575.

Speare, D. J. and Daley, J. 2003. Failure of vaccination in brook trout Salvelinus fontinalis against Loma salmonae (Microspora). Fish Pathology 38(1):27-28.

Speare, D. J., Daley, J., Markham, R. J. F., Sheppard, J., Beaman, H. J. and Sánchez, G. J. 1998c. Loma salmonaeassociated growth rate suppression in rainbow trout, Oncorhynchus mykiss (Walbaum), occurs during early onset xenoma dissolution as determined by in situ hybridization and immunohistochemistry. Journal of Fish Diseases 21(5):345-354.

Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. and Woese, C. R. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326(6111):411-414.

Wales, J. and Wolf, H. 1955. Three protozoan diseases of trout in California. Californian Fish and Game 41:183187.

Wheeler, Q. D. and Meier, R. 2000. Species concepts and phylogenetic theory. Columbia University Press. New York, NY. 230 pp.

Zhu, X., Wittner, M., Tanowitz, H. B., Cali, A. and Weiss, L. M. 1994. Ribosomal RNA sequences of Enterocytozoon bieneusi, Septata intestinalis and Ameson michaelis: Phylogenetic construction and structural correspondence. Journal of Eukaryotic Microbiology 41(3):204-209.


Figure 2.1: Polymerase chain reaction test (PCR-test) products in 1.5\% agarose gels showing positive results (a 272 bp band) of varying strengths for five ocean-caught Pacific salmon species (Oncorhynchus spp.). Positive results (presence of band) are shown in lanes 3, 4, 20, 21, $22,26,27,40,41,42,45,46,48,50,51,53,62,65$ and 74 where lanes $1-18=$ chum salmon, lanes 19-38 = coho salmon, lanes $39-57=$ chinook salmon, lanes $58-68=$ sockeye salmon, lanes $69-75=$ pink salmon. $5 \mu$ l of product was loaded into each lane. Lanes labeled " 1 kb " $=$ size marker, "+ve" = positive control L. salmonae DNA, and "-ve" = negative control.

Table 2.1: Wild Pacific salmon species examined for Loma salmonae using polymerase chain reaction (PCR) test, showing number of fish examined, number of L. salmonae-positive PCR results from DNA samples pooled from 5 fishes, and estimated range in actual prevalence based on PCR results.

| Host scientific <br> name | Common name | \# of fish <br> exam- <br> ined | \# of +'ve <br> PCR pools / <br> total \# of pools | Estimated <br> prevalence <br> $\%$ |
| :--- | :--- | :---: | :---: | :---: |
| Oncorhynchus keta | chum salmon | 88 | $2 / 18$ | $2.3-11.4$ |
| O. kisutch | coho salmon | 97 | $5 / 20$ | $5.2-25.8$ |
| O. shawytscha | chinook salmon | 95 | $10 / 19$ | $10.5-52.6$ |
| O. nerka | sockeye salmon | 51 | $2 / 11$ | $3.9-19.6$ |
| O. gorbuscha | pink salmon | 32 | $1 / 7$ | $3.1-15.6$ |
|  | TOTAL: | 363 | $20 / 75$ | $5.5-27.5$ |



Figure 2.2: Polymerase chain reaction test (PCR-test) products in 1.5\% agarose gel showing evidence of inhibition of the PCR reaction (due to inhibitors in template DNA) in several lanes. Left 2 lanes $=1 \mathrm{~kb}$ marker and positive control L. salmonae DNA; right lane $=$ negative control (water); lanes 1, 2, 3 and $4=$ separate sample templates showing "negative" PCR results; lanes $1 \mathrm{~s}, 2 \mathrm{~s}, 3 \mathrm{~s}$ and $4 \mathrm{~s}=$ positive "spikes", or samples in which L. salmonae positive control DNA was added to template DNA, showing that samples 1 and 3 are true negatives (and PCR inhibitors are not present), whereas sample 4 may be a false negative, as the positive spike failed to amplify (suggesting presence of PCR inhibitor), and sample 2 has a faint band (arrow), suggesting presence of some inhibitor in this sample.

Table 2.2: Loma salmonae isolates from laboratory, farms and wild populations for which ribosomal RNA and elongation factor-1 $\alpha$ genes were sequenced, this study. Table shows for each labeled isolate, its locality, host, notes, date, source, number of nucleotide positions sequenced for each gene, and numbered DNA differences corresponding to differences shown on Figs. 2.3 and 2.4. Wild isolates labeled with a * correspond to salmon pools from PCR-tests. $\mathrm{BC}=$ British Columbia, Canada, f. $=$ farm, $\mathrm{h} .=$ hatchery, chinook $=$ chinook salmon Oncorhynchus tshawytscha (Walbaum 1792), coho = coho salmon O. kisutch (Walbaum 1792), rainbow $=$ rainbow trout $O$. mykiss (Walbaum), pink = pink salmon $O$. gorbuscha (Walbaum 1792), sockeye = sockeye salmon O. nerka (Walbaum 1792). Notes: all laboratory isolates were from ethanol fixed gills except "spo" = purified spore concentrate; all farm/hatchery material was from non-broodstock, except "bro" = hatchery brood fish; all wild material was from prespawning or ocean phase fish except "spa" = returning spawner fish; Date is by month and year; Source refers to collector or supplier of tissue B = A. M. V. Brown, $\mathrm{S}=\mathrm{R}$. W. Shaw, $\mathrm{K}=$ M. L. Kent, D = S. C. Dawe. Numbered DNA differences shown in bold numbers represent shared differences (e.g. difference \#14 is shared among BA4, L6, CAL, Col1, Col2, and Lsc2).


| L. salmonae from laboratory-reared salmon |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SPO | Nanaimo BC | chinook | spo | 5.97 B | 1092 |  | - |
| N-14 | " | " |  | 7.97 S | 491 |  | - |
| C-24 | " | " |  | 9.97 S | 1080 |  | 969 |
| BA2 | " | " | spo | 11.97 B | 1082 |  | 1054 |
| BA3 | " | " | spo | 4.99 B | 1086 |  | - |
| BA4 | " | " | spo | 6.99 B | 1760 | 14,15 | 1040 4,5,6 |
| BA9 | " | " | spo | 4.00 B | 1829 | 2,4,9 | 10691 |

L. salmonae from salmon and trout in farms and hatcheries

| I-27 | Indian Bay f. BC | chinook |  | 10.97 S | 1081 | 10 |
| :--- | :--- | :--- | ---: | ---: | :--- | ---: |
| IB-27 | $"$ |  | 10.97 S | 1087 |  | - |
| KC-17 | Kuncchin h. BC | $"$ | bro | 11.97 S | 1098 | 959 |
| COR | Seashelt f. BC | coho | 2.98 S | 1083 | 954 |  |
| L6 | California f. USA | rainbow | 6.96 K | 1816 | $1, \mathbf{1 3 , 1 4}$ | - |
| CAL | $"$ | $"$ | 6.96 K | 1084 | $\mathbf{1 3 , 1 4}$ | - |
| Li | Idaho h. USA | $"$ | 3.96 K | 572 | 944 |  |
| Col1 | Colorado h.USA | $"$ | 3.98 K | 1111 | $3,6,7,8,12,13,14$ | - |
| Col2 | $"$ | $"$ | 3.98 K | 1111 | $\mathbf{1 2 , 1 3 , 1 4}$ | - |
| Lsc2 | Chile f. | coho | 12.00 K | 1812 | $5,11,14,16$ | 1069 |

L. salmonae from wild salmon

| S52* | Texada Isl. BC | chinook |  | 9.97 D | 514 |
| :--- | :--- | :--- | ---: | ---: | ---: |
| C237 | Barkley Sd. BC | coho | 5.97 B | 1081 | - |
| S21* $^{*}$ | Cape Mudge BC | $"$ | 9.97 D | 506 | - |
| S41* $^{*}$ | Ballenas Isl. BC | pink |  | 9.97 D | 508 |
| LOSO | Gr. Central L. BC | sockeye | spa | 11.96 S | 1090 |
| SP-24 | Sproat River BC | $"$ |  | 7.97 S | 1082 |
| ST-24 | Stamp River BC | $"$ |  | 7.97 S | 230 |
| S71* | Boundary B. BC | $"$ |  | 9.97 D | 1078 |
| SF-25 | Fulton River BC | $"$ | spa | 9.97 S | 1078 |

Table 2.3: Percent intraspecific or intra-isolate difference in ribosomal DNA sequence for Loma salmonae isolates for all material sequenced, or for individual isolates from laboratory and farm populations, compared to a single reference sequence from British Columbia (BC).

| Isolate | Number of rDNA <br> nucleotides <br> sequenced (all PCR <br> products or clones) | Intraspecific or intra- <br> isolate \% difference <br> from BC-form |  |  |
| :--- | :---: | :---: | :---: | :---: |
|  | SSU | LSU | SSU | LSU |
| L. salmonae (all) |  |  |  |  |
|  | 35320 | 21228 | 0.017 | 0.113 |
| L. salmonae from laboratory |  |  |  |  |
| BA4 (BC) | 4125 | 1482 | 0 | 0.101 |
| BA9 (BC) | 6624 | 1877 | 0.045 | 0.160 |
| L. salmonae from farms |  |  |  |  |
| I-27 (Idaho) | 987 | 567 | 0 | 0.176 |
| L6 (California) | 1833 | 1075 | 0.027 | 0.279 |
| CAL ( " ) | 1129 | 469 | 0 | 0.320 |
| Col1 (Colorado) | 1261 | 1441 | 0.079 | 0.659 |
| Col2 (Colorado) | 1080 | 672 | 0 | 0.521 |
| Lsc2 (Chile) | 4495 | 1989 | 0.022 | 0.151 |



Figure 2.3: Loma salmonae intraspecific differences in the small subunit (SSU), internal transcribed spacer (ITS) and large subunit (LSU) ribosomal DNA sequence, showing all variants, numbered in gray, against the reference sequence for the British Columbia-form isolate. A previously published $L$. salmonae sequence (Genbank accession number U78736) is shown with dots where nucleotides are identical to those of the BC-form isolate. Sequence is reported from $3^{\prime}$ to 5 ' and begins with the SSU. The ITS region is enclosed in a box. L.sal BC = BC-form isolate. L.sal U78736 = Genbank sequence.


Figure 2.4: Loma salmonae elongation factor-1 $\alpha(\mathrm{EF}-1 \alpha$ ) DNA and protein sequences, showing all variants, numbered in gray, against reference sequence for the most common isolate (BCform). Glugea plecoglossi from Genbank DNA and protein sequences (accession number D32139) are shown with dots where nucleotides or amino acids are identical to the BC-EF-1 $\alpha$ isolate. G.ple Prot $=$ Glugea plecoglossi EF-1 $\alpha$ Protein sequence, $L . s a l$ Prot $=$ proposed $L$. salmonae $\mathrm{BC}-\mathrm{EF}-1 \alpha$ protein sequence, G.ple $\mathrm{D} 32139=$ Glugea plecoglossi $\mathrm{EF}-1 \alpha$ Genbank sequence D32139. L.sal $\mathrm{BC}=\mathrm{BC}-\mathrm{EF}-1 \alpha \mathrm{DNA}$ sequence.

# Chapter 3: Description of five new species of Loma (Microsporidia) parasitizing Pacific fishes. 

## INTRODUCTION

Loma species, like other microsporidians, are tiny ( $\sim 2-6 \mu \mathrm{~m}$ ), spore-forming, singlecelled, intracellular parasites that undergo a number of vegetative divisions in contact with their host's cell cytoplasm or in sacs created by the parasite or host, eventually culminating in the formation of infective spores (for detailed review of microsporidian development see Vávra \& Larsson, 1999, and Cali \& Takvorian, 1999). Undescribed microsporidians resembling species of Loma (Morrison \& Sprague, 1981) were reported by Kent et al. (1998) in the gills of 5 Pacific fishes: Pacific cod Gadus macrocephalus Tilesius 1810, walleye pollock Theragra chalcogramma (Pallas 1811), Pacific tomcod Microgadus proximus (Girard 1854), lingcod Ophiodon elongatus Girard 1854, and sablefish Anoplopoma fimbria Pallas 1811, in British Columbia, Canada. These parasites were recognized as Loma species by the formation of large (up to $\sim 1 \mathrm{~mm}$ ), spore-filled, cyst-like nodules consisting of highly modified host-parasite complexes, called "xenomas", in the gills or other vascularized tissues of fishes. Other widespread Loma species cause serious losses to commercial fisheries, for example L. salmonae (Putz, Hoffman \& Dunbar, 1965) in salmon and trout (Oncorhynchus spp.) and L. camerounensis Fomena, Coste \& Bouix, 1992 in tilapia (Oreochromis niloticus (L.)) (Kent et al., 1989; Kent et al., 1995; Fomena et al., 1992; Shaw \& Kent, 1999). Xenomas and spores in other Loma species may occlude blood vessels, cause severe inflammatory response or death (Kent et al., 1989; Hauck, 1984; Speare et al., 1989; Shaw \& Kent, 1999). These recently recognized Loma-like infections in British Columbia occur in some commercially important fishes (lingcod, sablefish, pollock, and Pacific cod) and may be a threat to developing aquacultures (e.g. for sablefish).

Loma species, like most microsporidians, have historically been differentiated based on the host species and locality, and especially by the morphology of the most obvious stages, the spores. Spores alone are inadequate to distinguish species, as spore ultrastructure and dimensions in Loma species can vary as much within as they do among proposed species
(Kabata, 1959; Morrison \& Sprague, 1981a; Bekhti \& Bouix, 1985; Shaw et al., 1997; Azevedo \& Matos, 2002). This problem is exacerbated by the historical and sometimes present tendency to report sizes of spores or other structures as only a mean or a range of values, without any indication of the variance. In addition, spores shrink when fixed and other features are sensitive to measurement technique, making it difficult to compare studies by different researchers. Therefore species of Loma often must be distinguished using detailed transmission electron microscopic (TEM) study of features of the xenoma, spores, and particularly details in the developmental sequence or timing.

Loma species have some of the most basic of the microsporidian life cycles. They are directly transmitted and monosporous, with always unpaired, uninucleate cells. They undergo a set of vegetative divisions before forming stages (called sporonts) destined to form spores. Enough vegetative divisions occur in Loma species to fill and greatly extend the volume of the host cell, causing it to become an enlarged and highly modified host-parasite complex (the xenoma). As Loma species are usually found in tissues only when the xenomas are well underway, parasites tend to be seen more from this point on. Vegetative stages called meronts or merogonial plasmodia eventually finish "merogony" and begin "sporogony", the sporeproducing phase, which begins with the development of a thick surface coat and the creation of a sac (the parasitophorous vacuole) around dividing cells. The sac is created by the coalescence of vesicles derived from some unknown process in the host (or xenoma) cytoplasm (Morrison \& Sprague, 1981a; b; Lom \& Pekkarinen, 1999). These dividing stages, now called sporonts or sporogonial plasmodia, often produce species-characteristic numbers pre-spore stages, called sporoblasts, within which early spore structures form. Spores contain a variety of speciesdiagnostic features varying in numbers or sizes, such as the coiled polar filament, a large vacuolar space, and a thick, two-layered wall. There appear to be several species-diagnostic features in the material filling the xenoma cytoplasm and vesicles which coalesce around, and spill their contents into the space (episporontal space) enclosing the developing stages or spores (Bekhti \& Bouix, 1985, Lom \& Pekkarinen, 1999; Morrison \& Sprague, 1981a). Episporontal inclusions that appear to be species-diagnostic include the poorly understood "tubules" and "amorphous granular material" that possibly contribute some important material to the spore wall or space; or they help "communicate" material between the spore and xenoma cytoplasm (Morrison \& Sprague, 1981a; b; Vávra \& Larsson, 1999).

Species-diagnostic characters may be found at many points in microsporidian life-cycles. These characters have to be inferred, sometimes with great difficulty, from TEM sections. One
notoriously difficult character to discern in Loma species is the number of divisions in the final stage prior to spore formation, a process called sporogony, which leaves spores together in characteristic pairs, foursomes, or larger groups contained together in sporophorous or parasitophorous sacs. Often, cells can be deceptively beyond the plane of section. This problem has occurred with the type species, $L$. morhua Morrison \& Sprague, 1981, upon which the genus was defined (Morrison \& Sprague, 1981a; Lom \& Pekkarinen, 1999). Another problem that confuses species descriptions is the question of the origin (host vs. parasite), timing and mode of formation of the membranous sac around the parasites. Lom \& Pekkarinen (1999) argued that this character is important in determining whether Loma species are valid members of the group. They argued that valid members of the group should form this sac by coalescence of host vesicles, creating a host-derived "parasitophorous vacuole" (PV), whereas taxa with parasitederived sacs formed by membranous coat-delamination, creating parasite-derived "sporophorous vesicles"(SPV) are not valid Loma species. Lom \& Pekkarinen (1999) transferred a species to genus Loma from Glugea on the basis of this character. Lom \& Nilsen (2003) later found rDNA data placed this species, Loma acerinae (Jírovec, 1930) Lom \& Pekkarinen, 1999, not in the genus, but as a distant relative of Loma species. This suggested the character might not be as good as was thought. Lom \& Nilsen (2003) showed that several other species-diagnostic characters (e.g. early developmental stages and xenoma cytoplasmic features) have been insufficiently studied in this group, and need more thorough study.

The initial purpose of this study was to distinguish and describe these new Loma species ${ }^{1}$ by comparing them morphologically to the type species ${ }^{2}$, L. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981, and to the 11 other named species of Loma. Towards this end, spores were also collected from the type species' type host, haddock Melanogrammus aeglefinus (L.), as well as from the formerly recognized species L. morhua Morrison \& Sprague, 1981, considered to be synonymous with L. branchialis by Canning \& Lom (1986) and Lom (2002), from its type host, Atlantic cod Gadus morhua L., and type locality, Halifax, Nova Scotia, Canada. As a working species definition, I will use a combination of criteria based on

[^0]arguments in Wheeler \& Meier (2000). Here, the first criterion for distinguishing one species from another when in sympatry is that each must possess separate discrete characters or overlapping characters with a statistically separate mean. The second criterion will be that the suite of characters that provides evidence for species must agree across the sampled populations. The five new species studied here are potentially sympatric with other Loma species from 12 host fish species (Pacific cod, walleye pollock, Pacific tomcod, lingcod, sablefish, shiner perch, and six species of Pacific salmon and trout of Oncorhynchus spp.) in coastal British Columbian waters. The more complex hypothesis of two or more species, rather than one, will only be accepted if it fits the data better. The second goal of this study was to consider the comparative pathology, prevalence, ultrastructural, and developmental features of these new species, and their biological implications. The third objective of this study was to better characterize features that have been important traditionally in differentiating Loma species, but as yet have been inadequately examined. These features include merogonial and sporogonial stages, developmental divisions, the material in the xenoma cytoplasm that contributes to the sac around the parasite, and episporontal inclusions (Lom \& Pekkarinen, 1999; Lom \& Nilsen, 2003). The last, more practical objective of this study was to characterize differences among the species recognizable using basic microscopy, which could aid researchers in the field. Ribosomal DNA sequences are also reported in species descriptions to allow others to confirm identity of their specimens with this material.

## MATERIALS AND METHODS

## Specimen collection

Ten to 30 individuals from 77 species of fish were examined for signs of Loma-like infection in the gills (see Kent et al. 1998) and several fish species with possible signs of infection were sampled more extensively (see Table 3.1). Gills and other tissues were examined fresh at 400X and 1000X magnification and in many positive or suspicious cases for Loma spp. also by histological section, electron microscopy, and DNA sequence (described below). Table 3.1 includes fish species with suspected cases infections that were found at sea and in fish markets in Pacific and Atlantic Canada, and is broken into two groups to show those species for which gills proved positive or negative for Loma-like infection.

Fishes collected at sea were obtained between March 8, 1996 and May 16, 1999, by bottom trawling from several localities off the south, west and east coasts of Vancouver Island, British Columbia, as part of a government (Fisheries and Oceans Canada) shrimp survey. Trawling covered an area from 3 to 48 km distance from shore, 65 to 230 m depth, over a northsouth distance of about 300 km (from Juan de Fuca Strait lat/long 48.15 N 124.00 W , to the region off Barkley sound lat/long 48.50 N 125.20 W and up to Queen Charlotte Sound lat/long 51.20 N 129.00 W ). Fishes were killed with a blow to the head prior to examination.

Heads from Atlantic cod Gadus morhua L., haddock Melanogrammus aeglefinus (L.) and Pacific cod Gadus macrocephalus Tilesius 1810, were collected between 1996 and 1998 from markets in Nanaimo, British Columbia and Halifax, Nova Scotia, Canada.

Experimental laboratory-reared fishes used in transmission experiments were obtained and maintained as described in Ramsay et al. (2002) at the Pacific Biological Station, Nanaimo, Canada. These fishes were killed with an overdose of tricaine methanesulfonate (MS-222) and examined as described for other species.

## Tissue fixation and preparation

Tissue was examined in wet mounts as soon as possible post-mortem. Gill arches were clipped into 1 cm sections, thicker bony arches were cut away, and filaments were spread between slides and cover slips, in saline. Tissue from other organs was cut into smaller pieces and squashed between slides. Material for histological examination was fixed in Davidson's solution (Humason, 1979) then embedded in paraffin, sectioned and stained with hematoxylin and eosin. Tissue for transmission electron microscopy was fixed in 4\% glutaraldehyde for 12
to 24 hours, soaked in either Millonig's solution or Sörensen's solution ( 0.1 M ) for 24 hours and post fixed in $1 \%$ osmium tetraoxide for 1 hour. Tissue was embedded in standard Spurr's resin. Ultra thin sections were lifted onto copper grids and stained with $2 \%$ aqueous uranyl acetate followed by standard Reynold's lead citrate. For DNA study, tissues were fixed in $95 \%$ ethanol, or in some cases frozen.

## Prevalence and intensity

Gills were examined for xenomas first at low power ( 100 X or 400 X ), and then at highest power (1000X) to confirm negative or positive infections. About $6-8 \mathrm{~cm}$ of gill tissue from a single gill arch were examined from each fish. Prevalence (percent of hosts having xenomas in the gills) was determined by light microscopy from both wet mounts and histological sections. In some cases PCR was also used to confirm positive infection. Chi-square tests were performed to compare wet mount and histological prevalence estimates by combining data from both methods to generate an overall mean. This mean was used as the expected value in chisquare tests. Chi-square tests were also performed to compare prevalence between Loma species from different host taxonomic groups (family Gadidae or order Scorpaeniformes) using the pooled data from all species from each host group as the expected mean.

Intensity of infection was determined as described in Speare et al (1998a), by counting the number of well-oriented gill filaments and the number of xenomas on those filaments, giving a value of number of xenomas per gill filament (xpf).

## Xenoma size

Xenomas were measured from histological sections and resin-embedded sections by light microscopy (400X and 1000X) using an ocular micrometer that had been calibrated with a slide micrometer. Xenomas that were enlarged, not packed with spores, and without sharp walls may have been undergoing host-cell infiltration and granuloma formation, so might be altered greatly in size and shape from these processes. Such xenomas were not counted. Xenomas were also not counted if they appeared to be in "graze" cut (a cut through just an edge). Graze cuts could be recognized by the lack of sharpness of the wall, an unusually thick wall, and the lack of hostnucleus. Although these efforts were made to select mature xenomas cut closer to their middles rather than just through an edge (i.e. "graze" cuts), one would expect these size measurements to include a proportion of sub-medial cuts, producing average sizes that are smaller than the true average, but this bias would be expected to be the same for all species.

Statistical support for xenoma size differences was assessed by both parametric and nonparametric methods (two-sample t-test and F-tests to determine equivalence of variances; and Mann-Whitney test using normal approximation where necessary). If parametric and nonparametric tests produced a different result, both were reported. Statistical tests of xenoma size difference were performed in instances where xenoma size was similar (overlapping or close confidence intervals) and where xenomas appeared to arise from a similar cell or location in the gill.

## Measurements using transmission electron microscopy (TEM)

Ultrastructural features observed by transmission electron microscope (TEM) were photographed on a Zeiss 10C Transmission Electron Microscope and negatives were scanned at high resolution ( 1200 dpi ) and opened in Adobe Photoshop 6.0. The program's zoom feature was used to enlarge structures for measurement (e.g. $\sim 20 \mathrm{~cm}$ for spores and $\sim 2 \mathrm{~cm}$ for merogonial surface coats) while maintaining reasonable resolution. Structures were measured against the screen using a ruler with mm marks. Scale bars were calculated for each zoom level using magnification factors calculated for the electron microscope, and these were used to calculate actual sizes of structures measured on screen.

Round structures, such as vesicles, meronts and spores, that appeared to be cut through an edge ("graze" cuts) were not counted. Spores were easily recognized as being sectioned close to the middle at the longitudinal axis by internal features such as complete rows of roundpolar filaments. Graze cuts through meronts, sporonts and plasmodia, were less easily recognized; however, presence of the nucleus and particularly one or more relatively large nuclei were indicators of sections being more close to the middle. Graze cuts through smaller round structures, such as empty vesicles were also difficult to assess, so graze cuts were included. Surfaces of structures such as merogonial surface coats, endospores, xenoma walls, parasitophorous vacuoles could also be cut in graze. Graze cuts through surface structures were recognized by unusual thickness and an unusually poorly delineated, non-sharp edge.

Vesicles in the xenoma cytoplasm that were near meront, sporont, sporoblast, or spore may be destined to contribute material to the parasite surface, so these vesicle numbers were counted using a rule for counting purposes that vesicles were "associated" with a given cell if they were found within a space delimited by a line drawn halfway between the surface of the cell and adjacent cells. While this definition may not reflect the true destiny of vesicles with
respect to neighbouring cell surfaces, it allowed at least that counts could be made such that erroneous assumptions about vesicle destiny would be biased in the same way for all species.

Where possible, TEM features were measured for several individuals and several xenomas, as follows: seven xenomas in four individuals from Pacific cod; three xenomas in two individuals from walleye pollock and sablefish; two xenomas in one individual from Pacific tomcod and lingcod.

Statistical support for differences in size or numbers of TEM structures was assessed by analysis of variance (ANOVA) and Kruskal-Wallis tests (parametric and non-parametric, respectively) prior to their corresponding multiple-comparisons tests Tukey and Nemenyi tests, or where only two samples were to be compared, by an F-test for equal variances prior to a t-test and Mann-Whitney U-test, as described in Zar (1996), using Microsoft Excel. Only parametric test results were reported here where both parametric and non-parametric results agreed.

## Spore measurement and conversion factors

Spores from semi-thin resin-embedded sections, histological sections, and ethanol-fixed material in wet mount were photographed under light microscopy at 1000X magnification. Photographic negatives were scanned at high resolution (2000 dpi), opened in Adobe Photoshop 6.0, and spores were measured on screen as described above for TEM structures. Fresh and frozen spores were measured at 1000 X magnification using and ocular micrometer that had been calibrated using a slide micrometer to an estimated accuracy of $\pm 0.05 \mu \mathrm{~m}$.

Spore shrinkage due to fixation was estimated first for individual species by calculating the ratio spore sizes from each pair of fixation methods (e.g. fresh spore size mean/resinembedded spore mean) where spore sizes from two fixation methods were available. Length and width shrinkage were calculated separately. Then, mean spore shrinkage across several species was calculated for each pair of fixation methods. Average shrinkage across species was calculated from four species for resin-embedded to histological material (L. pacificodae n. sp. n $=12$ and 10 spores, $L$. wallae $\mathrm{n} . \mathrm{sp} . \mathrm{n}=11$ and 10 spores, $L$. kenti $\mathrm{n} . \mathrm{sp} . \mathrm{n}=12$ and 12 spores, $L$. lingcodae $\mathrm{n} . \mathrm{sp} . \mathrm{n}=10$ and 17 spores), from three species for resin-embedded to fresh material (L. pacificodae n. sp. $\mathrm{n}=12$ and 30 spores, $L$. lingcodae n. sp. $\mathrm{n}=10$ and 30 spores, $L$. salmonae $\mathrm{n}=10$ and 30 spores), and from two species for ethanol to frozen material (spores from Atlantic $\operatorname{cod} \mathrm{n}=13$ and 24, spores from haddock $\mathrm{n}=10$ and 35 ). Shrinkage was only available for one species for resin-embedded to ethanol material (a shrimp microsporidian $\mathrm{n}=$ 14 and 15 ), and fresh to frozen material (L. pacificodae $\mathrm{n} . \mathrm{sp} . \mathrm{n}=10$ and 30 spores). The
resulting shrinkage factors were used to convert spore sizes from fixed material into estimate fresh spore sizes where fresh spore sizes were unavailable.

## Polymerase chain reaction (PCR) and DNA sequencing

About 60 to 65 mg of ethanol-fixed tissue (approximately eight gill filaments) was used in each DNA extraction. Tissue was soaked for 15 minutes in lysis buffer ( 10 mM Tris, 1 mM EDTA, $10 \mathrm{mM} \mathrm{NaCl}, 1 \% \mathrm{SDS}$ ) to remove ethanol, and digested in buffer with $0.5 \mathrm{mg} / \mathrm{ml}$ proteinase K for $4-6$ hours at $37^{\circ} \mathrm{C}$ in a rotating incubator. The aqueous layer was extracted once from phenol, twice from phenol: chloroform: isoamyl alcohol (25:24:1), and once from chloroform: isoamyl alcohol (24:1). DNA was precipitated in cold $95 \%$ ethanol, washed twice with $70 \%$ ethanol, vacuum dried, resuspended in $40 \mu$ distilled water and stored for use at -20 ${ }^{\circ} \mathrm{C}$. Extracted genomic DNA was amplified by the polymerase chain reaction (PCR) in $25 \mu \mathrm{l}$ reactions using 0.3 to $0.8 \mu$ g genomic DNA, standard PCR buffer (Boehringer Mannheim, Germany), $2.5 \mathrm{mM} \mathrm{MgCl}_{2}$, and $0.2 \mathrm{mM} \mathrm{dNTP}, 15 \mathrm{pmol}$ of each primer, and 1 unit of Taq DNA polymerase (Boehringer Mannheim, Germany). Reactions were run in a Perkin Elmer Cetus DNA Thermal Cycler $480\left(1 \mathrm{~min} 95^{\circ} \mathrm{C}\right.$, 35 cycles of $95^{\circ} \mathrm{C} 50 \mathrm{sec}, 54^{\circ} \mathrm{C} 30 \mathrm{sec}, 72^{\circ} \mathrm{C} 90 \mathrm{sec}$, $5 \min 72^{\circ} \mathrm{C}$ ). Primers for PCR were specific to microsporidian sequences in the small subunit (SSU) and large subunit (LSU) ribosomal DNA (rDNA) enclosing the internal transcribed spacer (ITS). The forward primer, located in the $5^{\prime}$ half of the SSU rDNA was Seq1f ( $5^{\prime}$-CGT TGT AGT TCT AGC AGT-3') from Docker et al. (1997a) and reverse primer located in the $5^{\prime}$ half of the LSU rDNA was L580 ( $5^{\prime}$-GGT CCG TGT TTC AAG ACG G-3') from Vossbrinck et al. (1987).

PCR products were run in $1.5 \%$ agarose TBE or TAE gels to visualize products. The desired product was gel-excised, freeze-thaw extracted or $\beta$-agarase digested to remove agarose and then sequenced directly or after cloning.

Products for cloning were isolated in $0.8 \%$ agarose, ligated and cloned using the TOPO TA Cloning PCR Version 2.1 (Invitrogen Corp., Carlsbad, CA), and screened in $10 \mu \mathrm{IPCR}$ reactions using Taq DNA Polymerase (Invitrogen Corp., Carlsbad, CA) with standard reagents and screening primers M13-20 and M13 Rev (conditions: $94^{\circ} \mathrm{C}$ for $2 \mathrm{~min}, 34$ cycles of $92{ }^{\circ} \mathrm{C}$ for $45 \mathrm{sec}, 55^{\circ} \mathrm{C}$ for $45 \mathrm{sec} 72^{\circ} \mathrm{C}$ for 1 min 30 sec , followed by $72^{\circ} \mathrm{C}$ for 5 min ). Positive clones from master plates were grown in 3 ml of standard LB culture with 50 mM ampicillin by shaking at 220 rpm at $37^{\circ} \mathrm{C}$ overnight then isolated using the Rapid Plasmid Miniprep System (Gibco BRL, Gaithersburg, MD) following directions of the manufacturer.

Sequencing was performed using BigDye Terminator Version 3.1 on the ABI PRISM 377 DNA Automated Sequencer. Where possible, PCR and sequencing were repeated for several individuals of each species, and several clones or PCR products were sequenced, and the results were examined in both directions for artefacts of sequencing software, or background contamination.

Sequences were easily aligned by eye with other Loma species in Genbank and by using the secondary structure model of Encephalitozoon cuniculi Levaditi, Nicolau and Schoen, 1923 (L07255) in the rDNA structural database. The ITS region was also aligned with the help of Clustal W (Thompson et al., 1994). A single sequence from each species was selected and will be submitted as "type" sequence. Sequences were selected from the same host individual as the morphological type material (histological or TEM) used in the description (below), or where this was not possible, from an individual fish collected at the same time and geographic locality as the morphological type material.

## RESULTS

Loma-like infections from five hosts (Pacific cod, walleye pollock, Pacific tomcod, lingcod and sablefish) had ultrastructural features that placed them in genus Loma as well as features that distinguished all five as separate, new species of Loma. After briefly defining developmental stages, I formally describe these five new Loma species along with observations on ecology, pathology, and inferred developmental sequence.

## Definitions of developmental stages

Merogonial, or vegetative stages were defined here as in Lom and Pekkarinen (1999) and Lom and Nilsen (2003), as those stages having a thin surface coat of glycocalyx (no thicker than about 10 nm ) covering the plasmalemma. These stages must be in direct contact with a thin layer of host cytoplasm and covered with one or two sheets of host endoplasmic reticulum (RER) cisternae, inside which they may divide as uninucleate meronts or multinucleate merogonial plasmodia until the onset of sporogony, the stage leading to sporoblast and spore production.

Meronts were distinguished from merogonial plasmodia by the large nucleus nearly filling the cell, by an evenly round shape, and by their small ( $1.9 \mu \mathrm{~m}$ ) size. Dividing meronts and merogonial plasmodia could be distinguished by having two or more nuclei, by their greater cytoplasmic relative to nuclear area, and by their larger size (up to $5.3 \mu \mathrm{~m}$ wide). However, in many cases merogonial plasmodia could not be distinguished from meronts (or dividing meronts) because of the nature of sectioning, particularly in cases where plasmodia were cylindrical in shape.

Merogonial stages in transition to sporogonial stages were defined here as in Lom and Nilsen (2003) as stages having lost their host RER cisternae covering and having begun to increase their glycocalyx surface coat thickness (e.g. in this study, thickness was 12.6 to 27 nm ). Surface coat thickness was used as an indicator of the relative stage of sporogony in this study. At this point, early signs of parasitophorous vacuole (PV) formation, such as accumulation of vesicles at the surface of the merogonial stage, were often seen. Merogonial stages were seen in all five species ( $\mathrm{n}=9$ to 16 for each species, except $L$. richardi n . sp. where $\mathrm{n}=2$ ) meeting these criteria defined above: thin coat and host RER covering.

Sporogonial stages were defined here as beginning with the onset of PV formation around either a uninucleate meront (to form a sporont) or around a multinucleate merogonial
plasmodium (to form a sporogonial plasmodium). Typical sporogonial stages here had surface coats 22 to 25 nm thick or more. Sporonts or sporogonial plasmodia, by definition must be stages that will divide to produce sporoblasts, which then develop without further division into spores.

Sporoblasts were defined as stages having begun to form round polar filaments (PF) in cross-section, but not yet showing classical spore features, such as an electron-lucent endospore layer. Sporoblasts were enclosed in large, usually tubule-filled PV spaces, and sporoblast contents included vacuoles and various membranes, depending on the stage of development. Sporoblasts were always coated with two dark, thin, barely separated surface coat layers (measured together, about 36 to 46 nm ). Tubules in the PV space were measured and found here to be always roughly 40 nm wide but often significantly thicker (about 90 to 100 nm ) and bulbous at one end, and sometimes 500 to 800 nm long (Fig. 3.54). Tubule arrangement in the PV space was difficult to discern, but perhaps was oriented about poles (Fig. 3.74).

Spores were recognized by having many mature-spore features such as isofilar polar filament coils, exospore and endospore layers, lamellar and vesicular polaroplasts, posterior vacuoles and ribosomes packed into strands of polyribosomes (Fig. 3.75). Mature spores were recognized by having highly developed features listed above, as well as a fairly thick (measured here, at 37 to 58 nm ) electron-lucent inner wall or endospore layer compared to the thinner exospore layer.

## DESCRIPTIONS

## Loma pacificodae $\mathbf{n}$. $\mathbf{s p}$.

(Figs. 3.1-3.27, Tables 3.2-3.4)

Xenoma in gills of Pacific cod, Gadus macrocephalus Tilesius 1810. Xenoma of cellhypertrophy form with enlarged, branched, but not divided host nuclei. Xenomas of two shapes: elongate xenomas in central venous sinus of primary lamellae usually behind efferent blood vessel (Fig. 3.10), $120.1 \pm 14.1 \mu \mathrm{~m}$ wide ( $\mathrm{n}=43$ ), arising from fibroblast connective tissue; and round xenomas at base of secondary lamellae (Fig. 3.9) $47.3 \pm 15.0 \mu \mathrm{~m}$ wide $(\mathrm{n}=15)$ arising from cells of the pillar system. Xenomas or spores also found in gonads (in $>1 / 2$ of cases)(Fig. 3.11), spleen, heart (about $1 / 2$ of cases), gallbladder, and rarely liver and kidney. Xenoma wall
a smooth or interdigitated plasmalemma, covered by thick electron-lucent, granular, amorphous material in one layer mixed with collagen fibers and fibroblast inclusions, 1.5-2.0 $\mu \mathrm{m}$ thick (Figs. 3.12, 3.13). Collagenous material occasionally in middle of xenomas (Fig. 3.14). Developmental stages and spores intermixed, nuclei always unpaired. Developmental stages: uninucleate meronts having thin surface glycocalyx coat in direct contact with host cytoplasm undergoing binary fission within a host RER covering (Fig. 3.15); highly vacuolated merogonial plasmodia (Fig. 3.16); oblong, rounded, sporogonial plasmodia with smooth, thick glycocalyx coat (Fig. 3.17) enclosed within parasitophorous vacuole (PV) of host origin. Parasitophorous vacuole formation (Figs. 3.18, 3.19) before sporogony by coalescence of small, light (empty) 60 - 200 nm membrane-bound vesicles (Fig. 3.20) and by coalescence of tubule-filled vesicles assembled around later stages (Fig. 3.21). Tubules present in PV spaces before sporogony. Tubule-filled vesicles small, $0.203 \pm 0.033 \mu \mathrm{~m}$ wide ( $\mathrm{n}=16$ ), with few tubules $3 \pm 0.6$ per vesicle ( $\mathrm{n}=33$ ). Numerous tubule-filled vesicles in xenoma cytoplasm associated with each: sporoblast $8.0 \pm 2.0(\mathrm{n}=7)$, and spore $3.8 \pm 1.2(\mathrm{n}=10)$. Numerous tubules in PV spaces of sporoblast $19.3 \pm 14.0(n=4)$, and spore $9 \pm 3.1(n=12)$. Spores oval and slightly narrower at anterior end with sub-apically situated anchoring disc, typical lamellar and vesicular polaroplasts, singly-coiled polar filament and exospore with fine ridges on outer surface. Spores, fresh $5.5(4.8-6.0) \mu \mathrm{m}$ long $\mathrm{x} 3.0(2.7-3.2) \mu \mathrm{m}$ wide $(\mathrm{n}=10)$, or frozen $4.6(4.0-5.5)$ $\mu \mathrm{m}$ long $\mathrm{x} 2.5(2.0-3.0) \mu \mathrm{m}$ wide $(\mathrm{n}=30)$, or formalin fixed paraffin-embedded $3.58 \pm 0.15 \mu \mathrm{~m}$ long x $1.94 \pm 0.15 \mu \mathrm{~m}$ wide ( $\mathrm{n}=10$ ), or resin (Spurr's) embedded $3.48 \pm 0.23 \mu \mathrm{~m}$ long $\times 1.95 \pm$ $0.12 \mu \mathrm{~m}$ wide $(\mathrm{n}=12)$. Number of polar filament turns $17.9 \pm 0.9$ (range 16-22) $(\mathrm{n}=14)$. Posterior vacuole small ( $1 / 4$ of spore). Two spores per PV.

## Taxonomic summary

Type host: Pacific cod, Gadus macrocephalus Tilesius 1810.
Type locality: Barkley Sound, Vancouver Island, Canada.
Site of infection: Primarily central venous sinus of gills, in fibroblast connective tissue or lymph vessel, or at base of secondary lamellae. Secondarily in gonads, spleen, heart, gallbladder, liver and kidney.
Prevalence: $32.6 \%(\mathrm{n}=227)$.
Material deposited: Prior to publication, material will be deposited to appropriate national museum collection, and rDNA sequence will be submitted to Genbank.

Etymology: Named after common name of host, Pacific cod.

## Ecology

Overall prevalence estimated from histological sections was $41.3 \%$ (50/121) and from wet mounts 26.0 (40/154) (Table 3.2). Prevalence differences were observed across localities and over time as shown in Fig. 3.1, but did not appear to follow any particular pattern with respect to depth or distance from shore. North of a latitude of about 497.0 prevalence appeared to be lower, for example, as shown in Fig. 3.1 where the May 1997 collection data were broken up at this latitude to show this difference. Infections occurred in both male and female hosts, and in a range of sizes of hosts ( 34 to 61 cm ) with middle-sized to larger fish appearing to be more often and most heavily infected. Intensity of infection ranged from light ( 1 to 4 xenomas per gill arch) to heavy (up to 59 xenomas in 30 primary lamellae $=2.0 \mathrm{xpf})$.

## Pathology

Signs of infection with $L$. pacificodae n . sp. were sometimes seen grossly, upon examination of organs with heavy xenoma loads. Gills were most often infected, and secondarily xenomas or spores were found in gonads, spleen, heart, gallbladder, liver and kidney. Gills with heavy infections appeared pale, and other organs (e.g. liver, spleen or gonads) could be discoloured. In one heavy infection, the Loma-infected egg-filled ovaries were almost completely black, while many other tissues were speckled with black spots, presumably from melanization. Infected gills commonly showed xenomas in the central venous sinus, in fibroblast connective tissue or in the lymph vessel. Sometimes xenomas developed at the base of the secondary lamellae. In histological sections, heavily infected organs showed signs of host defense, most commonly, granuloma formation, and subsequent changes in tissue structure (e.g. secondary lamellar fusion) once a xenoma was cleared. In the most severe cases, $1 / 4$ to $1 / 2$ of the secondary lamellae were fused, and in active infections granulomas could be as common as intact xenomas. A typical granuloma in cod (Fig. 3.24) had recognizable xenoma walls but few remaining recognizable (not partially digested) spores (Figs. 3.25, 3.26). Xenomas in the central venous sinus of primary lamellae virtually always showed signs of host response (sparsely distributed spores, phagocytes engulfing spores, and wall less defined or intact), whereas xenomas in the bases of secondary lamellae tended to show no or few signs of host response (wall sharp and intact, spores densely packed, and no signs of host cells in xenoma) (Table 3.3). Under electron microscopy (Figure 3.2) fibroblasts could be seen infiltrating xenomas and spores were seen in various stages of being engulfed and digested.

## Remarks

This microsporidian from Pacific cod had features grossly and ultrastructurally consistent with placement in the genus Loma Morrison \& Sprague, 1981 (also see Lom and Nilsen, 2003), including: (1) the production of xenomas (host-parasitic complexes originating from infection of one host cell, producing greatly enlarged and changed ultrastructure, and cystlike appearance) in the gills; (2) xenomas of the cell-hypertrophy type with a centrally located host nucleus and a thick amorphous xenoma wall (host-cell coat); (3) all developmental stages and spores intermingled throughout xenomas; (4) parasite nuclei unpaired at all stages; (5) meronts or merogonial plasmodia with simple plasmalemmas coated by a thin glycocalyx coat in direct contact with host cell cytoplasm; (6) merogonial stages developing into plurinucleate plasmodia and sporogony polysporoblastic; (7) parasitophorous vacuoles with host-derived membranes formed around developing spores; and (8) polar filaments with coils arranged in one layer.

Loma pacificodae n . sp . was most similar to L. wallae n . sp. (described below) but could be distinguished by having: more vacuolated meronts or merogonial plasmodia; presence of a smooth, round, thick-coated sporogonial plasmodium stage; sporogony products not associated in chains with two pairs (daughters) remaining together; parasitophorous vacuole (PV) formation by coalescence of small, light vesicles only (rather than also by larger dark vesicles); tubules appearing in PV spaces early (rather than late); wider spores (resin width $t$-test $\mathrm{p}=$ 0.0283 ; but not by Mann-Whitney test $\mathrm{U}=93.5<\operatorname{Ucrit}_{0.05(1) 11,12} 94$; Fig. 3.6); more vesicles associated with sporoblast stages (Tukey test $\mathrm{q}=9.427>$ qcrit $_{0.05,30,5} 4.102$; Fig. 3.4); thinner xenoma walls that are less distinctly separated into layers and sometimes with interdigitated rather than smooth plasmalemma (Table 3.4, Fig. 3.13); two xenoma types, both of these being different in size from those formed by $L$. wallae $\mathrm{n} . \mathrm{sp}$. (cvs type was larger, t -test $\mathrm{p}=0.0014$; and secondary lamellae type was smaller, t-test $p=0.00058$; Fig. 3.2); absence of large balls of dark granular material contributing to PV space; and with low statistical support ( $p>0.1$ ) slightly longer (resin and histological sections) and wider (histological sections) spores. Loma pacificodae n . sp . was also closest to this species in rDNA sequence, but was distinguished by one transition, five transversions, and one indel across 818 alignment positions of SSU, ITS, and LSU rDNA. In other respects (Tables 3.3, 3.4, Figs. 3.2-3.6, 3.8), such as xenoma location, cell type, vesicle sizes, numbers, and tubule numbers, these two microsporidians appeared to be similar or not significantly different.

## Comparison of L. pacificodae n. sp. to other species examined in this study

While $L$. pacificodae n . sp. was similar in many ways to $L$. wallae n . sp . (described below), it also shared many features with other species examined in this study. For simplicity, feature comparisons across species are summarized in Tables 3.3 \& 3.4 and Figures 3.2 to 3.8, and further details such as statistical test results are given below.

Loma pacificodae n . sp. was similar to L. lingcodae n . sp. (described below), but could be distinguished by having: different xenoma location, cell type, shape and size (Table 3.3); different wall features (Table 3.4); presence of small (empty) vesicles that coalesce to form PV spaces; PV space formation earlier with respect to surface coat formation; presence of sporogonial plasmodia; two (rather than four) spores per PV space; longer and wider spores (converted lengths Tukey test $\mathrm{q}=11.357>$ qcrito $0.05, \infty, 7.17$ for; widths Tukey test $\mathrm{q}=8.050>$ qcrit $_{0.05, \infty, 7} 4.17$; Fig. 3.7); more polar filament turns (Tukey test $\mathrm{q}=5.704>$ qcrit $_{0.05,30,5} 4.102$; but not by Nemenyi test; Fig. 3.8); fewer tubules per vesicle (by Nemenyi test $\mathrm{Q}=4.263$ > Qcrito.05,4 2.639 ; but not by Tukey test; Fig. 3.3); and by nine transitions, 21 transversions, and eight indels over 820 alignment positions in SSU, ITS, and LSU rDNA.

Loma pacificodae n . sp. differentiated from $L$. kenti n . sp. (described below) in having: later PV formation; more vacuolated meronts; different vesicles contributing to PV space formation, these vesicles being smaller and more numerous (Fig. 3.3); presence of tubules in vesicles and PV spaces; longer and wider spores (converted lengths Tukey test $\mathrm{q}=10.856>$ qcrit $_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=6.764>$ qcrit $_{0.05, \infty, 7} 4.17$; Figs. 3.6, 3.7); more polar filament turns (Tukey test $\mathrm{q}=4.575>$ qcrit $_{0.0530,5} 4.102$; Fig. 3.8); and by five transitions, 18 transversions, and two indels over 820 alignment positions of SSU, ITS, and LSU rDNA.

Loma pacificodae n. sp. differentiated from L. richardi n . sp. (described below)(Tables $3.3,3.4$ ) in having: different xenoma size, shape, cell type, location, and wall features; as well as smaller vesicles (Tukey test $q=7.929>$ qcrit $0.05,60,53.977$ ); fewer tubules per vesicle (Tukey test $\mathrm{q}=12.548>$ qcrito.05,60,4 3.737 ; Fig. 3.3); more vesicles per sporoblast (Tukey test $\mathrm{q}=5.582>$ qcrito.05,30,5 4.102; but not Nemenyi test; Fig. 3.4), fewer tubules per sporoblast and spore PV (see Fig. 3.23)(for sporoblasts Tukey test $\mathrm{q}=7.544>$ qcrit $0.05,18,43$ 3.997; for spores Tukey test q $=17.021>$ qcrit $_{0.05,40,4} 3.791$; Fig. 3.5); shorter spores (converted lengths Tukey test $\mathrm{q}=4.343>$ qcrit $0.05, \infty, 7$.17; but not by Nemenyi test; Figs. 3.6, 3.7); more polar filament turns (Tukey test $q$ $=8.699>$ qcrit $_{0.05,30,5} 4.102$; Fig. 3.8); and by 7 transitions, 21 transversions and seven indels over 821 alignment positions of SSU, ITS, and LSU rDNA.

Loma pacificodae n. sp. spores were similar to spores collected from Atlantic cod and haddock gills at a Halifax fish market resembling L. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981. Loma pacificodae n . sp. spores were shorter and narrower than spores from Atlantic cod (converted lengths Tukey test $\mathrm{q}=8.833>$ qcrit $_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=$ $10.040>$ qcrit $_{0.05, \infty, 7} 4.17$, but widths not significantly different by Nemenyi test). Loma pacificodae n . sp. spores were shorter and narrower than spores from haddock (converted lengths Tukey test $\mathrm{q}=5.905>$ qcrit $_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=8.954>$ qcrit $_{0.05, \infty, 7} 4.17 ;$ but neither lengths nor widths significantly different by Nemenyi test; Fig. 3.7). Loma pacificodae n . sp. also differed from these two microsporidians in rDNA sequence by having: one transition, six transversions and two indels in 825 alignment positions compared to spore DNA from Atlantic cod; and three transitions, 16 transversions, and one indel in 819 alignment positions compared to spore rDNA from haddock.

## Comparison of L. pacificodae n. sp. to previously described Loma species

Because genus Loma may be polyphyletic (Lom \& Nilsen, 2003) and questions have been raised about the validity of several key species (e.g. L. branchialis and L. morhua considered synonyms by Canning \& Lom, 1986) I present comparisons between L. pacificodae n. sp. and all previously described species of Loma below. A summary table showing several of the major differences between this species, other new Loma species described below, and previously described Loma species is given in Appendix 11.

Loma pacificodae n . sp. was more similar to L. branchialis than to any other described Loma species. It could be distinguished from L. branchialis in having: different host species and geographic location; less dense material filling PV spaces; smaller xenomas (particularly compared to L. branchialis xenomas from haddock); probably smaller spores (particularly where L. branchialis spore sizes and fixation technique were taken into account); less invaginated walls (particularly compared to L. branchialis from haddock); spores always in pairs rather than singly; and an absence of occasional double-sized, binucleate spores. These species were similar in that both species occurred in hosts of genus Gadus, and were similar in having: two types of xenomas or similar shape, cell type, and location; similarly invaginated walls; similar host response; and similar polar filament turn number.

Loma pacificodae n. sp. was similar to L. salmonae (Putz, Hoffman \& Dunbar, 1965) Morrison \& Sprague, 1981 and L. embiotocia Shaw, Kent, Docker, Brown, Devlin \& Adamson, 1997 in that these species shared geographic localities (were sympatric), were similar in spore
size and some developmental features; however, they differed in host order (Gadiformes versus Salmoniformes and Perciformes). Compared to L. salmonae, L. pacificodae n. sp. had polar filaments appearing later during sporogony; fewer merogonial nuclei; smaller xenomas; different cell type forming xenoma; a marine rather than freshwater habitat; and perhaps more polar filament turns. Compared to L. embiotocia, L. pacificodae n. sp. had tubules arising in empty vesicles of host origin rather than from within granular substance of host origin, and perhaps larger spores. Loma pacificodae and L. salmonae also differed by seven transitions, 21 transversions, and five indels across 819 alignment positions of SSU, ITS, and rDNA.

Loma pacificodae n . sp. could be distinguished from L. fontinalis Morrison \& Sprague, 1983, L. mugili Ovcharenko, Sarabeev, Wita \& Czaplińska, 2000, L. diplodae Bekhti \& Bouix, 1985, Loma sp. of Bekhti, 1984, and L. trichiuri Sandeep \& Kalvati, 1985 by host order, geographic locality, and by having: sporogonic stages and sporoblasts within a PV space, and smaller xenoma size compared with L. fontinalis and L. mugili; different location in the gills and different sporogonic vacuole formation compared with L. diplodae; different location in the gills and different wall compared with Loma sp. of Bekhti; and different spore shape, fewer merogonial nuclei and thinner wall compared with L. trichiuri.

Loma pacificodae n. sp. differed from L. acerinae (Jírovec, 1930) Lom \& Pekkarinen, 1999, L. dimorpha Loubès, Maurand, Gasc, De Buron \& Barral, 1984, L. boopsi Faye, Toguebaye \& Bouix, 1995, and L. myrophis Azevedo \& Matos, 2002, L. camerounensis Fomena, Coste \& Bouix, 1992 in having: xenomas primarily in the gills (rather than primarily in the gut); different host order; different geographic locality; and other developmental and ultrastructural features. Loma pacificodae n . sp. also differed from the latter four species (but not $L$. acerinae) and L. diplodae in having: membranes of host origin (or "true" parasitophorous vacuoles) around sporogonic stages formed by coalescence of vesicles rather than parasitederived sporophorous vesicles resulting from blistering or delamination of material from the sporont surface.

## Loma wallae n. sp.

(Figs. 3.2-3.8, 3.26-3.36, Tables 3.2-3.4)

Xenoma in central venous sinus of gills of walleye pollock, Theragra chalcogramma (Pallas 1811). Xenoma of cell-hypertrophy form with enlarged, branched, but not divided host
nuclei. Xenomas arise from fibroblast connective tissue. Xenomas rounded ovals (Fig. 3.28), often septate or divided (Fig. 3.29), and $82.9 \pm 16.9 \mu \mathrm{~m}$ wide $(\mathrm{n}=25)$. Xenomas or spores also found in gallbladder ( 5 of 6 cases), gonads (in $>1 / 2$ of cases), spleen, heart (about $1 / 2$ of cases), and rarely liver and kidney. Xenoma wall a smooth plasmalemma covered by thick electronlucent, granular, amorphous material in three layers with collagenous fibers and fibroblast inclusions, $2.5 \mu \mathrm{~m}$ thick (Fig. 3.30). Developmental stages and spores intermixed, nuclei always unpaired. Developmental stages: uninucleate meronts with thin glycocalyx coat without host RER covering (or covering is disappearing), with densely packed cytoplasmic contents undergoing binary fission (Figs. 3.31, 3.32) or cylindrical merogonial plasmodia dividing in chains, daughters remaining together (Figs. 3.33, 3.34); sporoblasts within parasitophorous vacuoles (PV) of host origin in chains of four, in pairs of two closely associated PV spaces (Figs. 3.33, 3.34). Parasitophorous vacuole formation before sporogony by coalescence of small, light (empty) 60-200 nm membrane-bound vesicles and large 400-600 nm dark balls of granular material (Fig. 3.35) and by coalescence of tubule-filled vesicles around later stages. Tubules not present in PV spaces until appearance of highly vacuolated sporoblasts (Fig. 3.36). Tubule-filled vesicles small, usually arranged in a line, $0.260 \pm 0.046 \mu \mathrm{~m}$ wide ( $\mathrm{n}=10$ ), with few tubules $4.0 \pm 1.0$ per vesicle $(n=23)$. Few tubule-filled vesicles in xenoma cytoplasm associated with each: sporoblast $1.9 \pm 0.8(\mathrm{n}=9)$, and spore $3.6 \pm 1.7(\mathrm{n}=14)$. Few tubules in PV spaces of sporoblast $8.0 \pm 3.8(n=7)$, and spore $5.9 \pm 2.5(n=10)$. Spores oval and slightly narrower at anterior end with sub-apically situated anchoring disc, typical lamellar and vesicular polaroplasts, singly-coiled polar filament and exospore with fine ridges on outer surface. Spores fixed in formalin and paraffin-embedded $3.47 \pm 0.15 \mu \mathrm{~m}$ long x $1.93 \pm 0.09 \mu \mathrm{~m}$ wide ( $\mathrm{n}=10$ ), or fixed in glutaraldehyde and resin (Spurr's) embedded $3.37 \pm 0.14 \mu \mathrm{~m}$ long x $1.81 \pm 0.05 \mu \mathrm{~m}$ wide $(\mathrm{n}=11)$. Number of polar filament turns $18.6 \pm 1.8$ (range 16-21) ( $\mathrm{n}=$ 5). Posterior vacuole size small (1/4-1/3 of spore). Two spores per PV, in pairs (Fig. 3.36).

## Taxonomic summary

Type host: walleye pollock (= Alaska pollock), Theragra chalcogramma (Pallas 1811).
Type locality: Barkley Sound, Vancouver Island, Canada.
Site of infection: Central venous sinus of gills, in fibroblast connective tissue. Secondarily in gonads, spleen, heart, liver, kidney and especially gallbladder.

Prevalence: 28.3\% ( $\mathrm{n}=145$ ).

Material deposited: Prior to publication, material will be deposited to appropriate national museum collection, and rDNA sequence will be submitted to Genbank.

Etymology: Named after common name of type host, walleye pollock.

## Ecology

Prevalence estimated from histological sections was $35.6 \%$ (16/45) and from wet mounts 24.4 (32/131) (Table 3.2). Infections occurred in both male and female hosts, and in a range of sizes of hosts ( 12 to 48 cm ). Intensity of infection ranged from light ( 1 to 4 xenomas per gill arch) to heavy (up to 25 xenomas per 37 primary lamellae $=0.68 \mathrm{xpf}$ ).

## Pathology

Xenomas appeared to develop primarily in the central venous sinus of gills, in fibroblast connective tissue, and secondarily in gonads, spleen, heart, liver, kidney and especially gallbladder. Signs of host immune response to $L$. wallae n . sp. infection were similar to those described for $L$. pacificodae n . sp. (described above), but were generally less severe.

## Remarks

This microsporidian from walleye pollock had features consistent with placement in the genus Loma Morrison \& Sprague, 1981 similar to those listed for L. pacificodae n. sp. (see remarks for $L$. pacificodae n . sp . above).

Loma wallae n . sp. was similar to L. pacificodae n . sp. (described above) but could be distinguished by having: denser merogonial stages; an absence of rounded sporogonial stages; products of sporogony remaining together in chains; later appearance of tubules in PV space; fewer vesicles per sporoblast; only one type of xenoma that was different in size; presence of balls of dark granular material contributing to PV spaces; perhaps smaller spore sizes (see $L$. pacificodae n . sp. description and remarks above for statistical support) and rDNA sequence.

## Comparison of $L$. wallae n. sp. to other species examined in this study

While $L$. wallae n . sp. and $L$. pacificodae n . sp . were most similar to one another, they were similar to other species examined in this study. As before, for clarity, the feature comparisons across species are summarized in Tables $3.3 \& 3.4$ and Figures 3.2 to 3.8, and further details such as statistical test results are given below.

Loma wallae n . sp . could be distinguished from $L$. kenti n . sp. (described below) by many of the same features shared with L. pacificodae n . sp . (described above), for example, by having: later PV formation; different vesicles contributing to PV space formation, these vesicles being smaller and more numerous (Fig. 3.3); presence of tubules in vesicles and PV spaces;
longer and wider spores (converted lengths Tukey test $\mathrm{q}=8.752>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=5.035>$ qcrit $_{0.05, \infty, 7} 4.17$, but width difference not significant by Nemenyi test; Figs. 3.6, 3.7); more polar filament turns (Tukey test $q=6.893>$ qcrit $_{0.0530,5} 4.102$; Fig. 3.8); and by five transitions, 16 transversions, and two indels across 818 alignment positions of SSU, ITS, and LSU rDNA.

Similarly, $L$. wallae n . sp. differed from L. lingcodae n . sp. (described below) in features shared with L. pacificodae n . sp., for example by having: different xenoma location, cell type, shape and size (Table 3.3); different wall features (Table 3.4); presence of small (empty) vesicles that coalesce to form PV spaces; PV space formation earlier with respect to surface coat formation; two (rather than four) spores per PV space; fewer vesicles per sporoblast (Tukey test $\mathrm{q}=8.429>\mathrm{qcrit}_{0.05,30,5} 4.102$ ); longer and wider spores (converted lengths Tukey test $\mathrm{q}=9.186$ $>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=6.255>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$, but width difference not significant by Nemenyi test; Fig. 3.7); more polar filament turns (Tukey test $q=5.792>$ qcrit ${ }_{0.05,30,5} 4.102$; but not by Nemenyi test; Fig. 3.8); and by eight transitions, 19 transversions, and eight indels across 819 alignment positions of SSU, ITS, and LSU rDNA.

Differences between $L$. wallae n . sp. and $L$. richardi n . sp . (described below) were similar to differences described for L. pacificodae n . sp. (above), such as different xenoma size, shape, cell type, location, and wall features (Tables 3.3,3.4); as well as smaller vesicles (Tukey test $\mathrm{q}=6.083>$ qcrit $_{0.05,60,5} 3.977$ ); fewer tubules per vesicle (Tukey test $\mathrm{q}=10.667>$ qcrit $_{0.05,60,4} 3.737$; Fig. 3.3); fewer tubules per PV for sporoblasts Tukey test $q=9.859>$ qcrit $_{0.05,18,4} 3.997$, and for spores Tukey test $q=17.287>$ qcrit $_{0.05,40,4} 3.791$ (Fig. 3.5); shorter spores (converted lengths Tukey test $q=5.845>q^{c r i t}{ }_{0.05, \infty, 7} 4.17$, but not by Nemenyi test; Figs. 3.6, 3.7); more polar filament turns (Tukey test $q=8.205>$ qcrit $_{0.05,30,5} 4.102$; Fig. 3.8); and by seven transitions, 17 transversions, and seven indels across 819 alignment positions of SSU, ITS and LSU rDNA.

Loma wallae n. sp. spores were similar to spores collected from Atlantic cod and haddock gills at a Halifax fish market resembling L. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981. Loma wallae n. sp. spores were shorter and narrower than spores from Atlantic cod (converted lengths Tukey test $\mathrm{q}=10.698>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; converted widths Tukey test $\mathrm{q}=$ $11.564>$ qcrit $_{0.05, \infty, 7} 4.17$ ); and shorter and narrower than spores from haddock (converted lengths Tukey test $q=7.955>$ qcrit $_{0.05, \infty, 7} 4.17$; converted widths Tukey test $q=10.608>$ qcrit $_{0.05, \infty, 7} 4.17$; Fig. 3.7). Ribosomal DNA differences between $L$. wallae n . sp. and spores
from these hosts were: six transversions and one indel in 823 alignment positions compared to spores from Atlantic cod; and two transitions and 13 transversions in 817 alignment positions compared to spores from haddock.

## Comparison of $\boldsymbol{L}$. wallae n. sp. to previously described Loma species

Loma wallae n. sp. could be distinguished from described species of Loma by differences listed for L. pacificodae n . sp . (described above; see remarks for L. pacificodae n . sp.) except that: compared to $L$. branchialis, $L$. wallae n . sp. had only one xenoma type; smooth rather than highly invaginated xenoma wall; and different host genus (same host family).

## Loma kenti n. sp.

(Figs. 3.2-3.8, 3.37-3.49, Tables 3.2-3.4)

Xenoma at base of secondary lamellae, throughout primary lamellae in gills of Pacific tomcod, Microgadus proximus (Girard 1854). Xenoma of cell-hypertrophy form with enlarged, branched, but not divided host nuclei. Xenomas arise from endothelial cells or blood vessels. Xenomas round (Fig. 3.37) and $142.3 \pm 24.0 \mu \mathrm{~m}$ wide ( $\mathrm{n}=8$ ). Xenomas or spores also found in gonads (in $>1 / 2$ of cases), spleen, heart (about $1 / 2$ of cases), gallbladder, and rarely liver and kidney. Xenoma wall an undulating plasmalemma ( $0.5-1 \mu \mathrm{~m}$ between peaks) covered by thick electron-lucent, granular, amorphous material in three layers (innermost darkest, outermost lightest) interspersed with collagen fibers, 1.3-3.6 $\mu \mathrm{m}$ thick (Figs. 3.38, 3.39). 600 nm vesicle inclusions in wall. Developmental stages and spores intermixed, nuclei always unpaired. Developmental stages: plurinucleate (three or more nuclei) merogonial plasmodia (nuclei in cluster rather than in a chain), with densely packed cytoplasmic contents (Figs. 3.40-3.43), with thin, patchily distributed surface glycocalyx coats with beginnings of parasitophorous vacuole (PV) formation (Figs. 3.41, 3.42), or with fully formed large PV spaces of host origin (Fig. 3.43), usually associated in pairs, and not covered in host RER (or covering is disappearing); oblong or round, sporogonial plasmodia with smooth, thick glycocalyx coats (Figs. 3.46, 3.47) dividing by multiple fission (a "clover-leaf" Fig. 3.46). Parasitophorous vacuole formation at merogonial stages by coalescence or emptying of vesicles containing dark balls of granular material within a vacuolar space (Figs. 3.38, 3.42, 3.43, 3.48). Dark material-filled vesicles, sometimes with tubule-like structures visible within the granular dark material (Fig. 3.48) large $0.715 \pm 0.212 \mu \mathrm{~m}$ wide $(\mathrm{n}=15$ ). No tubule-filled vesicles. No tubules in PV space (or rarely).

Spores oval and slightly narrower at anterior end with sub-apically situated anchoring disc, typical lamellar and vesicular polaroplasts, singly-coiled polar filament and exospore with fine ridges on outer surface. Spores fixed in formalin and paraffin-embedded $2.97 \pm 0.13 \mu \mathrm{~m}$ long x $1.75 \pm 0.27 \mu \mathrm{~m}$ wide ( $\mathrm{n}=12$ ), or fixed in glutaraldehyde and resin (Spurr's) embedded $2.94 \pm$ $0.11 \mu \mathrm{~m}$ long $\mathrm{x} 1.52 \pm 0.05 \mu \mathrm{~m}$ wide $(\mathrm{n}=12$ ). Number of polar filament turns $14.8 \pm 0.4$ (range 14-16) $(\mathrm{n}=12)$. Posterior vacuole size large (just $<1 / 2$ of spore). One or rarely two spores per PV (Fig. 3.49).

## Taxonomic summary

Type host: Pacific tomcod, Microgadus proximus (Girard 1854).
Type locality: Barkley Sound, Vancouver Island, Canada.
Site of infection: Base of secondary lamellae and blood vessels of gills, in endothelial cells.
Secondarily in gonads, spleen, heart, gallbladder, liver and kidney.
Prevalence: $14.4 \%$ ( $\mathrm{n}=419$ ).
Material deposited: Prior to publication, material will be deposited to appropriate national museum collection, and rDNA sequence will be submitted to Genbank.
Etymology: Named after fish pathologist and parasitologist, Dr. Michael L. Kent.

## Ecology

Prevalence estimated from histological sections was $5.3 \%$ (3/57) and from wet mounts 14.4 (60/409)(Table 3.2). Infections occurred in both male and female hosts, and in a range of sizes of hosts ( 12 to 14 cm ). Intensity of infection ranged from light ( 1 to 4 xenomas per gill arch) to moderately heavy (up to 9 xenomas per 70 primary lamellae $=0.13 \mathrm{xpf}$ ).

## Pathology

Xenomas developed primarily in the gills at the base of secondary lamellae and blood vessels, in endothelial cells, and spores or xenomas were also found sometimes in the gonads, spleen, heart, gallbladder, liver and kidney. Signs of host immune response to $L$. kenti n . sp. infection were rarely seen, but were similar to those described for L. pacificodae n . sp . (above), but less severe.

## Remarks

This microsporidian from Pacific tomcod had features consistent with placement in the genus Loma Morrison \& Sprague, 1981 similar to those listed for L. pacificodae n. sp. (see remarks for $L$. pacificodae n . sp. above).

Loma kenti n . sp. was distinguished from L. pacificodae n. sp. (described above) by having earlier parasitophorous vacuole (PV) formation; less vacuolated meronts; different
vesicles contributing to PV space formation, these vesicles being larger and having dark material suspended in a vacuole (not tubules) and being less numerous; no tubules in vesicles or PV spaces; shorter and narrower spores; fewer polar filament turns; and differences in rDNA (see $L$. pacificodae n . sp. description and remarks above for statistical support). Similarly, L. kentin. sp. differed from $L$. wallae n . sp. (described above) by features the latter species shared with $L$. pacificodae n . sp.; as well as by five transitions, 16 transversions, and two indels across 818 alignment positions of SSU, ITS, and LSU rDNA (see $L$. wallae n. sp. description and remarks above for stàtistical support).

## Comparison of $\boldsymbol{L}$. kenti n . sp. to other species examined in this study

Because L. kentin. sp. shared a number of features with other species examined in this study, feature comparisons across species are summarized in Tables $3.3 \& 3.4$ and Figures 3.2 to 3.8, and further details such as statistical test results are given below.

Loma kenti n. sp. was similar in spore size (Figs. 3.6,3.7) to L. lingcodae n. sp. (described below), but could be distinguished by having: earlier PV formation while merogonial surface coat is thin; different vesicles contributing to PV space formation with dark material (rather than tubules) suspended in a vacuole; no tubules in either vesicles or PV spaces; larger vesicle size (Tukey test $\mathrm{q}=7.798>\mathrm{qcrit}_{0.05,60,5} 3.977$; Fig. 3.3); fewer vesicles per sporoblast (Tukey test $q=5.6>$ qcrit $_{0.05,30,5} 4.102$; but not by Nemenyi test); fewer vesicles per spore (Tukey test $\mathrm{q}=4.084>$ qcrit $_{0.05,40,5} 4.039$; Fig. 3.4); and by 12 transitions, nine transversions, and eight indels across 813 alignment positions of SSU, ITS, and LSU rDNA.

Loma kenti n . sp. could be distinguished from L. richardi n . sp. (described below) in many of the same ways described for the species from lingcod, for example, in having: earlier PV formation; different vesicles; no tubules; as well as by having shorter and narrower spores (converted lengths Tukey test $\mathrm{q}=12.914>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=5.151>$ qcrit $_{0.05, \infty, 7} 4.17$, but width difference not significant by Nemenyi test; Figs. 3.6, 3.7); and by 11 transitions, eight transversions, and seven indels across 812 alignment positions of SSU, ITS, and LSU rDNA.

Loma kenti n . sp. spores were smaller (Figs. 3.6, 3.7) than spores collected from Atlantic cod and haddock gills at a Halifax fish marked resembling L. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981. Compared to spores from Atlantic cod, L. kentin. sp. spores were shorter and narrower (converted lengths Tukey test $\mathrm{q}=20.132>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=17.182>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$ ), and similarly, they were shorter and narrower compared to
spores from haddock (converted lengths Tukey test $\mathrm{q}=18.152>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; lengths Tukey test $\mathrm{q}=16.726>$ qcrit $_{0.05, \infty, 7} 4.17$; Fig. 3.7).

## Comparison of $L$. kenti n. sp. to previously described Loma species

Loma kenti n . sp. could be distinguished from described species of Loma by features listed for L. pacificodae n. sp. (described above; see remarks for L. pacificodae n. sp.) but with the following differences: compared to all described species L. kenti n . sp. had different vesicles contributing to PV spaces; earlier appearance of PV spaces; and compared to L. branchialis, $L$. kenti n . sp. had only one xenoma type; a simple, undulating (rather than highly invaginated) wall; different host genus (same host family); whereas compared to $L$. salmonae and $L$. embiotocia, L. kenti n . sp. had differences shared with L. pacificodae n . sp . but more similar polar filament turn number.

## Loma lingcodae n. sp.

(Figs. 3.2-3.8, 3.50-3.58, Tables 3.2-3.4)

Xenoma in secondary lamellae at tips of primary lamellae in gills of lingcod Ophiodon elongatus Girard 1854. Xenoma of cell-hypertrophy form with enlarged, branched, but not divided host nuclei. Xenomas arise from endothelial cells or pillar system. Xenomas round or oval (Fig. 3.50) and $37.4 \pm 3.9 \mu \mathrm{~m}$ wide ( $\mathrm{n}=56$ ). Xenomas or spores also found in gonads (in 7 of 8 cases), spleen, heart (about $1 / 2$ of cases), gallbladder, and rarely liver and kidney. Gonads primary site in two cases. Xenoma wall a finely undulating plasmalemma (200-400 nm between peaks) covered by thick electron-lucent, granular, amorphous material in one layer, $0.34-0.5 \mu \mathrm{~m}$ thick (Figs. 3.51, 3.52). 50 nm vesicle inclusions in wall. Developmental stages and spores intermixed, nuclei always unpaired. Developmental stages: meronts and merogonial plasmodia with loosely packed cytoplasmic contents and host RER covering (Figs. 3.53, 3.54) and with three or more spindle plaques (Fig. 3.55), with thin or thick (up to 38 nm ) surface glycocalyx coats (perhaps sporogonial plasmodia) in direct contact with host cytoplasm rather than in a parasitophorous vacuole (PV). Dark granular material in balls accumulated around merogonial stages (Fig. 3.56). Parasitophorous vacuole formation in late sporogony before or as sporoblasts form, by coalescence of tubule-filled vesicles. Tubule-filled vesicles small, $0.277 \pm$ $0.027 \mu$ m wide $(\mathrm{n}=19)$, with few tubules $6.1 \pm 0.7$ per vesicle $(\mathrm{n}=28)$. Numerous tubule-filled vesicles in the host cytoplasm associated with each: sporoblast $7.0 \pm 1.3(n=9)$, and spore $4.8 \pm$
$1.0(\mathrm{n}=10)$. Numerous tubules in PV spaces of sporoblast $17 \pm 6.4(\mathrm{n}=7)$, and spore $11.8 \pm$ $2.9(\mathrm{n}=23)$. Spores oval and slightly narrower at anterior end with sub-apically situated anchoring disk, typical lamellar and vesicular polaroplasts, singly-coiled polar filament and exospore with fine ridges on outer surface. Spores fresh $4.6(3.8-5.4) \mu \mathrm{m}$ long $\times 2.8(2.5-3.0)$ $\mu \mathrm{m}$ wide ( $\mathrm{n}=10$ ), or frozen $4.8(4.0-5.0) \mu \mathrm{m}$ long $\mathrm{x} 2.1(2.0-2.5) \mu \mathrm{m}$ wide $(\mathrm{n}=30)$, or fixed in formalin and paraffin-embedded $2.91 \pm 0.07 \mu \mathrm{~m}$ long $\times 1.55 \pm 0.06 \mu \mathrm{~m}$ wide ( $\mathrm{n}=17$ ), or fixed in glutaraldehyde and resin (Spurr's) embedded $3.00 \pm 0.12 \mu \mathrm{~m}$ long x $1.65 \pm 0.09 \mu \mathrm{~m}$ wide $(\mathrm{n}=10)$. Number of polar filament turns $15 \pm 0.7$ (range $14-16)(\mathrm{n}=6)$. Posterior vacuole size large (just $<1 / 2$ of spore). Four spores per PV (Fig. 3.57).

## Taxonomic summary

## Type host: lingcod Ophiodon elongatus Girard 1854.

Type locality: Barkley Sound, Vancouver Island, Canada.
Site of infection: Secondary lamellae of gills, primarily at tips of primary filaments, in endothelial cells. Secondarily in spleen, heart, liver, gallbladder, kidney and especially gonads. Prevalence: 21.4\% ( $\mathrm{n}=210$ ).
Material deposited: Prior to publication, material will be deposited to appropriate national museum collection, and rDNA sequence will be submitted to Genbank.
Etymology: Named after common name of host, lingcod.

## Ecology

Prevalence estimated from histological sections was $16.7 \%$ (26/156) and from wet mounts 21.1 (42/199)(Table 3.2). Infections occurred in both male and female hosts, and in a range of sizes of hosts ( 45 to 82 cm ), with middle-sized to larger fish appearing to be more often and most heavily infected. Intensity ranged from light ( 1 to 4 xenomas per gill arch) to heavy (up to 65 xenomas in 28 primary lamellae $=2.3 \mathrm{xpf}$ )

## Pathology

Xenomas appeared to develop primarily at tips of primary filaments, in endothelial cells of gills, and xenomas or spores were secondarily found in spleen, heart, liver, gallbladder, kidney and especially gonads. Signs of host immune response to $L$. lingcodae n. sp. were rarely seen, but where observed, granulomas resembled those described for L. pacificodae n. sp. (described above).

## Remarks

This microsporidian from lingcod had features consistent with placement in the genus Loma Morrison \& Sprague, 1981 similar to those listed for L. pacificodae n . sp . (see remarks for L. pacificodae n . sp. above).

Loma lingcodae n. sp. was most similar to L. richardin. sp. (described below) but could be distinguished by having: xenoma wall with fine undulations rather than smooth; xenoma wall slightly thicker; slightly larger xenomas (t-test $\mathrm{p}=0.023$; but not by Mann-Whitney test; Fig. 3.2); uninucleate meronts; later formation of PV spaces after surface coat had greatly thickened; no sporogonial stages surrounded by PV spaces; no large dark vesicles contributing to PV space formation; four spores per PV space (rather than two); smaller vesicles (Tukey test $\mathrm{q}=6.921>$ $q^{c r i t} t_{0.05,60,5} 3.977$ ); fewer tubules per vesicle (Tukey test $q=9.116>$ qcrit $_{0.05,60,4} 3.737$; Fig. 3.3); more vesicles per sporoblast (Tukey test $q=4.528>$ qcrit $_{0.05,30,5} 4.102$; but not by Nemenyi test; Fig. 3.4); fewer tubules per spore PV (Tukey test $\mathrm{q}=18.675>$ qcrit $_{0.05,40,4} 3.791$; Fig. 3.5); shorter and narrower spores (resin) (lengths t-test $p=1.05 \times 10^{-7}$; widths t-test $p=0.00302$; Fig. 3.6). Loma lingcodae n . sp. was also most similar to this species in rDNA sequence, but could be differentiated by 3 transitions, one transversion, and three indels across 814 alignment positions of SSU, ITS, and LSU rDNA.

## Comparison of L. lingcodae n. sp. to other Loma species examined in this study

Loma lingcodae n. sp. was also similar to other species examined in this study. Tables $3.3 \& 3.4$ and Figures 3.2 to 3.8 , provide a summary of these comparative differences, and further details such as statistical test results are given below.

Loma lingcodae n . sp. was distinguished from L. pacificodae n . sp. (described above) by having different xenoma location, cell type, shape and size (Table 3.3); different wall features (Table 3.4); lacking small (empty) vesicles that coalesce to form PV spaces; PV space formation later with respect to surface coat formation; lack of sporogonial plasmodia; four (rather than two) spores per PV space; shorter and narrower spores (Figs. 3.6, 3.7); fewer polar filament turns (Fig. 3.8); fewer tubules per vesicle (Fig. 3.3); and rDNA differences (see L. pacificodae n. sp. description and remarks above).

Loma lingcodae n . sp. was distinguished from $L$. wallae n . sp. (described above) by many features similar to those of L. pacificodae n . sp . and from L. wallae n . sp ., as well as more vesicles per sporoblast (Fig. 3.4); and rDNA differences (see L. wallae n. sp. description and remarks above for statistical support). Loma lingcodae n. sp. had spores of similar size as those from L. kenti n . sp. (described above) but differed in having: later PV formation; different vesicles contributing to PV space formation (see difference from L. pacificodae n . sp . above);
presence of tubules in vesicles and PV spaces; smaller vesicles (Fig. 3.3); more vesicles per sporoblast and spore (Fig. 3.4); and rDNA differences (see L. kenti n. sp. description and . remarks above for statistical support).

Loma lingcodae n. sp. spores were smaller (Figs. 3.6, 3.7) than spores collected from Atlantic cod and haddock at a Halifax fish market resembling L. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981. Compared to spores from Atlantic cod, L. lingcodae n. sp. spores were shorter and narrower (converted lengths Tukey test $\mathrm{q}=20.922>$ qcrit $_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=18.804>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$ ), and shorter and narrower than spores from haddock (converted lengths Tukey test $\mathrm{q}=19.008>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; widths Tukey test $\mathrm{q}=18.537>$ qcrit $_{0.05, \infty, 7} 4.17$; Fig. 3.7).
Comparison of L. lingcodae n. sp. to previously described Loma species
Loma lingcodae n. sp. could be distinguished from described species of Loma by features listed for L. pacificodae n . sp. (described above; see remarks for L. pacificodae n . sp .) except that $L$. lingcodae n. sp. was more similar to $L$. salmonae and $L$. embiotocia than to $L$. branchialis. Loma lingcodae n . sp. was similar to L. salmonae and L. embiotocia in xenoma cell type and polar filament turn number, but different in having: four rather than two spores per PV, smaller spores, and a finely undulating wall plasmalemma. In other features, these species differed as listed for L. pacificodae n . sp. (above) and compared to L. salmonae, L. lingcodae n . sp. had four transitions, five transversions, and three indels across 808 alignment positions of SSU, ITS, and LSU rDNA.

Compared to $L$. branchialis, L. lingcodae n. sp. had one xenoma type; a simple, smooth wall with amorphous light contents; different host order; and four rather than two spores per PV; and fewer polar filament turns, but in other ways L. lingcodae n. sp. differed from remaining described species of Loma by the same features listed for L. pacificodae n . sp. (above).

## Loma richardi n. sp.

(Figs. 3.2-3.8, 3.59-3.73, Tables 3.2-3.4)

Xenoma in tips of secondary lamellae throughout primary lamellae in gills of sablefish Anoplopoma fimbria (Pallas 1811). Xenoma of cell-hypertrophy from with enlarged, branched, but not divided host nuclei. Xenomas arise from endothelial cells or pillar system. Xenomas round and $33.1 \pm 1.4 \mu \mathrm{~m}$ wide $(\mathrm{n}=5)$ (Figs. $3.59,3.60$ ). Xenomas or spores also found in
gonads (in $>1 / 2$ of cases), spleen, heart (about $1 / 2$ of cases), gallbladder, and rarely liver and kidney. Xenoma wall a smooth plasmalemma covered by thick electron-lucent, granular, amorphous material in one or two layers, $0.24 \mu \mathrm{~m}$ thick (Figs. $3.61-3.63$ ). 50 nm vesicle inclusions in wall. Developmental stages and spores intermixed, nuclei always unpaired. Developmental stages: merogonial plasmodia with thin, patchily distributed surface glycocalyx coats and loosely packed cytoplasmic contents in direct contact with host cytoplasm, and covered in host RER (Fig. 3.64); sporonts with concentric rings of ER cisternae (Fig. 3.65) and sporogonial plasmodia with thick surface glycocalyx coats within large parasitophorous vacuoles (PV) of host origin, sometimes dividing by binary fission (Fig. 3.66), sometimes at different developmental stages within a single PV space (Figs. 3.67, 3.68). Parasitophorous vacuole formation early, before sporogony by coalescence of small, light (empty), membranebound vesicles (Figs. 3.69-3.71), and by coalescence of tubule-filled vesicles at later stages. Tubule-filled vesicles large, $0.658 \pm 0.133 \mu \mathrm{~m}$ wide $(\mathrm{n}=16$ ), with numerous tubules $15.6 \pm 4.7$ per vesicle $(\mathrm{n}=20)$. Few tubule-filled vesicles in the host cytoplasm associated with each: sporoblast $3.5 \pm 1.3(\mathrm{n}=4)$, and spore $2.5 \pm 1.0(\mathrm{n}=8)$. Great numbers of tubules in PV spaces of sporoblast $90.3 \pm 40.6(\mathrm{n}=4)$, and spore $52 \pm 7.2(\mathrm{n}=16)$. Spores oval and slightly narrower at anterior end with sub-apically situated anchoring disk, typical lamellar and vesicular polaroplasts, singly-coiled polar filament and exospore with fine ridges on outer surface. Spores fixed in glutaraldehyde and resin (Spurr's) embedded $3.81 \pm 0.15 \mu \mathrm{~m}$ long x $1.93 \pm 0.15 \mu \mathrm{~m}$ wide $(\mathrm{n}=10)$. Number of polar filament turns $13.5 \pm 1.2$ (range $11-15)(\mathrm{n}=6)$. Posterior vacuole size large (just <1/2 of spore). Two spores per PV (Fig. 3.72).

## Taxonomic summary

Type host: sablefish (blackcod), Anoplopoma fimbria (Pallas 1811).
Type locality: Barkley Sound, Vancouver Island, Canada.
Site of infection: Secondary lamellae of gills throughout filaments, in endothelial cells.
Secondarily in gonads, spleen, heart, gallbladder, liver and kidney.
Prevalence: 13.2\% ( $\mathrm{n}=197$ ).
Material deposited: Prior to publication, material will be deposited to appropriate national museum collection, and rDNA sequence will be submitted to Genbank.

Etymology: Named after late mathematician, Dr. Richard L. W. Brown.

## Ecology

Prevalence estimated from histological sections was $3.9 \%$ ( $5 / 128$ ) and from wet mounts 13.2 (26/197)(Table 3.2). Infections occurred in both male and female hosts, and in a range of sizes of hosts ( 27 to 43 cm ). Intensity of infection was generally low (at most 8 xenomas per gill arch).

## Pathology

Xenomas developed in the secondary lamellae of gills throughout the filaments, within endothelial cells, and xenomas or spores were found sometimes also in gonads, spleen, heart, gallbladder, liver and kidney. Signs of host immune response to $L$. richardi n . sp . were not seen.

## Remarks

This microsporidian from sablefish had features consistent with placement in the genus Loma Morrison \& Sprague, 1981 similar to those listed for L. pacificodae n . sp . (see remarks for L. pacificodae n. sp. above).

Loma richardi n . sp. was most similar to L. lingcodae n . sp. (described above) but could be distinguished by having: a smooth (rather than undulating) xenoma wall; xenoma wall slightly thinner; slightly smaller xenomas; lack of uninucleate meronts; earlier formation of PV spaces while surface coats of merogonial plasmodia are still thin; presence of sporogonial stages surrounded by PV spaces; presence of large dark vesicles contributing to PV space formation; two spores per PV space (rather than four); larger vesicles and more tubules per vesicle (Fig. 3.3); fewer vesicles per sporoblast (Fig. 3.4); more tubules per spore PV (Fig. 3.5); longer and wider spores in both resin and converted spore measurements (Figs. 3.6, 3.7) (see L. lingcodae n. sp. description and remarks above for statistical support) and different rDNA sequence

## Comparison of L. richardi n. sp. to other Loma species examined in this study

Although L. richardin. sp. was similar to L. lingcodae n. sp. (described above), it was similar in some ways to other species examined in this study, as summarized in Tables 3.3 \& 3.4 and Figures 3.2 to 3.8, and further described below.

Loma richardi n . sp . was distinguished from L. pacificodae n . sp . (described above) by having: different xenoma size, shape, cell type, location, and wall features; larger vesicles and more tubules per vesicle (Fig. 3.3); fewer vesicles per sporoblast (Fig. 3.4); more tubules per sporoblast and spore PV (Fig. 3.5, see Fig. 3.7); longer spores (Figs. 3.6, 3.7); fewer polar filament turns (Fig. 3.8); and different rDNA sequence (see L. pacificodae n. sp. description and remarks above for statistical support).

Loma richardi n . sp. was distinguished from $L$. wallae n . sp . (described above) by having different xenoma size, shape, cell type, location, and wall features (Tables 3.3, 3.4); as well as
larger vesicles and more tubules per vesicle (Fig. 3.3); more tubules per sporoblast and spore PV (Fig. 3.5); longer spores (Figs. 3.6, 3.7); fewer polar filament turns (Fig. 3.8); and different rDNA sequence (see $L$. wallae n . sp. description and remarks above for statistical support).

Loma richardi n . sp. differed from $L$. kenti n . sp . (described above) by having later PV formation; different vesicles; presence of tubules in both vesicles and PV spaces; as well as by having wider and longer spores (Figs. 3.6, 3.7); and different rDNA sequence (see L. kenti n. sp. description and remarks above for statistical support).

Loma richardi n. sp. spores were similar to those from Atlantic cod and haddock at a fish market in Halifax resembling L. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981. Compared to spores from Atlantic cod and haddock, L. richardi n . sp . spores were narrower (converted widths compared to Atlantic cod spores Tukey test $\mathrm{q}=8.028>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; compared to haddock spores Tukey test $\mathrm{q}=6.956>\mathrm{qcrit}_{0.05, \infty, 7} 4.17$; but these differences were not significant by Nemenyi test; Fig. 3.7). Loma richardi n. sp. rDNA sequence revealed eight transitions, 18 transversions, and seven indels across 825 alignment positions compared to spores from Atlantic cod; and 10 transitions, 13 transversions, and six indels over 819 alignment positions compared to spores from haddock.

## Comparison of $L$. richardi n. sp. to previously described Loma species

Loma richardi n . sp. could be distinguished from described species of Loma by features that differed for L. pacificodae n. sp. (described above; see remarks for L. pacificodae n. sp.) with the exception that, like $L$. lingcodae $\mathrm{n} . \mathrm{sp} ., L$. richardi n . sp . was more similar to $L$. salmonae and L. embiotocia than to L. branchialis. Loma richardi n. sp. was similar to $L$. salmonae and L. embiotocia in xenoma cell type and polar filament turn number, but different in other ways listed for L. pacificodae n . sp. Loma richardi n . sp . also differed in rDNA sequence from L. salmonae by: three transitions, four transversions and one indel across 808 alignment positions. Compared to L. branchialis, L. richardi n . sp. had one xenoma type; a simple, smooth wall with amorphous light contents; different host order; and fewer polar filament turns, and in other ways differed from this species in ways listed for L. pacificodae n. sp. (above).

## Inferred developmental sequences of new Loma species

Below, I present inferred developmental sequences to help summarize and extend the understanding of differences observed among these new Loma species; however, such differences were inferred from only a few individuals or xenomas (seven xenomas in four individuals from Pacific cod, three xenomas in two individuals from walleye pollock and
sablefish, two xenomas in one individual from Pacific tomcod and lingcod) and therefore some stages may have been missed. As in many studies of microsporidia, earlier stages (early merogonial phases) were likely missed in this material, due to the bias from study of more advanced infections in visibly xenoma-filled tissues. I may have also missed stages that are rare, and therefore these developmental sequences should be regarded as a beginning, rather than a final analysis of the developmental process in these five species.

## Development of Loma pacificodae n. sp.

Merogony proceeds as meronts divide by binary fission within a host RER covering (Fig. 3.15) to form merogonial plasmodia. Meronts and merogonial plasmodia are highly vacuolated (compared with stages in L. wallae n. sp. and other species described below) (Fig. 3.16). Sometimes small, double-membrane-bound vesicles, about 76.5-88.2 nm wide with amorphous dark contents were observed in the merogonial cytoplasm (Fig. 3.27). These small structures resemble mitochondrial homologs observed by Williams et al. (2002). Stages become increasingly ,vacuolated as host RER disappears and surface coat thickens. These stages remain highly vacuolated as parasitophorous vacuoles (PV) form (Figs. 3.18, 3.19) by coalescence of small (60-200 nm) membrane-bound vesicles with light (or empty) contents, which have gathered around merogonial plasmodia (Fig. 3.20). Merogonial plasmodia become sporogonial plasmodia as PV spaces grow, their contents become less dense, and surface coats thicken. Sporogonial plasmodia grow in size, lose PV space thickness and develop a smooth surface and rounded oblong shape, in contrast to the rippled surface of earlier stages (Fig. 3.17). A similar stage was observed in L. kenti n . sp. but not in L. wallae n . sp., L. lingcodae n . sp . or L. richardi n . sp. An invagination in the surface of this stage into which tubules seem enter is comparable to stages seen by Lom \& Pekkarinen (1999; Fig. 25), and may suggest exchange of material from xenoma cytoplasm to parasite surface. Small, tubule-filled vesicles continue to assemble around these stages and coalesce to increase the PV space (Fig. 3.21), until a final division producing two sporoblasts, which grow into spores (two per PV).

## Development of Loma wallae n. sp.

Uninucleate meronts (or cross-sections through cylindrical multinucleate merogonial plasmodia) with contents more densely packed with ribosomes than those in L. pacificodae n . sp., without RER covering (or with covering in the process of disappearing) undergo binary fission (Fig. 3.31) and daughter cells remain together (Figs. 3.32, 3.33) in chains. Large (400600 nm ), balls of amorphous, dark material in direct contact with the cytoplasm or within membrane bound vesicles (Fig. 3.35) as well as some smaller, light vesicles, like those described
above in L. pacificodae n. sp., accumulate at the surface of merogonial stages and empty their contents onto the surface of cells to form parasitophorous vacuoles (PV). No tubules were seen in PV spaces until the products of sporogonial plasmodial fission developed into highly vacuolated early sporoblasts (Fig. 3.36). Small vesicles containing small numbers of tubules, often arranged in a line, later coalesce to add volume to PV spaces of sporoblasts and spores. Late sporoblasts and spores were always seen as chains of four, in two closely associated PV spaces, each containing two cells (Fig. 3.36).

## Development of Loma kentin. sp.

Multinucleate merogonial plasmodia with dense contents were often closely associated in pairs (suggesting prior binary fission). Plasmodia have three nuclei (or $>$ three) in a cluster, rather than a line or chain as in L. wallae n . sp. (described above) (Fig. 3.40). These stages with thin and patchily distributed surface coats are beginning to form parasitophorous vacuoles (PV) (Figs. 3.41, 3.42), earlier than in L. pacificodae n . sp. and $L$. wallae n . sp. The PV space forms by the emptying of vesicles containing balls of dark, amorphous material at the surface of these stages). There were no small, light (empty) vesicles like those seen in L. pacificodae n. sp. and L. wallae n . sp. surrounding merogonial stages involved in PV formation. The dark material shows some resemblance to material seen in vesicles in L. wallae n. sp., except that in the present species dark material was surrounded by large spaces rather than being in direct contact with the cytoplasm (Figs. 3.38, 3.48). Darker tubule-like structures are sometimes visible in the dark balls (Fig. 3.48) but tubules are never present in vesicles and never or rarely appear in PV spaces. After thick PV spaces have formed around thinly coated stages, these plasmodia continued to divide by binary fission (Figs. 3.44, 3.45). Surface coats continue to thicken as stages become sporogonial plasmodia, which then "round-up" like stages seen in L. pacificodae n. sp. (see Lom \& Pekkarinen 1999) to have a round shape, smooth, thick surface coat, and lose most of the PV space. These rounded sporogonial plasmodia divided by multiple fission in rosette or "cloverleaf" form (Fig. 3.46) to produce four daughters. These daughters form sporoblasts and eventually spores, one per PV, or rarely two per PV (Fig. 3.49).
Development of Loma lingcodae n. sp.
Merogonial plasmodia and small, round meronts, loosely packed with ribosomes, and covered in host RER (Figs. 3.53, 3.54) begin to accumulate dark balls filled with amorphous material, similar to those described in $L$. wallae n. sp. at their surfaces (Fig. 3.56). No small, light (empty) vesicles similar to those described in L. pacificodae n . sp. and L. wallae n . sp . were seen. In one case a meront was seen with three structures resembling spindle plaques (Fig.
3.55). Meronts also contain small (e.g. 76.6 nm ), round, double-membrane bound vesicles with dark centers similar to those seen in L. pacificodae n . sp. and resembling those hypothesized to be possible mitochondrial homologs (Fig. 3.58)(Williams et al., 2002). Host RER disappears and surface coat thickens substantially (up to 38 nm ), prior to parasitophorous vacuole (PV) formation, thus PV spaces form later than in other species described above. Parasitophorous vacuoles form by coalescence of large, tubule-filled vesicles around sporoblasts. No sporonts or sporogonial plasmodia (according to the definition given here) were seen. Instead, thickerwalled plasmodia without PV spaces would be the only stage prior to sporoblast formation. Spores are found four per PV (Fig. 3.57).

## Development of Loma richardi n. sp.

Prior to sporogony, merogonial plasmodia with thin, patchily distributed surface coats, in direct contact with cytoplasm and covered by host RER cisternae (Fig. 3.64) begin to accumulate around them small, light (empty) membrane-bound vesicles like those seen in $L$. pacificodae n . sp. These small vesicles coalesce at the cell surface to create the parasitophorous vacuole (PV), thus the PV arises earlier than in L. pacificodae n . sp. and later than in L. kenti n . sp. (Figs. 3.69-3.71). Host RER cover disappears, the surface coat thickens, PV spaces grow, cytoplasmic contents become less dense and filled with large vacuoles like those seen in $L$. pacificodae n . sp. but not seen in $L$. lingcodae n . sp. Large, tubule-filled vesicles coalesce to empty their contents into PV spaces of sporogonial stages. These stages divide by binary fission (Fig. 3.66). Concentric rings of endoplasmic reticulum (ER) and obvious nucleoli are seen in these thick-walled sporonts Fig. 3.65) inside large, tubule-packed PV spaces. Sporonts develop into sporoblasts (two per PV) at slightly different rates (Figs. 3.67, 3.68) and ultimately form spores, two per PV (Fig. 3.72).

## DISCUSSION

This study described five new species of Loma from British Columbian marine fishes based on morphological and developmental characteristics, rDNA, and the occurrence in different hosts.

## Comparative ecology and pathology of new Loma species

## Prevalence

Loma pacificodae n . sp. and L. wallae n . sp. were both found in approximately $1 / 3$ of their respective hosts, which was similar to the level reported by Kent et al. (1998) and strengthens the suggestion that these are widespread and important pathogens (see pathology below) of Pacific cod and walleye pollock, respectively. The observation that the least prevalent species, L. richardi n. sp., occurred in about $13 \%$ of hosts examined, suggested that all these parasites are widespread in these British Columbian fishes. Prevalence differences among species (e.g. L. kenti $\mathrm{n} . \mathrm{sp}$. and $L$. richardi n . sp. were half as prevalent as $L$. pacificodae n . sp ., whereas L. lingcodae n. sp. was intermediate between these) pointed to potential differences in these hosts' susceptibilities, movements, these parasites' virulences, or their life cycles. Differences in host susceptibility were also suggested by the result that L. pacificodae n . sp . appeared more prevalent in Pacific cod from southern Vancouver Island compared to the northern west coast of Vancouver Island. Such differences in host susceptibility among hosts from different geographic localities were demonstrated in another Loma species, L. salmonae (Shaw et al., 2000a, b; Shaw et al., 2001).

The prevalence of Loma species from Atlantic cod and haddock from Nova Scotia, which are type host and locality of $L$. morhua Morrison \& Sprague, 1981, and type host of $L$. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981, respectively, were significantly lower (see chi-square Table 3.2) than those from the Pacific Loma species in closely related gadids (Pacific cod), suggesting there could be a difference in host susceptibility between Atlantic and Pacific Loma hosts. Another explanation could be a difference in the seasonal timing of xenoma maturation. For example, xenoma production may coincide with the reproduction of the host (Shaw \& Kent, 1999). Alternatively, perhaps the Atlantic Loma species prevalences were underestimated because xenomas were destroyed or decayed in the fish heads during handling at sea or in the market.

This study showed that prevalence estimates could be affected by technique. Prevalence estimates were higher in histology than in wet mounts from Loma species with xenomas in the central venous sinus, while the reverse was true (wet mount prevalence was higher) for species with xenomas primarily in the secondary gill lamellae. This difference was statistically significant for $L$. richardi n . sp., the species with the smallest xenomas, suggesting that wet mounts were more sensitive and should be used in surveys for this species. Also, PCR amplification produced lower prevalence estimates (data not shown), presumably because the PCR primers used were designed for more general targets (shorter) rather than for sensitivity (longer) like those designed for survey of L. salmonae (Docker et al.1997a).

This study also included a survey of other Pacific fishes for Loma species, with particular emphasis on species in the same host groups as those found to be infected in the survey of Kent et al. (1998); however, only one additional species of fish (pile perch) was observed with Loma-like infection, although for some species the sample size was small. This result suggested, as did Kent et al. (1998) that these new Loma species may be host-specific, or that if they use other reservoir hosts, the infection did not form xenomas in the gills.

## Pathology

The pathogenesis of these Loma species appeared to be similar to that observed for other Loma species (Speare et al., 1989; Morrison \& Sprague, 1981a; c; Morrison, 1983; Shaw \& Kent, 1999). This suggested these hosts may potentially develop resistance to reinfection after prior exposure, depending on conditions (e.g. temperature), as has been shown for L. salmonae (Speare et al., 1998b; Beaman et al., 1999; Kent et al., 1999; Shaw et al., 2001). This should be investigated in these new Loma species. Response to infection was most obvious in $L$. pacificodae n . sp. and $L$. wallae n . sp. but was rarely seen in the other species, suggesting, as with the differences in prevalence (above), that host response may be different in these species. Site of infection seemed to show a relationship with prevalence and host response. For example, species with xenomas in the central venous sinus (L. pacificodae n . sp . and $L$. wallae $\mathrm{n} . \mathrm{sp}$.) rather than the secondary lamellae had more signs of host response and a higher prevalence of infection. Where xenomas commonly occurred in both sites, as in L. pacificodae n. sp., xenomas in the central venous sinus more often showed host cell infiltration and phagocytosis than did xenomas in the pillar system of the secondary lamellae, suggesting host response and site of infection were associated; however, this difference in response could be due to the relative timing of xenoma formation over the course of infection, as some host-overcome
xenomas in the central venous sinus were seen adjacent to smaller, secondary xenomas at the base of the secondary lamellae.

The presence of all of these Loma species in the gonads raises the possibility that they could transmit vertically, as has been suggested for L. salmonae and L. morhua (Morrison, 1983; Docker et al., 1997a) and many other microsporidia (Dunn et al., 2001). Vertical transmission, if it exists in these new Loma species, could have implications for virulence. For example, models and experimental studies predict that greater vertical transmission relative to horizontal transmission often selects for reduced virulence (Lipsitch et al., 1995; Mangin et al., 1995; Lipsitch et al., 1996; Agnew \& Koella, 1997; Koella \& Agnew, 1997; Dunn \& Smith, 2001).

## New features

The suggestion that there were mitochondrial homologs in microsporidia (Williams et al., 2002) prompted the examination here for merogonial cytoplasms containing reduced mitochondria-like structures similar to those of Williams et al. (2002). The small (76.5-88.2 nm ) double-membrane-bound vesicles shown in Figure 3.15, with contents slightly darker than surrounding cytoplasm, seemed to resemble mitochondrial homologs of Williams et al. (2002).

## Considerations regarding character interpretation

## Relative character variability in new Loma species

This study suggested some characters varied more within species than others. For example, characters with high intraspecific variability were: number of vesicles per sporoblast or spore, tubules per sporoblast or spore, spore size and polar filament turns. Characters that appeared to show lower intraspecific variability were: tubules per vesicle, xenoma size, number of spores per PV, and vesicle size. The number of tubules per sporoblast or spore would be expected to vary within species because tubules change dramatically with development. Vesicles per sporoblast or spore also change dramatically over the course of development and may also be somewhat artificial, due to measurement method (see Materials and Methods) rather than biological differences.

## Characters that may be correlated

Some characters may be genetically linked and so may not vary independently. Data are usually not sufficient for statistical correlation tests of these character suites; however, such correlations would be of functional and taxonomic interest. For example, we may expect some characters to evolve in concert, such as: spore length and width; spore size, number of polar filament turns, and posterior vacuole size; polar filament length, mode of infection
(autoinfection vs. horizontal infection), and the cell type in which xenomas occur. Others have proposed that polar filament length (or number of coils) may be associated with mode of infection, since spores with different polar filament length tend to be specialized to infect different host cells, tissues or different host species (Cali \& Takvorian, 1999).

In this study, xenoma size and number of developmental stages in the developmental sequence seemed to show a relationship. For example, L. kenti n . sp . and L. pacificodae n . sp . (cvs form) had significantly larger xenomas and also displayed sporogonial ("rounded up") stages not present in the other three species. This suggests the number of developmental divisions could be finite and could indicate xenoma size. Similarly, the reverse could also be true, that xenoma size might indicate the number of developmental stages. The observation that spores tended to fill the volume of mature xenomas (in Loma species, but not in Glugea species in which developmental stages and spores occupy different regions of mature xenomas) would suggest that division is not so much "asynchronous", as Lom \& Pekkarinen (1999) suggested, but proceeds to a species-characteristic climax, at which point xenomas are $>90 \%$ full of spores, having a characteristic (diagnostic) size, and this size may be informative, providing hints about the underlying developmental sequence. When these sequences are better characterized by further study, it may be possible to use xenoma size not only for species-recognition, but also as an indicator of plasticity in development where this is suspected (e.g. in the case of $L$. pacificodae n . sp . where two distinct xenoma sizes appear to arise by some form of developmental plasticity).

## Spore shrinkage factors

Spore shrinkage from fixation was found to be substantial (e.g. up to a factor of 0.6 ) and differed considerably among different fixation methods. This suggests comparisons of spore sizes among studies that have used different fixatives should always be interpreted cautiously. However, sometimes comparisons among such studies may be helpful or necessary. For example, when describing new species by reference to publications in which only spores were measured, or when presented with material that is decayed (e.g. cod heads) for which electron microscopy is impossible. For this reason, it may be helpful to estimate relative spore shrinkage factors and extrapolate to estimate spore sizes for other material fixed in this study, or to use, for example, the "fresh" size estimates to compare to those of others. It would be less appropriate to apply such shrinkage factors to material fixed by others, since differences in fixation, handling, and measurement technique could be significant. Here, shrinkage was greatest with glutaraldehyde followed by Spurr's resin-embedding of spores ( 1.5 x smaller in length, 1.35 x
smaller in width), followed by histological sections, ethanol-fixed, and frozen spores ( $\sim 1.2 \mathrm{x}$ smaller in both length and width), suggesting some consistent loss of volume occurs once spores are infiltrated or mechanically damaged (e.g. freezing). Perhaps in such cases the posterior vacuole becomes reduced.

## Characters potentially determined by the host

For Loma species examined here, many characters could be phenotypically plastic and dependant on the host. In the past, some have speculated that this could be true for features of the xenoma, which is a host-parasite complex (Lom \& Pekkarinen, 1999). Characters that would be least likely to be phenotypically plastic depending on the host are meront surface, meront contents, spore shape and size, spore contents, and number of divisions during development. Characters that would be most likely to depend on the host are xenoma size, xenoma wall features, xenoma cytoplasmic contents including the sac or vacuole around the parasite that is supposedly of host origin (Bekhti \& Bouix, 1985; Lom \& Pekkarinen, 1999). If the parasite can vary depending on the host species or environment in which it finds itself, and if most features of the xenoma (internal cytoplasmic features and site or external gross morphology) are determined by the host, then data gathered for these five new Loma species might represent a single phenotypically plastic species, rather than separate species. However, Lom and Pekkarinen (1999) argued against this, suggesting the xenoma is strongly determined by parasite rather than host and remains a good diagnostic character. As evidence, Lom \& Pekkarinen (1999) referred to the observation that xenomas of well-differentiated microsporidians, Glugea spp. and Loma spp., infect identical cells in similar hosts (a neutrophil granulocyte), yet produce xenomas with completely different walls, cytoplasmic organization, and other features. Lom \& Nilsen (2003) also maintained that all evidence so far suggests features of xenomas depend more on the intrinsic characters of the parasite, rather than the host. This perspective supports observations from this study, showing xenoma size and tubule-filled vesicles (hypothesized to be of "host" origin by Lom \& Pekkarinen, 1999) seem to be good characters for distinguishing Loma species and are better than many traditional characters like spores or polar filaments (discussed above) that are more difficult to measure and sometimes overlap significantly between species.

## Questions answered by transmission experiments

Experimental transmission of spores to alternate hosts can help directly demonstrate host specificity of a parasite, as well as indirectly examine the potential for reproductive isolation
between parasites, particularly if parasites undergo reproduction within the host, as do Loma species. Shaw et al. (1997) and Shaw et al. (2000a) demonstrated that L. embiotocia and $L$. salmonae will not infect reciprocal hosts, and so are presumably reproductively isolated. DNA sequence differences between L. salmonae and L. embiotocia (Shaw et al. 1997; and Chapter 4; see Appendix 12 for name equivalents) support the results from transmission data, suggesting these are distinct species. These two species overlap in spore size, polar filament turns, and many other features, but like other Loma species, are distinct in host preference and gene sequences (i.e. they are cryptic species) (Brown \& Kent, 2002). This strengthens the argument that, despite the similarity in morphology between L. pacificodae n . sp . and $L$. wallae n . sp., their occurrence in different hosts may be important. Transmission studies are still lacking for all but one of these new Loma species. Only L. lingcodae n. sp. was demonstrated to be nontransmissible to two fish species, chinook salmon and shiner perch, which are hosts of $L$. salmonae and L. embiotocia (respectively) (Shaw \& Kent, 1999; R. W. Shaw, personal communication). While this result suggests these three species are separate, it cannot help answer the question of whether $L$. lingcodae n . sp. is separate from the other four new species presented here. Similar transmission experiments were attempted with L. pacificodae n . sp . spores (referred to in Shaw \& Kent, 1999), and although this species would not infect salmon or shiner perch, the control for spore viability, transmission back to a Pacific cod host, was not performed due to lack of suitable laboratory fish (R. W. Shaw, personal communication). While laboratory transmission studies may be helpful if they show species cannot be transmitted to alternate hosts, results may not reflect transmissibility in nature.

## Ultrastructure vs. DNA data and species concepts in Loma

Results here showed $L$. kenti n . sp. from Pacific tomcod was different from the other new species, particularly in the nature of its vesicles, developmental timing, and spore features, whereas the other four new Loma species fell into two sister-species pairs. The first pair, $L$. pacificodae n . sp. from Pacific cod and $L$. wallae n . sp . from walleye pollock were similar to one another but distinguished by differences in developmental stages and vesicles (see remarks, above). The second pair, L. lingcodae n. sp. from lingcod and L. richardin. sp. from sablefish were similar but could be distinguished by xenoma features, vesicle and spore sizes (see remarks, above). These results were consistent with molecular (rDNA and EF-1 $\alpha$ ) sequence data presented here and in depth in Chapter 4 (See Appendix 12 for name equivalences between chapters; and see Chapter 4, Figs. 4.4, 4.9), which suggest L. kenti n . sp. is an outlier among the
five new species, whereas the other four species form two close sister-pairs (one pair in each of clades A and B of Fig. 4.4, Chapter 4).

However, genetic results presented in Chapter 4 (Figs. 4.2 to 4.9), suggested partners of the two sister-species (L. pacificodae n. sp. and L. wallae n . sp.; L. lingcodae n. sp. and $L$. richardi $\mathrm{n} . \mathrm{sp}$.) were genetically virtually indistinguishable, having as much molecular variation within as between sister species. Thus, in Chapter 4, data did not statistically support separation of $L$. pacificodae n . sp . from $L$. wallae n . sp., or $L$. lingcodae n . sp . from $L$. richardi n . sp . However, analyses in Chapter 4 also could not reject the validity of all four species (e.g. see Chapter 4, Table 4.10, trees \#2, 3, 6, 17, and Table 4.11, tree \#4). Two alternative interpretations of these results must be considered.

Alternative 1: based on molecular similarity, these two pairs should be grouped into two species, rather than four (i.e. L. pacificodae n. sp. to L. wallae n. sp. should be considered conspecific and $L$. lingcodae n . sp. to $L$. richardi n . sp. should be considered conspecific). This alternative implies there are significant morphological polymorphisms in these species that correspond to the host. Such polymorphism could have arisen by selection under partial allopatry created by different host species environments. This allopatry would have to be incomplete, enabling sufficient gene flow to prevent speciation. Alternately, the differences could be due to phenotypic plasticity in different hosts. There is no evidence that a single species of microsporidia can display this degree of phenotypic plasticity, and no evidence that Loma species (or other fish-parasitic species) have sex and can effectively maintain gene flow under any conditions.

Alternative 2: based on morphological differences and overall molecular conservation, these two pairs are four valid species (i.e. L. pacificodae n. sp., $L$. wallae n. sp., L. lingcodae n. sp . and $L$. richardin. sp. are valid). This alternative implies the inability of the molecular regions sequenced (rDNA and partial EF) to distinguish two pairs of valid species. The best explanation is that these are recent species. Ribosomal DNA and especially the ITS region is not normally expected to be invariant among well differentiated species, since rDNA tends to accumulate mutations quickly in less-constrained regions such as spacer regions, loops and terminal regions of stems (Hancock \& Dover, 1988; Hancock \& Dover, 1990; Hancock, 1995). Similarly, the single-copy EF gene sequence would be expected to accumulate variation quickly at 3 rd codon positions. This high rate of mutation at less critical nucleotide positions should reveal divergence soon after gene flow has stopped or become negligible accompanying speciation. Explanations for unexpectedly low genetic divergence among good morphospecies
include an unusually low background mutation rate, unusually conserved gene regions or genes, species hybridization (reticulate speciation), insufficient length of gene examined to observe evolutionary divergence, or recent speciation. There is evidence for all of these explanations except low background mutation rate in Loma species (discussed in detail in Chapter 4). In addition, for any molecular marker, differences may not yet be fully fixed in the population for recent species; so, sampling method (number and distribution of individuals; number of nucleotide sites and loci) will affect the ability to separate species.

When recent speciation is suspected, it is perhaps even more important than usual to define "species" for the given taxon. As stated in the Introduction, here the operational species concept is derived from a synthesis of arguments in Wheeler \& Meier (2000) in two parts. The first criterion for distinguishing one species from another when in sympatry is that each must possess separate discrete characters or overlapping characters with a statistically separate mean. The second criterion is that the suite of characters that provides evidence for species must agree across the sampled populations. It is not clear whether Loma species undergo sexual reproduction; therefore, reproduction-based concepts requiring species to be breeding groups reproductively isolated from other such groups, i.e. the Biological Species Concept of Mayr (1940) (and see arguments Wheeler \& Meier, 2000), or independent reproductive (evolutionary) lineages (Templeton, 1989; Templeton, 1994; Sites \& Crandall, 1997), are most appropriate. The concept used here offers working criteria for evaluating species and is consistent with either reproduction-based or lineage-based theoretical species concepts. These five new Loma species were found in potential contact (often collected in the same trawls), or sympatry, as specified by the working definition. Morphological data suggested these species differ in both discrete and continuously distributed characters. Some of these differences were supported statistically, whereas others (e.g. developmental timing differences and differences in stages) were observed only qualitatively. Genetic analyses in Chapter 4 separated these five new species into three groups (as discussed above), but did not consistently separate two sister-species pairs from their partners. However, analyses in Chapter 4 failed to reject these species, and so does not negate the evidence of the present study, which supports the more complex hypothesis of two species, rather than one for each sister-species pair (L. pacificodae n . sp . and $L$. wallae $\mathrm{n} . \mathrm{sp}$.; $L$. lingcodae n . sp. and $L$. richardi $\mathrm{n} . \mathrm{sp}$.).

The second argument for five separate Loma species in Pacific cod, walleye pollock, Pacific tomcod, lingcod and sablefish, after morphological differences, is their occurrence in different hosts. While host-specificity has not been demonstrated experimentally between these
two pairs of sister-species, host has been shown to be an important character distinguishing other close relatives in genus Loma and relatives (Lom \& Nilsen, 2003) where transmission experiments were performed (Shaw et al., 1997; Shaw \& Kent, 1999; and see Lom \& Nilsen, 2003). For Loma species, experienced taxonomists have suggested species be defined using multiple morphological characters, including most importantly: host species, PV formation, spore size, xenoma size, xenoma wall, cell origin of xenoma formation, form and abundance of episporontal inclusions (Lom \& Pekkarinen, 1999; Weiss \& Vossbrinck, 1999; Lom, 2002). This study confirmed that such characters form clusters separated by discontinuities, and that these correspond to host-group. As Mallet (1995) pointed out, it is "not clear why we need to improve on a good taxonomist's or naturalist's definition" of species; however, host group and any associated morphological characters could be misleading if a parasite uses multiple host species or if there is host-dependent phenotypic plasticity. So, even though the assumption of host specificity was used to form testable hypotheses at the start of this study, this does not mean host is a valid character. Host specificity was not directly tested here; however, it was indirectly examined by surveying potential alternate hosts for presence of the parasite. Plasticity was also indirectly examined by sampling a wide range of characters and developmental stages across many individuals (cells) and host isolates. Ideally, the influence of host on these characters should be examined by experimental transmission to alternate hosts for each species of Loma; however, this is may be technically difficult or impossible in some cases. In the few cases where it is possible (see above), the results may not closely reflect transmission in nature.

The combination of character differences among these five species would seem to present a picture of relationships among species, but this does not necessarily represent evolutionary relationships. Clearly there is insufficient knowledge to model character evolution in Loma species. For example, it is not known how spore size or number of polar filament turns changes over time in a lineage. For this reason, phenetics or numerical taxonomy (Sokal \& Sneath, 1963; Sneath \& Sokal, 1973; Clifford \& Stephenson, 1975), in which taxa are clustered based on morphological similarities to infer phylogenetic or evolutionary relatedness, is unsuitable. It would result in erroneous groupings, especially in the microsporidia, due to homoplasy (convergent evolution) of species under common environmental selection pressures - a phenomenon that is widespread (if not dominant) in evolution. Despite the pervasiveness of homoplasy in the microsporidia (see Chapter 1, and Weiss \& Vossbrinck, 1999) the argument for separating these five species which display statistically supported differences in multiple characters and host is justified because they are close relatives (see Chapter 4) in potential
sympatry. These species may be an example of "adaptive radiation" as defined by Schluter (2000), although convincing demonstration of this would require experiment, or for a start, more species studied very closely in nature over a sufficiently long time.

## FURTHER INVESTIGATION

Further studies are needed to assess the transmissibility of these species to alternate host species. Intraspecific variation could be better characterized by sampling more individuals across their geographic range. At present, the actual geographic ranges of these parasites are not known beyond the sampling areas, namely on the coasts of Vancouver Island, British Columbia, from as far south as Juan de Fuca Strait to as far north as Queen Charlotte Sound. Experimental studies should also examine, in particular, the possibility of phenotypic plasticity in the characters by experimental transmission to different host cells, host individuals, populations or species, and under different conditions. Further study of the developmental sequence of these species should also help address the question of whether there may be sporogony stages (or other stages) that were missed in this study. Early development and transmission will also be of interest, given the intriguing results from studies of Loma salmonae, which appears to be highly resistant in the environment and persistent at low-levels in its hosts, making it a difficult commercial pathogen to control. As with L. salmonae (Sánchez et al., 2001), there may be promise for development vaccines against these new Loma species.

## ACKNOWLEDGEMENTS

The Natural Sciences and Engineering Research Council of Canada supported this work (strategic grant 582073 to M. L. Adamson). I sincerely thank Jim Boutillier at the Pacific Biological Station and the crew of the research vessel W. E. Ricker for assistance in identifying and collecting Pacific fishes. I am grateful to Elaine Humphrey and technicians at the BioImaging Facility, UBC, and to Susan Shinn for her help with ultrathin sectioning for TEM. I thank Stewart Johnson and Laura Brown at the National Research Council in Halifax, for their assistance with collection of Atlantic cod and haddock heads, and I thank others in their nearby laboratories for kindly tolerating the excruciating smell.

## LITERATURE CITED

Agnew, P. and Koella, J. C. 1997. Virulence, parasite mode of transmission, and host fluctuating asymmetry. Proceedings of the Royal Society of London Series B 264:9-15.

Azevedo, C. and Matos, E. 2002. Fine structure of a new species, Loma myrophis (Phylum Microsporidia), parasite of the Amazonian fish Myrophis playrhynchus (Teleostei, Ophichthidae). European Journal of Protistology 37:445-452.

Beaman, H. J., Speare, D. J., Brimacombe, M. and Daley, J. 1999. Evaluating protection against Loma salmonae generated from primary exposure of rainbow trout, Oncorhynchus mykiss (Walbaum), outside of the xenoma-expression temperature boundaries. Journal of Fish Diseases 22:445-450.

Bekhti, M. and Bouix, G. 1985. Loma salmonae (Putz, Hoffman et Dunbar, 1965) et Loma diplodae n. sp., microsporidies parasites de branchies de poissons téléosteens: implantation et données ultrastructurales. Protistologica 21(1):47-59.

Brown, A. M. V. and Kent M. L. 2002. Molecular diagnostics for Loma salmonae and Nucleospora salmonis (microsporidia) In Molecular diagnostics of salmonid diseases. Cunningham, C. O. (ed.). Kluwer Academic Publishers, Dordrecht p. 267-283.

Cali, A. and Takvorian, P. M. 1999. Developmental morphology and life cycles of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 85-128.

Canning, E. U. and Lom, J. 1986. The microsporidia of vertebrates. Academic Press, London. 289 pp.
Clifford, H. T. and Stephenson, W. 1975. An introduction to numerical classification. Academic Press, New York. 229 pp .

Docker, M. F., Devlin, R. H., Richard, J., Khattra, J. and Kent, M. L. 1997a. Sensitive and specific polymerase chain reaction assay for detection of Loma salmonae (Microsporea). Diseases of Aquatic Organisms 29(1):41-48.

Dunn, A. M. and Smith, J. E. 2001. Microsporidian life cycles and diversity: the relationship between virulence and transmission. Microbes and Infection 3:381-388.

Dunn, A. M., Terry, R. S. and Smith, J. E. 2001. Transovarial transmission in the microsporidia. Advances in Parasitology 48:57-100.

Faye, N., Toguebaye, B. S. and Bouix, G. 1995. On the cytology and development of Loma boopsi n. sp. (Microspora, Glugeidae), parasite of Boops boops (Pisces, Teleostei, Sparidae) from the coasts of Senegal. Archiv Fur Protistenkunde 146:85-93.

Fomena, A., Coste, F. and Bouix, G. 1992. Loma camerounensis new species (Protozoa: Microsporida) a parasite of Oreochromis niloticus Linnaeus 1757 Teleostei Cichlidae in fish-rearing ponds in Melen Yaounde Cameroon. Parasitology Research 78(3):201-208.

Hancock, J. M. 1995. The contribution of DNA slippage to eukaryotic nuclear 18 S rRNA evolution. Journal of molecular evolution 40:629-639.

Hancock, J. M. and Dover, G. A. 1990. 'Compensatory slippage' in the evolution of ribosomal RNA genes. Nucleic Acids Research 18(20):5949-5954.

Hancock, J. M. and Dover, G. A. 1988. Molecular coevolution among cryptically simple expansion segments of eukaryotic 26S/28S rRNAs. Molecular Biology and Evolution 5(4):377-391.

Hauck, A. K. 1984. Mortality and associated tissue reactions of chinook salmon, Oncorhynchus tshawytscha (Walbaum), caused by the microsporidian Loma sp. Journal of Fish Diseases 7:217-229.

Humason, G. L. 1979. Animal tissue techniques. 4th ed. W. H. Freeman and Company, San Francisco, California, 470 p.

Kabata, Z. 1959. On two little-known microsporidia of marine fishes. Parasitology 49:309-315
Kent, M. L., Elliot, D. G., Groff, J. M. and Hedrick, R. P. 1989. Loma salmonae (Protozoa: Microspora) infections in seawater reared coho salmon Oncorhynchus tshawytscha. Diseases of Aquatic Organisms 20:231-233.

Kent, M. L., Dawe, S. C. and Speare, D. J. 1995. Transmission of Loma salmonae (Microsporea) to chinook salmon in sea water. Canadian Veterinary Journal 36:98-101.

Kent, M. L., Dawe, S. C. and Speare, D. J. 1999. Resistance to reinfection in chinook salmon Oncorhynchus tshawytscha to Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 37:205-208.

Kent, M. L., Traxler, G. S., Kieser, D., Richard, J., Dawe, S. C., Shaw, R. W., Prosperi-Porta, G., Ketcheson, J. and Evelyn, T. P. T. 1998. Survey of salmonid pathogens in ocean-caught fishes in British Columbia, Canada. Journal of Aquatic Animal Health 10:211-219.

Koella, J. C. and Agnew, P. 1997. Blood-feeding success of the mosquito Aedes aegypti depends on the transmission route of its parasite Edhazardia aedis. Oikos 78:311-316.

Lipsitch, M., Nowak, M. A., Ebert, D. and May, R. M. 1995. The population dynamics of vertically and horizontally transmitted parasites. Proceedings of the Royal Society of London. Series B. Biological Sciences 260:321-327.

Lipsitch, M., Siller, S. and Nowak, M. A. 1996. The evolution of virulence in pathogens with vertical and horizontal transmission. Evolution 50(5):1729-1741.

Lom, J. 2002. A catalogue of described genera and species of microsporidians parasitic in fish. Systematic Parasitology 53:81-99.

Lom, J. and Nilsen, F. 2003. Fish microsporidia: fine structural diversity and phylogeny. International Journal for Parasitology 33:107-127.

Lom, J. and Pekkarinen, M. 1999. Ultrastructural observations on Loma acerinae (Jirovec, 1930) comb. nov. (Phylum Microsporidia). Acta Protozoologica 38:61-74.

Loubès, C., Maurand, J., Gasc, C., de Buron, I. and Barral, J. 1984. Étude ultrastructurale de Loma dimorpha n. sp., microsporidie parasite de poissons Gobiidae languedociens. Protistologica 14(4):579-589.

Mallet, J. 1995. A species definition for the modern synthesis. Trends in Ecology and Evolution 10:294-299.
Mangin, K. L., Lipsitch, M. and Ebert, D. 1995. Virulence and transmission modes of two microsporidia in Daphnia magna. Parasitology 111(2):133-142.

Mayr, E. 1940. Speciation phenomena in birds. The American Naturalist 74:249-278.
Morrison, C. M. and Sprague, V. 1981a. Electron microscopical study of a new genus and new species of microsporida in the gills of Atlantic cod, Gadus morhua L. Journal of Fish Diseases 4:15-32.

Morrison, C. M. and Sprague, V. 1981b. Light and electron microscope study of microsporida in the gill of haddock, Melanogrammus aeglefinus (L.). Journal of Fish Diseases 4:179-184.

Morrison, C. M. and Sprague, V. 1981c. Microsporidian parasites in the gills of salmonid fishes. Journal of Fish Diseases 4:371-386.

Morrison, C. M. and Sprague, V. 1983. Loma salmonae (Putz, Hoffman and Dunbar, 1965) in the rainbow trout, Salmo gairdneri Richardson, and L. fontinalis sp. nov. (Microsporidia) in the brook trout, Salvelinus fontinalis (Mitchill). Journal of Fish Diseases 6:345-353.

Nemeczek, A. 1911. Beiträge zur Kenntnis der Myxo- und Microsporidien der Fishce. Archiv für Protistenkunde 22:143-169.

Ovcharenko, N. O., Sarabeev, V. L., Wita, I. and Czaplińska, U. 2000. Loma mugili sp. n., a new microsporidium from the gills of grey mullet (Mugil soiuy). Vestnik zoologii 34(4-5):9-15.

Putz, R. E., Hoffman, G. L. and Dunbar, C. E. 1965. Two new species of Pleistophora (Microsporidia) from North American fish with a synopsis of Microsporidia of freshwater and euryhaline fishes. Journal of Protozoology 12(2):228-236.

Quicke, D. L. J. 1993. Principles and techniques of contemporary taxonomy. Blackie Academic \& Professional. London. 311 pp .

Ramsay, J. M., Speare, D. J., Dawe, S. C. and Kent, M. L. 2002. Xenoma formation during microsporidial gill disease of salmonids caused by Loma salmonae is affected by host species (Oncorhynchus tshawytscha, $O$. kisutch, O. mykiss) but not by salinity. Diseases of Aquatic Organisms 48:125-131.

Sánchez, J. G., Speare, D. J., Markham, R. J. F. and Jones, S. R. M. 2001. Experimental vaccination of rainbow trout against Loma salmonae using a live low-virulence variant of $L$. salmonae. Journal of Fish Biology 59:442-448.

Sandeep, B. V. and Kalvati, C. 1985. A new microsporidian, Loma trichiuri n. sp., from the gill of a marine fish, Trichiurus savala Cuv. (Trichiuridae). Indian Journal of Parasitology 9(2):257-259.

Shaw, R. W. and Kent, M. L. 1999. Fish microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 418-446.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000a. Viability of Loma salmonae (Microsporidia) under laboratory conditions. Parasitology Research 86:978-981.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2001. Phagocytosis of Loma salmonae (Microsporidia) spores in Atlantic salmon (Salmo salar), a resistant host, and chinook salmon (Oncorhynchus tshawytscha), a susceptible host. Fish \& Shellfish Immunology 11:91-100.

Shaw, R. W., Kent, M. L., Brown, A. M. V., Whipps, C. M. and Adamson, M. L. 2000b. Experimental and natural host specificity of Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 40: 131-136.

Shaw, R. W., Kent, M. L., Docker, M. F., Brown, A. M. V., Devlin, R. H. and Adamson M. L. 1997. A new species of Loma (Microsporea) in shiner perch (Cymatogaster aggregata). Journal of Parasitology 83(2): 296-301.

Schluter, D. 2000. The ecology of adaptive radiation. Oxford University Press, Oxford, UK. 288 pp.
Sites, J. W. and Crandall, K. A. 1997. Testing species boundaries in biodiversity studies. Conservation Biology 11(6):1289-1297.

Sneath, P. H. A. and Sokal, R. R. 1973. Numerical taxonomy, W. H. Freeman, San Francisco, California. 573 pp.
Sokal, R. R. and Sneath, P. H. A. 1963. Principles of numerical taxonomy, W. H. Freeman \& Co. San Francisco. 359 pp .

Speare, D. J., Arsenault, G. J. and Buote, M. A. 1998a. Evaluation of rainbow trout as a model for use in studies on pathogenesis of the branchial microsporidian Loma salmonae. Contemporary Topics in Laboratory Animal Science 37:55-58.

Speare, D. J., Beaman, H. J., Jones, S. R. M., Markham, R. J. F. and Arsenault, G. J. 1998b. Induced resistance in rainbow trout, Oncorhynchus mykiss (Walbaum), to gill disease associated with the microsporidian gill parasite Loma salmonae. Journal of Fish Diseases 21(2):93-100.

Speare, D. J., Brackett, J. and Ferguson, H. W. 1989. Sequential pathology of the gills of coho salmon with a combined diatom and microsporidian gill infection. Canadian Veterinary Journal 30:571-575.

Templeton, A. R. 1989. The meaning of species and speciation: a genetic perspective. pgs. 3-27 in Otte, D. and Endler, J. A. (eds) Speciation and its consequences. Sinauer Associates, Sunderland, Massachusetts.

Templeton, A. R. 1994. The role of molecular genetics in speciation studies. pgs. 455-477 in Schierwater, B., Streit, B., Wagner, G. P. and DeSalle, R. (eds) Molecular ecology and evolution: approaches and applications. Bierkauser Verlag, Basel, Switzerland.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.

Vávra, J. and Larsson, J. I. R. 1999. Structure of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 7-84.

Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. and Woese, C. R. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326(6111):411-4.

Weiss, L. M. and Vossbrinck, C. R. 1999. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 129-171.

Wheeler, Q. D. and Meier, R. 2000. Species concepts and phylogenetic theory. Columbia University Press. New York, NY. 230 pp.

Williams, B. A. P., Hirt, R. P., Lucocq, J. M. and Embley, M. T. 2002. A mitochondrial remnant in the microsporidian Trachipleistophora hominis. Nature 418:865-869.

Zar, J. H. 1996. Biostatistical analysis. Prentice Hall, Upper Saddle River, New Jersey. 662 pp.

Table 3.1: Fishes examined for Loma-like infections in this study, showing common, scientific and family names, and total number examined. \# = number of individuals examined.

| Common name | Scientific name | Family | \# |
| :---: | :---: | :---: | :---: |
| Species examined with Loma infections: |  |  |  |
| Pacific cod | Gadus macrocephalus Tilesius, 1810 | Gadidae | 227 |
| walleye pollock | Theragra chalcogramma (Pallas, 1811) | Gadidae | 145 |
| Pacific tomcod | Microgadus proximus (Girard, 1854) | Gadidae | 419 |
| lingcod | Ophiodon elongatus Girard, 1854 | Hexagrammidae | 210 |
| sablefish | Anoplopoma fimbria (Pallas, 1811) | Anoplopomatidae | 197 |
| Atlantic cod | Gadus morhua L. | Gadidae | 427 |
| haddock | Melanogrammus aeglefinus (L.) | Gadidae | 232 |
| shiner perch | Cymatogaster aggregata Gibbons, 1854 | Embiotocidae | 94 |
| coho salmon | Oncorhynchus kisutch (Walbaum, 1792) | Salmonidae | 8 |
| pile perch | Rhacochilus vacca (Girard, 1855) | Embiotocidae | 1 |
| Species examined with suspicious gill appearance found not to have Loma infection: |  |  |  |
| Pacific hake | Merluccius productus (Ayres, 1855) | Merlucciidae | 218 |
| Pacific herring | Clupea pallasi Valenciennes, 1847 | Clupeidae | 97 |
| eulachon | Thaleichthys pacificus (Richardson, 1836) | Osmeridae | 81 |
| bigfin eelpout | Lycodes cortezianus Gilbert, 1890 | Zoarcidae | 21 |
| whitebait smelt | Allosmerus elongatus (Ayres, 1854) | Osmeridae | 15 |
| American shad | Alosa sapidissima (Wilson, 1812) | Clupeidae | 11 |
| whitespotted greenling | Hexagrammos stelleri Tilesius, 1810 | Hexagrammidae | 8 |

Table 3.2: Prevalence estimates for Loma species from various fishes, showing estimates from histological sections, wet mounts and both methods combined. Numbers in parentheses are total number of fish examined. Significant differences from the chi-squared test are shown by asterisk $\left({ }^{*}\right)$ beside the probability ( p -val) from the chi-squared test, for which expected values $(\exp \mathrm{H}+\mathrm{W})$ were obtained by pooling histology and wet mount or for between-species chisquared test expected values (exp allG or $\mathrm{L}+\mathrm{B}$ ) from pooling all species from gadoid hosts or order Scorpaeniformes (lingcod and sablefish) hosts, respectively. Histological estimates were obtained, as described, by S. C. Dawe. Wet mount estimates were made by M. L. Kent, S. C. Dawe, and A. Brown.

| Species | Histology $\% \quad(\mathrm{~N})$ | alence <br> Wet mount \% (N) | Chi-square test for wet mounts vs. histology $\exp$ mean p-value | Overall prevalence wet mount \& histology combined $\% \quad(\mathrm{~N})$ | Chi-square test for species differences in host group: <br> 1: Gadidae <br> 2: Scorpaeniformes exp <br> mean p-value |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| L. pacificodae n. sp. | 41.3 (121) | 26.0 (154) | $32.7>0.05$ | 32.6 (227) | $\begin{aligned} & \text { 1: } \\ & 14.4 \end{aligned}$ | $\begin{aligned} & <0.001^{*} \\ & <0.001^{*} \end{aligned}$ |
| L. wallae n . sp . | 35.6 (45) | 24.4 (131) | $27.3>0.1$ | 28.3 (145) |  |  |
| L. kenti n . sp. | 5.3 (57) | 14.4 (409) | $13.3>0.05$ | 14.4 (419) |  | $>0.5$ |
| spores Atlantic cod | - | 4.4 (427) | - - | 4.4 (427) |  | $<0.001 *$ |
| spores haddock | - | 6.5 (232) | - - | 6.5 (232) |  | <0.001* |
| L. lingcodae n. sp. | 16.7 (156) | 21.1 (199) | $19.2>0.1$ | 21.4 (210) | $\begin{array}{\|l\|} \mathbf{2 :} \\ 17.4 \end{array}$ | $\begin{aligned} & >0.1 \\ & >0.1 \end{aligned}$ |
| L. richardi n . sp . | 3.9 (128) | 13.2 (197) | $9.5<0.05^{*}$ | 13.2 (197) |  |  |
| L. embiotocia |  | 8.5 (94) | - - | 8.5 (94) | - | - |
| L. salmonae | - | 12.5 (8) | - - | 12.5 (8) |  |  |



Figure 3.1: Prevalence differences of $L$. pacificodae n. sp. across geographic regions or years, showing prevalence to be lowest on the northern west coast and highest in Nanaimo (east coast) and southern west coast of Vancouver Island. Small filled and open circles and small "x"s represent trawl locations from southern 1997, northern 1997, and 1998 sampling trips, respectively. Large triangles represent several 1997 trawl locations from which rDNA sequence for L. pacificodae n. sp. was obtained. Other Pacific Loma species were sampled from similar areas around Vancouver Island, continuously from the southern to mid-northern and northern coasts.

Table 3.3: Xenoma shape, location in gills, cell type, and degree of host response. cvs = central venous sinus, $\mathrm{bv}=$ blood vessel, eff $=$ efferent, $1^{\circ} \& 2^{\circ}=$ primary and secondary lamellae, respectively, fibr $=$ fibroblast.

| Species | Xenoma <br> Shape | Xenoma <br> Location | Cell type | Degree of <br> host <br> response |
| :--- | :--- | :--- | :--- | :--- |
| L. pacificodae n. sp. | elongate <br> or <br> round, <br> septate | in cvs, behind eff <br> bv or <br> at base of $2^{\circ}$ | fibr of lymph vessel <br> or <br> pillar cells | massive <br> or <br> little |
| L. wallae n. sp. | rounded <br> oval, <br> septate | cvs, connective <br> tissue | fibr connective <br> tissue | moderate |
| L. kenti n. sp. | round | base of $2^{\circ}$, <br> throughout $1^{\circ}$ | endothelial cells or <br> bv | little seen |
| L. lingcodae $\mathrm{n} . \mathrm{sp}$. | oval | tip of $2^{\circ}$, distal part <br> of $1^{\circ}$ | endothelial, pillar | little seen |
| L. richardi $\mathrm{n} . \mathrm{sp}$. | round | tip of $2^{\circ}$, <br> throughout $1^{\circ}$ | endothelial, pillar | none seen |

Table 3.4: Xenoma wall features in new Loma species.

| Species | Xenoma plasmalemma | Wall <br> thick- <br> ness <br> in $\mu \mathrm{m}$ | \# of lay ers | Layer description | Wall inclusions |
| :---: | :---: | :---: | :---: | :---: | :---: |
| L. pacificodae $\mathrm{n} . \mathrm{sp}$. | smooth or interdigitated | $\begin{aligned} & 1.5 \\ & -2.0 \end{aligned}$ | 1 | collagen fibers evenly intermixed with light granular material | fibroblasts |
| L. wallae n . sp . | smooth | 2.5 | 3 | inner: light, little collagen mid: granular with collagen outer: mostly collagen | fibroblasts |
| L. kenti $\mathrm{n} . \mathrm{sp}$. | undulating | $\begin{aligned} & 1.3 \\ & -3.6 \end{aligned}$ | 3 | inner: dark, amorphous mid: granular, vesicles, collagen outer: light, much collagen | 600 nm vesicles |
| L. lingcodae $\mathrm{n} . \mathrm{sp}$. | finely undulating | $\begin{aligned} & \hline 0.34 \\ & -0.5 \end{aligned}$ | 1 | evenly light, finely granular material, 2 membranes around outer wall | 50 nm vesicles |
| L. richardi n. sp. | smooth | 0.24 | $\begin{aligned} & 1 \\ & \text { or } \\ & 2 \end{aligned}$ | evenly light, finely granular, 2 membranes on outer wall, sometimes a middle membrane | $\begin{aligned} & 50 \mathrm{~nm} \\ & \text { vesicles } \end{aligned}$ |



Figure 3.2: Mean xenoma size in new Loma species. Error bars represent 95\% confidence intervals on either side of the mean. cvs = xenomas in the central venous sinus, $2 \mathrm{nd}=$ xenomas at the base of secondary lamellae. Number of xenomas measured: L. pacificodae n . sp. $\mathrm{cvs}=43$, $L$ pacificodae $\mathrm{n} . \mathrm{sp} .2 \mathrm{nd}=15, L$. wallae $\mathrm{n} . \mathrm{sp} .=25$, $L$. lingcodae $\mathrm{n} . \mathrm{sp} .=56$, and $L$. richardi n . $\mathrm{sp} .=5$.


Figure 3.3: Mean tubule-filled vesicle size and number of tubules per vesicle in new Loma species. Actual tubule number is 10 x the value on the y -axis. Error bars represent $95 \%$ confidence intervals on either side of the mean. Number of vesicles counted for size measurements: L. pacificodae $\mathrm{n} . \mathrm{sp} .=16, L$. wallae $\mathrm{n} . \mathrm{sp} .=10, L$. kenti $\mathrm{n} . \mathrm{sp} .=15$, L. lingcodae $\mathrm{n} . \mathrm{sp} .=19, L$. richardi $\mathrm{n} . \mathrm{sp} .=16$. Number of vesicles counted for tubule numbers: $L$. pacificodae $\mathrm{n} . \mathrm{sp} .=33$, L. wallae $\mathrm{n} . \mathrm{sp} .=23, L$. kenti $\mathrm{n} . \mathrm{sp} .=0$, L. lingcodae $\mathrm{n} . \mathrm{sp} .=28, L$. richardi $\mathrm{n} . \mathrm{sp} .=20 .{ }^{*}$ For $L$. kenti $\mathrm{n} . \mathrm{sp}$. vesicles did not contain tubules. ves size $=$ vesicle size, tu per ves = tubules per vesicle.


Figure 3.4: Mean number of tubule-filled* vesicles per sporoblast and spore in new Loma species (* in L. kenti n. sp. there were no tubules). Error bars represent $95 \%$ confidence intervals on either side of the mean. Number of sporoblasts and spores (respectively) counted were: L. pacificodae $\mathrm{n} . \mathrm{sp} .=7$ and $10, L$. wallae $\mathrm{n} . \mathrm{sp} .=9$ and $14, L$. kenti $\mathrm{n} . \mathrm{sp} .=8$ and $12, L$. lingcodae $\mathrm{n} . \mathrm{sp} .=9$ and $10, L$. richardi $\mathrm{n} . \mathrm{sp} .=4$ and 8 . ves per $\mathrm{sb}=$ vesicles per sporoblast, ves per spo $=$ vesicles per spore.


Figure 3.5: Mean number of tubules per sporoblast and spore parasitophorous vacuole (PV) in new Loma species. Error bars represent $95 \%$ confidence intervals around the mean. Number of sporoblast and spore (respectively) PV spaces counted were: L. pacificodae n. sp. $=4$ and $12, L$. wallae $\mathrm{n} . \mathrm{sp} .=7$ and $10, L$. kenti $\mathrm{n} . \mathrm{sp} .=0$ and $0^{*}, L$. lingcodae $\mathrm{n} . \mathrm{sp} .=7$ and $23, L$. richardi n . $\mathrm{sp} .=4$ and 16. ${ }^{*}$ For $L$. kenti n. sp. there were rarely any tubules in PV spaces. tu per $\mathrm{sb}=$ number of tubules per sporoblast, tu per spo = number of tubules per spore.


Figure 3.6: Mean spore size from glutaraldehyde fixed, resin (Spurr's) embedded spores in semi-thin sections for new Loma species. Error bars represent 95\% confidence intervals around the mean. Number of spores measured: L. pacificodae $\mathrm{n} . \mathrm{sp} .=12, L$. wallae $\mathrm{n} . \mathrm{sp} .=11, L$. kenti n. $\mathrm{sp} .=12, L$. lingcodae n. sp. $=10, L$. richardi $\mathrm{n} . \mathrm{sp} .=10$.


Figure 3.7: Mean spore sizes from resin embedded, histological, and ethanol-fixed material converted into estimated fresh spore size using conversion factors for new Loma species. Error bars represent $95 \%$ confidence intervals on either side of the mean. Numbers of spores measured from resin embedded and histological sections (respectively) were: L. pacificodae n . $\mathrm{sp} .=12$ and $10, L$. wallae $\mathrm{n} . \mathrm{sp} .=11$ and $10, L$. kenti $\mathrm{n} . \mathrm{sp} .=12$ and 12 , L. lingcodae $\mathrm{n} . \mathrm{sp} .10$ and $17, L$. richardi $\mathrm{n} . \mathrm{sp} .=10$ and 24 . Number of spores measured from ethanol fixed material was: spores from Atlantic cod $=24$, spores from haddock $=35$.


Figure 3.8: Mean number of polar filament turns in new Loma species. Error bars represent $95 \%$ confidence intervals on either side of the mean. Number of spores counted: L. pacificodae $\mathrm{n} . \mathrm{sp} .=14, L$. wallae $\mathrm{n} . \mathrm{sp} .=5, L$. kenti $\mathrm{n} . \mathrm{sp} .=12$, L. lingcodae $\mathrm{n} . \mathrm{sp} .=6, L$. richardi $\mathrm{n} . \mathrm{sp} .=$ 6.


Figures 3.9-3.12: Loma pacificodae n. sp. Fig.3.9 Light micrograph of 2 round xenomas (arrow) at base of secondary lamella of gill. Scale bar $=100 \mu \mathrm{~m}$. Fig. 3.10 Light micrograph of oblong xenoma in central venous sinus of gill. Scale bar $=100 \mu \mathrm{~m}$. Fig. 3.11. Light micrograph of spores (arrow) in gonads (resin sections) and dark pigment granules (p) suggesting this may be a melanomacrophage. Scale bar $=10 \mu \mathrm{~m}$. Fig. 3.12 Transmission electron micrograph of xenoma wall with granular material interspersed with collagenous fibers laid over a smooth plasmalemma (arrow). Scale bar $=1 \mu \mathrm{~m}$.


Figures 3.13-3.16: Loma pacificodae n. sp. transmission electron micrographs. Fig. 3.13 Xenoma wall with interdigitated plamalemma (arrows) and numerous opposing layers of collagen fibrils. Fig. 3.14 Collagen intrusion (arrow) into middle of xenoma. Fig. 3.15 Meronts or merogonial plasmodium undergoing binary fission within host rough endoplasmic reticulum (RER)(arrow). $\mathrm{n}=$ nuclei of dividing meronts. Fig. 3.16 Highly vacuolated meronts or merogonial plasmodia with earliest stages of parasitophorous vacuole formation (arrow). $\mathrm{n}=$ nuclei. Scale bars $=1 \mu \mathrm{~m}$.


Figures 3.17-3.20: Loma pacificodae n. sp. transmission electron micrographs. Fig. 3.17 Oblong, thick-walled, sporogonial plasmodium with a tubule-filled invagination (arrow). Fig. 3.18 Meronts or merogonial plasmodia at various stages of development. Early meronts within host rough endoplasmic reticulum (RER) indicated with " $n$ " for nuclei. Smoother-walled stages beginning to accumulate tubule-filled vesices (large arrows) at surfaces and developing stacked ER (e) cisternae. Dark circles (small arrow) are cross-sections of extruded polar filaments from nearby spores. Fig. 3.19 Parasitophorous vacuole (PV) formation around highly vacuolated meronts or merogonial plasmodia with thickened surface coats ( $\mathrm{n}=$ nuclei of meronts, $\mathrm{pv}=$ early PV space). Fig. 3.20 Close-up of small light, empty vesicles (arrows) gathered around meront surface. Scale bars $=1 \mu \mathrm{~m}$.


Figures 3.21-3.24: Loma pacificodae n. sp. Fig. 3.21 Transmission electron micrograph (TEM) of tubule-filled vesicles (arrow) and parasitophorous vacuole (PV) around an early spore. Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.22 TEM of tubule-filled vesicles and tubules within PV of a sporoblast (sb). Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.23 TEM showing 2 spores sharing a PV (pv). Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.24 Light micrograph of granuloma with spores (arrows) engulfed by host phagocytes. Scale bar $=10 \mu \mathrm{~m}$.


Figures 3.25-3.28: Transmission electron micrographs (TEM) and light micrograph (LM) of Loma spp. Fig. 3.25 TEM of host response to Loma pacificodae n. sp. showing fibroblasts ( $\mathrm{fn}=$ fibrobrast nuclei), lysosomes ( L ), and spores being engulfed and digested (arrows $=$ spores being digested and spore wall remnanats). Scale bar $=10 \mu \mathrm{~m}$ Fig. 3.26 TEM of gonadal tissue (ovaries) of Pacific cod with L. pacificodae n. sp. infection, showing host response ( $\mathrm{fn}=$ fibroblast nuclei), pigment granule formation (p), and spores being digested (arrows) containing lysosomes (L). Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.27 TEM of $L$. pacificodae n . sp. meronts or merogonial plasmodia (m) showing 2 double membrane-bound $75-90 \mathrm{~nm}$ vesicles with slightly dark contents (arrows) in cytoplasm of meronts (possibly analogous to mitochondrion remnants of Williams et al. 2002), and showing details of cells surface. Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.28 LM of Loma wallae n. sp. xenoma in central venous sinus of gill. Scale bar $=100 \mu \mathrm{~m}$.


Figures 3.29-3.32: Loma wallae n. sp. Fig. 3.29 Light micrograph of 2 nearby xenomas packed densely with spores (resin). Scale bar $=10 \mu \mathrm{~m}$. Fig. 3.30 Transmission electron micrograph (TEM) of multi-layered xenoma wall, showing smooth plasmalemma, collagenous layers, and fibroblast inclusions (arrow). Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.31 TEM of meront or merogonial plasmodium undergoing binary fission, without any obvious host rough endoplasmic reticulum (RER) covering (arrow). Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.32 TEM showing paired meronts or merogonial plasmodia with dense contents (packed ribosomes), surrounded by dark granular vesicles (arrows). Scale bar $=1 \mu \mathrm{~m}$.


Figures 3.33-3.36: Loma wallae n. sp. transmission electron micrographs. Fig. 3.33 Products of quadrinucleate cylindrical merogonial plasmodium - a chain of 4 sporoblasts $(\mathrm{sb})$ at similar stages of development within 2 parasitophorous vacuoles closely associated with each other. Wall layers can be seen in upper left (arrow). Fig. 3.34 Products of quadrinucleate cylindrical merogonial plasmodium - a chain of 4 early sporoblasts (es) at similar stages of development within 2 parasitophorous vacuoles closely associated with each other. Note also wall layers on right (arrow). Fig. 3.35 Numerous dark balls of amorphous material in xenoma cytoplasm (large arrows), and small tubule-filled vesicles containing few tubules (small arrows). Fig. 3.36. Balls of dark granular material (large arrows). Tubule-filled vesicles (small arrows) common around spores (s) but rare around sporoblasts (es, sb). Scale bars $=1 \mu \mathrm{~m}$.


Figures 3.37-3.40: Loma kenti n. sp. Fig. 3.37 Light micrograph of xenoma in secondary lamella of gill. Scale bar $=100 \mu \mathrm{~m}$. Fig. 3.38 Transmission electron micrograph (TEM) of xenoma periphery showing multilayered, undulating wall with a dark inner layer (large arrow) and lighter outer layers. There are also numerous vesicles containing dark amorphous material enclosed within a space in the xenoma cytoplasm (small arrows). Scale $\mathrm{bar}=1 \mu \mathrm{~m}$. Fig. 3.39 TEM of xenoma wall showing undulating plasmalemma (arrow), dark inner layer, and collagenous outer layers. Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.40 TEM of multinucleate merogonial plasmodium ( $\mathrm{n}=$ nuclei) with cytoplasm densely packed with ribosomes. Scale bar $=1 \mu \mathrm{~m}$.


Figures 3.41-3.44: Loma kenti n. sp. transmission electron micrographs. Fig. 3.41 A merogonial plasmodium undergoing nuclear division ( $\mathrm{n}=$ nuclei) with very thin surface coat (arrow) in direct contact xenoma cytoplasm. Fig. 3.42 Several merogonial plasmodia beginning to form parasitophorous vacuoles (PV) by the coalescing of dark material filled vesicles at their surfaces (arrows). Fig. 3.43 Merogonial plasmodia with dense cytoplasmic contents and thin surface coats in large PV spaces (pv) formed by dark-material-filled vesicles (arrow). Fig. 3.44 Meronts or merogonial plasmodium just after dividing within a large PV space ( $\mathrm{n}=$ nuclei). Scale bars $=1 \mu \mathrm{~m}$.


Figures 3.45-3.48: Loma kenti n. sp. transmission electron micrographs. Fig. 3.45 Meront or merogonial plasmodium dividing. Fig. 3.46 "Clover-leaf" shaped sporogonial plasmodium with thick surface coat, smooth, round shape and reduced parasitophorous vacuole (PV) space (arrow) undergoing division by budding. Fig. 3.47 Three products (sp) from budding of a sporgonial plasmodium with thick surface coats, smooth, round shape and reduced PV space. Fig. 3.48 Close-up of dark-material-filled vesicles contributing to PV space, showing granular, tubule-like structures within amorphous dark material. Scale bars $=1 \mu$ m.


Figures 3.49-3.52: Transmission electron micrographs (TEM) and light micrographs (LM) of Loma spp. Fig. 3.49 TEM of Loma kenti n. sp. showing typically 1 spore per parasitophorous vacuole (PV). Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.50 LM of Loma lingcodae n. sp. showing small, round xenomas in tips of secondary lamellae of gills. Scale bar $=100 \mu \mathrm{~m}$. Fig. 3.51 TEM of $L$. lingcodae n. sp. showing xenoma wall with finely undulating plasmalemma (arrow), and light finely grannular wall. Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.52 TEM of $L$. lingcodae n . sp. showing details of xenoma wall. Scale bar $=1 \mu \mathrm{~m}$.


Figures 3.53-3.56: Loma lingcodae n. sp. transmission electron micrographs. Fig. 3.53 A merogonial plasmodium (m) with thin surface coat (arrow) in direct contact with xenoma cytoplasm. Fig. 3.54 Details of the edge of a meront or merogonial plasmodium (m) showing loosely packed ribosomes (small arrow), in direct contact with xenoma cytoplasm ( $\mathrm{n}=$ nucleus of meront); also a long-section of a tubule (large arrow) can be seen in a graze cut of a parasitophorous vacuole (PV) space (pv). Fig. 3.55 Details of a meront showing 3 spindle plaques (s) within the double-layered nucleus (n). Small arrows show meront plasmalemma. Fig. 3.56 Amorphous granular material in membrane-bound vesicles in xenoma cytoplasm (large arrow) and numerous tubules in a tubule-filled vesicle (small arrow) that is coalescing at spore surface. Scale bars $=1 \mu \mathrm{~m}$.


Figures 3.57-3.60: Transmission electron micrographs (TEM) and light micrograph (LM) of Loma spp. Fig. 3.57 TEM of Loma lingcodae n. sp. showing 3 spores per parasitophorous vacuole (PV) showing coninuous PV spaces between spores (arrows) in 2 cases (labelled "a" and "b"). Presumably a 4th spore lies beyond the plane of section. Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.58 Close-up of meront cytoplasm from L. lingcodae $\mathrm{n} . \mathrm{sp}$. showing a double membrane-bound $75-90 \mathrm{~nm}$ vesicle (large arrow) with slightly dark amorphous contents (possibly analogous to mitochondrion remnants of Williams et al. 2002). Small arrow shows endoplasmic reticulum (ER). Scale bar $=1 \mu \mathrm{~m}$. Fig. 3.59 LM of Loma richardi n . sp. xenoma in secondary lamella of gill filament showing small size of these xenomas (resin). Scale bar $=10 \mu \mathrm{~m}$. Fig. 3.60. Close-up of xenoma of $L$. richardi n . sp.showing densely-packed mature spores. Scale bar $=10 \mu \mathrm{~m}$.


Figures 3.61-3.64: Loma richardi n. sp. transmission electron micrographs. Fig. 3.61 Periphery of a xenoma showing light, smooth xenoma wall (large arrow), and spores within large parasitophorous vacuoles (pv) containing numerous tubules (small arrows). Fig. 3.62. Details of xenoma wall showing light, amorphous contents and small vesicle inclusions within wall (arrow). Fig. 3.63 A 2-layered xenoma wall with a vesicle in wall (arrow). $\mathrm{xc}=$ xenoma cytoplasm. $\mathrm{h}=$ host cell adjacent to xenoma full of dark, oblong mitochondria. Fig. 3.64 Part of 2 merogonal plasmodia (m) with thick surface coats (large arrow) and partly coved in host endoplasmic reticulum (small arrows). Scale bars = $1 \mu \mathrm{~m}$.


Figures 3.65-3.68: Loma richardi n. sp. transmission electron micrographs. Fig. 3.65 Sporont with thick surface coat (large arrow) enclosed in a tubule-filled parasitophorous vacuole (PV) (small arrows), with cytoplasm containing numerous well-developed cisternae of endoplasmic reticulum (er). $\mathrm{n}=$ nucleus. Fig. 3.66 Stage with thick surface coat undergoing binary fission within a small PV space. Fig. 3.67 Details of membranes surrounding tubule-filled vesicles and PV spaces showing extra membranes around tubule-filled vesicles (large arrows) that are presumably part of the host endoplasmic reticulum. Fig. 3.68 Two cells at different points in their development within the same tubule-filled PV space. Scale bars $=1 \mu \mathrm{~m}$.


Figures 3.69-3.72: Loma richardi n. sp. transmission electron micrographs. Fig. 3.69 A meront ( m ) ( $\mathrm{n}=$ nucleus) with a parasitophorous vacuole (PV) beginning to form at its surface by the coalescence of light (empty) vesicles (arrows). A tubule-filled PV space (pv) is associated with the more advanced developmental stage to the lower right. Fig. 3.70 part of a merogonial plasmodium (m) with light, empty vesicles (arrows) at its surface. Fig. 3.71 Small light vesicles (arrows) at meront (m) surface. Fig. 3.72 Arrow shows 2 spores per PV. Scale bars $=1 \mu \mathrm{~m}$.


Figures 3.73-3.75: Transmission electron micrographs of Loma spp. Fig. 3.73 Loma richardi n . sp. showing large numbers of tubules in vesicles (small arrows) and parasitophorous vacuoles (PV)(large arrows). Fig. 3.74 A graze-cut of the PV space of a spore of Loma pacificodae n . sp. showing the orientation of tubules apparently about the longitudinal poles (large arrows). Large elongate structures (small arrows) are extruded polar filaments from nearby spores. Fig. 3.75. Details of polyribosomes (small arrows) lined up along posterior vacuole (pv) of a spore of Loma lingcodae n . sp . Scale bars = 1 $\mu \mathrm{m}$.

## Chapter 4: Phylogenetic species- and genusboundary tests in Loma (Microsporidia) using three genetic loci: rDNA, EF-1 $\alpha$, and RPB1.

## INTRODUCTION

Microsporidians of genus Loma Morrison \& Sprague, 1981 are small, spore forming, single-celled eukaryotes that parasitize gills or other tissues of fishes. Loma species, like related genera Glugea Thélohan, 1891 and Pseudoloma Matthews, Brown, Larison, Bishop-Stewart, Rogers \& Kent, 2001, form conspicuous, spore-filled host-parasite complexes, called xenomas, which eventually release infectious spores into the environment. Loma species cause morbidity and mortality in wild and farmed fishes worldwide (Morrison \& Sprague, 1981c; Morrison, 1983; Hauck, 1984; Poynton, 1986; Fomena et al., 1992; Shaw \& Kent, 1999) and potentially pose a threat to aquaculture on east and west coasts of Canada where they infect a range of cod and relatives (family Gadidae), salmon and trout (Salmonidae), and other marine fish species. New species of Loma have recently been observed in commercially important fishes off the west coast of Vancouver Island, British Columbia (Kent et al., 1998; described in Chapter $3^{1}$ ). These new species increase the list that presently includes 14 named Loma species (see Appendix 11), that have been distinguished based on host and geographic locality, and to a lesser extent by small differences in morphological characters (Lom, 2002; Lom \& Nilsen, 2003; and Appendix 11). Past studies have shown that morphological differences are sometimes inadequate for distinguishing valid species of Loma (Shaw et al., 1997; Lom \& Pekkarinen, 1999; Lom, 2002; Lom \& Nilsen, 2003), whereas molecular data often provide valuable diagnostic and taxonomic data. Hence, the present study aims to distinguish species using molecular sequence data from a range of Loma species, emphasizing new and poorly known species, including five species described in the previous chapter. For this purpose, this study uses a molecular phylogenetic

[^1]hypothesis-testing approach. This approach requires that species form separate clades (this is explained in further detail in Appendix 5) in addition to fitting the working species concept (defined later in this section). Molecular data are evaluated using statistical tests of alternative phylogenies developed by Shimodaira and Hasegawa (2001). Then, these results are considered in light of morphological, host and other evidence for species.

Recent reviews of the morphology (Lom, 2002) and molecular rDNA relationships (Lom \& Nilsen, 2003) of microsporidia from fishes emphasize the importance of including molecular data with all species descriptions or other studies of microsporidia. For example, molecular studies show that species with different morphology (e.g. presence or absence of xenomas, presence or absence of diplokaryotic nuclei) or different hosts or sites in the host can be very closely related, while species with similar morphology may be only distant relatives (Baker et al., 1997; Nilsen et al., 1998; Nilsen, 2000; Pomport-Castillon et al., 2000; Cheney et al., 2001; Bell et al., 2001; Lom, 2002; Lom \& Nilsen, 2003). Such results confirm the long-held observation (Sprague, 1977; Sprague et al., 1992) that the evolution of morphology, nuclear/chromosomal behaviours and life cycles of microsporidia are poorly understood even by experts in the field. Consequently, newer studies employing molecular tools are essential to help distinguish and define species and reveal relationships among species.

For Loma species in particular, molecular data may help define taxa with significant overlap in traditional species-diagnostic characters like spore size and polar filament turn number. For example, Shaw et al. (1997) distinguished two Loma species having sympatric hosts and overlapping morphology, L. embiotocia and L. salmonae, based on ribosomal DNA (rDNA) sequence differences shared across geographic localities. Shaw \& Kent (1999) and Shaw et al. $(2000 \mathrm{c})$ confirmed this sequence-based result experimentally by showing these two species could not infect reciprocal hosts. Inability to infect reciprocal hosts is good evidence of a barrier to gene flow between parasites like these, which must undergo repeated reproduction cycles within a host. Laboratory transmission studies may be impractical for hosts that are difficult to raise or maintain in the laboratory, like sablefish (Anoplopoma fimbria Pallas, 1814). More importantly, results under laboratory conditions may not realistically represent nature, so additional genetic studies from natural populations are crucial to understanding speciesboundaries.

One of the most important cases of species with such character-overlap occurs in two species of Loma from Atlantic Canada that are among the oldest and best-studied species of the genus, Loma branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981 and a close sister-
species, L. morhua Morrison \& Sprague, 1981. When Morrison \& Sprague (1981a) erected the new genus Loma, they named L. morhua from Atlantic cod (Gadus morhua L.) as its type; however, these species are so similar that Canning \& Lom (1986) considered these species synonymous, and considered the older name (L. branchialis) to represent the type species. Historically, there has been uncertainty over the validity of the $L$. morhua (i.e. the synonymy of the two species), and therefore also uncertainty over which name represents the type species of genus Loma. When genus Loma was erected, new combination L. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981 was created for a second species, from haddock (Melanogrammus aeglefinus L.), originally named Nosema branchiale Nemeczek, 1911 and later Glugea branchiale (Nemeczek, 1911) Lom \& Laird, 1976. Using transmission electron microscopy, Morrison \& Sprague (1981a; b) noticed differences in spore sizes, xenoma size, and xenoma walls between these microsporidia from host species, Atlantic cod and haddock, and used these as a basis to suggest two valid Loma species. Morrison \& Sprague (1981a) proposed that L. branchialis may prefer haddock to Atlantic cod as a host, whereas L. morhua may prefer Atlantic cod. Several other authors (see Kabata, 1959; Lom \& Laird, 1976) observed variation in spore and xenoma size that seemed to correspond with host (related gadids such as haddock, Atlantic cod, or other Atlantic cod subspecies), although sometimes there appeared to be two classes of spore size in a single host species (Kabata, 1959) perhaps depending on geographic locality.

Under the assumption that there are two valid species, historical records of $L$. branchialis from both hosts may represent a mixture of data from L. morhua and L. branchialis erroneously lumped together. However, Canning \& Lom (1986) observed that all such differences may be due to fixation method alone, hence these authors, and later Lom (2002), chose to recognize only the first named species, $L$. branchialis as valid, thus making $L$. morhua a junior synonym, and $L$. branchialis the type species. The problem stemmed partly from the fact that Morrison \& Sprague (1981a) considered the large difference in spore size and different host between Nemeczek's (1911) L. branchialis and their material (L. morhua) to be sufficient to erect a separate species, but were uncertain about the identity of any of their material with original records for $L$. branchialis because their spores were never as large as those reported by Nemeczek (1911) (Morrison \& Sprague, 1981a, b; Morrison, 1983; Morrison \& Marryatt, 1986). Furthermore, the original specimens of $L$. branchialis were collected at two distant locations, northeastern North America and Russia, and Nemeczek (1911) did not specify which locality represented the type locality or the locality from which spore measurements were taken.

Kabata (1959) suggested Nemeczek's (1911) material might represent two species of Loma, though it is difficult to evaluate this suggestion, as Nemeczek's (1911) material is no longer available, and his single measurement of "typical spore size" cannot be statistically re-analyzed. The present study investigates spore shrinkage due to fixation, to compare material collected here to that in reported in previous studies, and combines this data with molecular data to assess the validity of the two species $L$. branchialis and $L$. morhua, at least one of which must be the type species of the genus Loma.

Recent studies also suggest genus-boundaries for Loma and related genera require reexamination. For example, a molecular study by Lom \& Nilsen (2003) demonstrated that $L$. acerinae (Jirovec, 1930) Lom \& Pekkarinen, 1999, a species recently transferred to this genus based on morphological similarity, is probably not a valid member of the genus. Hence, the genus will be re-examined using sequences from all species of Loma for which molecular data are available, including species available here (above) and Loma sp. of Nilsen (2000) from the fourbeard rockling, Enchelyopus cimbrius (L.), and L. acerinae (Jirovec, 1930) Lom \& Pekkarinen, 1999 (from ruffe, Gymnocephalus cernuиs L.). While the "genus group" is more ambiguous in its essential meaning than the "species group", genus boundaries can be examined here using similar criteria as for species boundaries, wherein members of a genus should form a supported monophyletic clade; however, this clade must also include the type species. Molecular data from these species may also be used to infer the relationships between these species to examine biogeographic, host-parasite co-evolutionary and morphological evolutionary patterns. If Loma species, like most microsporidia, are narrowly host-specific, they might be expected to co-evolve to a large extent with their hosts, and therefore host and parasite phylogenies should be largely congruent. Alternatively, phylogenies might suggest that Loma species have switched hosts. For example, recent host-switching would be implied if Loma species from Pacific salmonids and gadids were more closely related to one another than they were to Loma species from Atlantic/eastern Canadian salmonids and gadids, whereas under coevolution, Loma species from salmonids and gadids should group according to host. Morphological traits can be plotted on phylogenetic trees to look for evidence of repeated, independent loss or evolution of features, or other similar potentially adaptive changes.

Obtaining and interpreting molecular data that enable species to be characterized and distinguished is not necessarily straightforward, as it will depend on how species are to be defined, the extent of prior knowledge of a species' geographic range, and whether species are sexual or asexual. Historically such knowledge has been lacking for most microsporidians,
particularly for Loma species. Furthermore, microsporidians can be technically difficult to isolate for molecular analysis, being found intracellularly, mixed with host tissue, making it necessary to either (1) isolate spores by careful layering and centrifugation on density-specific gradients (e.g. Percoll) or (2) amplify their DNA from host tissue using microsporidian-specific primers. The first method requires large quantities of fresh infected tissue (e.g. sometimes many hosts) and often selects for the most resistant spores. These can be difficult to break open without also shearing the DNA into small, unsequenceable fragments. The second method tends to work better on material with a higher percentage of pre-spore stages as the DNA is more easily extracted from these; however, collection by light microscopy tends to select for tissue with mostly mature spores. In this case, the type of tissue fixation, DNA extraction, presence of inhibitors, and the specificity of the primers can be critical to obtaining useable results (see Appendix 20).

Even for better-known multicellular organisms, gathering sufficient data to distinguish species can be difficult. For example, studies recommend one should characterize at least five independent nuclear loci (according to Wu, 1991), at least 50 individuals (as recommended by Crandall \& Templeton, 1993), and survey individuals broadly and thoroughly across a species' geographic range (see Sites \& Crandall, 1997). However, some studies have successfully determined species-boundaries using only multi-copy ribosomal DNA (rDNA) or fewer than five independent single-copy markers (see Rodriquez-Robles \& de Jesus-Escobar, 2000; Van Oppen et al., 2000; Bradley \& Baker, 2001; Diekmann et al., 2001; Puorto et al., 2001; Bernardi et al., 2002; Chen et al., 2002; Lopez et al., 2002; Van Oppen et al., 2002). These studies used various approaches, from simple \% divergence within versus among potential species, to more sophisticated statistical approaches combining molecular and morphological data or estimating gene flow. In these studies authors usually knew something about the reproductive biology and ploidy, understood the number of copies or segregating behaviour of the molecular locus, and were able to define their species concept and their criteria for hypothesis testing (Sites \& Crandall, 1997), thereby increasing their power to analyze and interpret genetic results. Reproductive biology, genetic and chromosomal behaviour, and geographic distribution of Loma species are poorly known; therefore this study employed a simple approach to speciesboundary questions.

This approach centers on an underlying species definition presented in Chapter 3 (based on arguments in Wheeler \& Meier, 2000) as follows: the first criterion for distinguishing one species from another when in sympatry is that each must possess separate discrete characters or
overlapping characters with a statistically separate mean, and the second criterion is that the suite of characters that provides evidence for species must agree across the sampled populations. However, the working species definition for molecular sequence data examined here is slightly different. It also requires that molecular data from all isolates of a valid species must be able to form a monophyletic group. Those isolates (or sequences) that fall outside the boundaries of the proposed species-clade, rather than grouping monophyletically with the predicted clade, are not members of the species. However, sequences may fall outside the proposed clade by chance alone. So, if a proposed species- or genus-group does not form a strongly supported monophyletic clade in all analyses, isolates can be forced together to create the hypothetical clade. The forced-together clade can be compared to the un-forced arrangement using treecomparison statistics in Shimodaira \& Hasegawa (2001). Results will be evaluated by identifying monophylies or paraphylies that can be statistically rejected, thus violating the species-criteria as defined here (see Appendix 5 for further explanation). Although this approach will be used to re-evaluate molecular evidence for species and species-boundaries, it will not be used as the sole basis for describing or delimiting species. Instead, species will be defined and described using all available data (e.g. including both molecular and ultrastructural data as in Chapter 3).

While host was used to form testable hypotheses in this study, presence of the parasite in a host does not indicate specificity for this host; however, many microsporidia appear to be highly host-specific. Host is often correlated with other important features because hosts of these obligate intracellular parasites act as islands separated by temporal and spatial barriers. Proper tests of host-specificity would involve surveying all potential alternate hosts for the parasite (e.g. using genetic markers) over a range of times and locations or experimental transmission to these alternate hosts. Some potential alternate hosts have been surveyed (Kent et al., 1998; Shaw \& Kent, 1999; and Chapter 3) and show that Pacific Loma species may be rather host-specific. In contrast, transmission experiments have been difficult, and have been informative for only a few Loma species (Shaw et al., 1997; Shaw \& Kent, 1999; Shaw et al., $2000 \mathrm{a} ; \mathrm{b} ; \mathrm{c}$ ). This is because they require significant effort and may not be possible for some species. Yet, to determine whether host-correlated morphological differences reflect valid characters of the parasite (i.e. not phenotypically plastic, depending on the host), ideally we should test host specificity. Here, molecular data are used to examine host-use.

As suggested by Sites \& Crandall (1997), our ability to assess biologically meaningful species-boundaries depends critically on the thoroughness of sampling is across the range.

Similarly, our ability to evaluate genus-boundaries will depend on how broadly species are sampled across the genus. Therefore, in this study, Loma species were collected from as wide a range of hosts and localities as possible. Species sampled include L. branchialis and L. morhua, and several others with overlap in morphological, host or geographic features (see Table 1.1 in Chapter 1, and Appendices 11 \& 12), including seven undescribed Loma species, Loma sp. from brook trout (Salvelinus fontinalis Mitchill, 1814), Loma sp. from surf bream (Acanthopagrus latus Günther, 1859), Loma sp. from Pacific cod (Gadus macrocephalus Tilesius, 1810), Loma sp. from Pacific tomcod (Microgadus proximus Girard, 1854), Loma sp. from walleye pollock (Theragra chalcogramma Pallas, 1811), Loma sp. from lingcod (Ophiodon elongatus Girard, 1854), and Loma sp. from sablefish (Anoplopoma fimbria Pallas, 1814), two described species, L. salmonae (from salmon and trout, Oncorhynchus spp.) and L. embiotocia (from shiner perch, Cymatogaster aggregata Gibbons, 1854), and the type species and its close sister species (or conspecific) from Atlantic gadids. Wherever possible, several isolates were collected across the geographic range of the species. An "isolate", which consists of a DNA sample from a piece of infected host tissue, could be a mixture of genetically similar lineages or several distinctly different strains. However, PCR and cloning of single-gene copies in this study allows "individuals" and individual copies of multi-copy genes to be separated from these multi-lineage isolates prior to analysis.

Lower-level molecular systematics are particularly susceptible to the effects of factors that cause gene-trees and species-trees to be incongruent, for example, incomplete lineagesorting, therefore, the most widely accepted solution is to examine several different independent loci. For the interspecific level (species-boundaries), two independent loci (including four gene regions) were chosen because they could be easily compared with published studies, and because of their potentially high sequence variation: ribosomal small subunit (SSU), internal transcribed spacer (ITS) and partial large subunit (LSU) genes, and partial elongation factor1alpha (EF-1 $\alpha$ ) gene (Cho et al., 1995; Kamaishi et al., 1996; Moreira et al., 1999; Weiss \& Vossbrinck, 1999). For the intra- and inter-generic level (genus-boundaries) ribosomal DNA (rDNA) was examined along with another independent locus known to produce a phylogeny consistent with morphological and rDNA phylogenies: RNA polymerase largest subunit II (RPB1) (Cheney et al., 2001).

Consequently, species- and genus-boundaries were examined for Loma species, employing the criteria for species discussed above. The first goal was to evaluate the validity of L. morhua as a separate species from L. branchialis using molecular and morphological data
from spores, and estimating spore shrinkage so these can be compared to previous studies. The second goal was to test the validity of described and undescribed Loma species, including five new species described in the previous chapter ${ }^{2}$, using two genetic loci, rDNA and EF-1 $\alpha$. The third goal of this study was to investigate genus-boundaries for Loma using partial SSU rDNA and RPB1 gene sequences from members of genus Loma and related genera. Finally, this study examines relationships among species of Loma, and compares these with what is known of the host relationships, to infer patterns in biogeography and possible host-parasite co-evolution.

[^2]
## MATERIALS AND METHODS

## Specimen collection

Six species of Oncorhynchus (Pacific salmon and rainbow trout) and nine other fish species with Loma-like infections (Table 4.1) were collected for DNA analysis of the parasites. See Appendix 12 for name equivalences with those species described in Chapter 3. Where possible, several host individuals or hosts from different geographic localities within the species' range were collected in order to look for genetic variation across this range. Isolates (individual hosts) from which DNA sequences were obtained are listed in Table 4.1 (see Appendix 13 for key to isolate names). All fishes collected, unless otherwise specified (below), were killed with a blow to the head before gills and sometimes other tissues were removed with a fresh, pre-sterilized blade, and examined under dissection microscopy for the presence of xenomas and fixed in $95 \%$ ethanol for DNA analysis. Care was taken to bleach or ethanolwash, then ethanol-flame all forceps proceeding and following handling of each individual tissue sample, to prevent cross-contamination among specimens.

Loma salmonae isolates from salmon and rainbow trout (Oncorhynchus spp.) in the laboratory (Pacific Biological Station (PBS), Nanaimo, British Columbia, Canada) were maintained as described in Kent et al. (1995), Shaw et al. (1998) and Shaw et al. (2000a), and gills were either fixed in ethanol or homogenized and run through Percoll (SIGMA, Ronkonkoma, NY) gradients to purify spores. Other isolates from salmon and trout were collected from the wild by hook and line, seine, or bottom trawling from several localities in British Columbia, or were kindly donated from farms and hatcheries in British Columbia and elsewhere (Table 4.1).

Loma species from several other fishes (Pacific cod Gadus macrocephalus Tilesius 1810, walleye pollock Theragra chalcogramma (Pallas 1811), Pacific tomcod Microgadus proximus (Girard 1854), lingcod Ophiodon elongatus Girard 1854, sablefish Anoplopoma fimbria (Pallas 1811), and shiner perch Cymatogaster aggregata Gibbons 1854) were collected at sea by bottom trawling from several localities off the south, west and east coasts of Vancouver Island, British Columbia. Trawling covered an area from 3 to 48 km distance from shore, 65 to 230 m depth, over a north-south distance of about 300 km (from Juan de Fuca Strait lat/long 48.15 N 124.00

W, to the region off Barkley sound lat/long 48.50 N 125.20 W and up to Queen Charlotte Sound lat/long 51.20 N 129.00 W ).
D. Barker (Marine Institute, Newfoundland, Canada) provided Loma species from Atlantic cod Gadus morhua L. near St. John's, Newfoundland. These fish were presumably infected from local wild sources while being held in sea net pens. Other Atlantic cod and haddock Melanogrammus aeglefinus (L.) were caught aboard commercial vessels in the area approaching Halifax harbour, Nova Scotia, Canada, heads with gills were purchased from a market in Halifax, Nova Scotia.
J. G. Sánchez (Atlantic Veterinary College, Prince Edward Island, Canada) donated Brook trout Salvelinus fontinalis (Mitchill, 1814) gills with Loma-like infections. These infected gills were obtained from brook trout that underwent experimental exposure (per os) to Loma salmonae spores originally isolated from laboratory salmon in British Columbia (PBS), and became infected in a third trial following two previous, identical trials in which fish did not exhibit infection. These brook trout were raised in the laboratory as part of a separate study (Sánchez et al., 2001a), and reportedly originated from a hatchery in Prince Edward Island (P.E.I.), Canada with no history of infection. The fish showed no signs of infection (xenomas) in the gills when initially examined live, by the naked eye or by low power dissecting microscope, according to those authors.
R. Adlard provided gills from a Loma species from Australian surf bream, Acanthopagrus australis (Günther, 1859), in Queensland, Australia.

Two other microsporidians were also provided by M. L. Kent and used in DNA sequencing and analysis as outgroups: Microsporidium prosopium Kent, Docker, Khattra, Vossbrinck, Speare \& Devlin, 1999 from mountain whitefish Prosopium williamsoni (Girard, 1856) from British Columbia, Canada, and an unnamed Glugea species from starry flounder Platichthys stellatus (Pallas) in Oregon, USA, in which will be presently referred to as Glugea sp. STAR in this study.

## Spore size measurements for Loma spp. in Atlantic cod and haddock

Spores from haddock and Atlantic cod in Halifax, Nova Scotia, Canada from which DNA sequences were obtained in this study (sequences labeled L. branchialis HAD and L. morhua ATL, respectively) were measured and compared to sizes reported in other studies (see Table 4.3). Measurement error was reduced by photographing spores under 1000 X magnification under light microscopy, scanning the photographs, and enlarging the spore images to $\sim 15 \mathrm{~cm}$ for
measurement on screen. From L. morhua (in Atlantic cod), 24 spores were measured, and from L. branchialis (in haddock) 35 spores were measured, and $95 \%$ confidence intervals were calculated around the mean. Because spores were fixed by a variety of methods (frozen, formalin, glutaraldehyde fixed and resin-embedded) in previous studies it was important to consider the effects of shrinkage due to fixation. Shrinkage due to fixation was estimated for this material by measuring spores from several species of Loma that had been fixed by more than one method, and inferred shrinkage from the differences in mean spore sizes, then combined the estimated shrinkage for each pair of fixation methods across all species of Loma. These estimated shrinkage factors were used to estimate fresh, frozen, and resin-embedded spore sizes from the $95 \%$ ethanol fixed spores from Atlantic cod and haddock. Length and width shrinkage estimates were always calculated separately.

## DNA isolation

DNA was isolated from about 60 mg of ethanol-fixed tissue from a single gill arch or an equivalent amount of other tissue for each fish, following a pre-soak for 15 minutes in lysis buffer ( 10 mM Tris, 1 mM EDTA, $10 \mathrm{mM} \mathrm{NaCl}, 1 \% \mathrm{SDS}$ ) to remove excess ethanol. DNA was also isolated from single xenomas from both frozen and ethanol-fixed tissue for some isolates. To isolate single xenomas, gills with light infection (few xenomas) were examined carefully with both incident and transmitted light under a dissecting scope to find xenomas well separated from neighbouring xenomas. Single xenomas were excised along with some host tissue using a fresh, sterile blade each time, and then were examined under a cover slip at 400X magnification to confirm the absence of other, small xenomas in adjacent tissue. DNA isolation from purified spore concentrates required bead beating, following the procedure of Docker et al. (1997a), to break open spores prior to DNA extraction.

Bead-beaten spores, ethanol-fixed gills or single xenomas were then digested in 5-10x volume lysis buffer with $0.5 \mathrm{mg} / \mathrm{ml}$ proteinase K for $4-6$ hours at $37^{\circ} \mathrm{C}$ in a rotating incubator. DNA was extracted once with phenol, twice with phenol:sevag (phenol: chloroform: isoamyl alcohol 25:24:1) and once with sevag, precipitated in cold $95 \%$ molecular-grade ethanol, washed twice with $70 \%$ ethanol, vacuum dried, and the pellet resuspended in $40 \mu \mathrm{l}$ distilled water and stored for use at $-20^{\circ} \mathrm{C}$. All stock reagents and templates after the first phenol stage were handled using fresh, pre-sterilized aerosol tips, and extractions were performed in small batches with periodic cleaning of equipment and surfaces to reduce the chance of contamination among samples.

## Polymerase chain reaction (PCR)

Primers for PCR amplification of ribosomal DNA (rDNA), elongation factor-1 alpha (EF-1 $\alpha$ ), and RNA polymerase largest subunit (RPB1) are shown in Table 4.2. Primers designed in this study were created using alignments of a large number of microsporidian and non-microsporidian sequences from Genbank, and designed to match conserved regions within only the microsporidia, where possible, or conserved regions across a larger group of available taxa if no reasonably general microsporidian-specific priming sites were apparent.

PCR was performed in a Perkin Elmer Cetus DNA Thermal Cycler 480 in $25 \mu \mathrm{l}$ reactions with roughly 0.3 to $0.8 \mu$ g genomic DNA, standard PCR buffer, $2.5 \mathrm{mM} \mathrm{MgCl}_{2}, 0.2$ mM dNTP, 15 pmol of each primer, and 1-3 units of Taq DNA polymerase (Invitrogen Corp., Carlsbad CA). Conditions were: $95^{\circ} \mathrm{C}$ for 1 min , followed by 35 cycles of $95^{\circ} \mathrm{C}$ for $50 \mathrm{sec}, 54$ ${ }^{\circ} \mathrm{C}$ for 30 sec (or as low as $50^{\circ} \mathrm{C}$ for difficult amplifications), $72^{\circ} \mathrm{C}$ for 90 sec , and a final extension of $72{ }^{\circ} \mathrm{C}$ for 5 min . In all PCR-runs pre-sterilized aerosol plug-tips were used for all template and reagent handling, and all runs included a positive control (microsporidian DNA) and a negative control (water) to check for contamination. PCR products were visualized in $1.5 \%$ agarose TBE or TAE gels, and the correct-sized product was then excised and freeze-thaw extracted or digested with $\beta$-agarase remove agarose for direct sequencing or cloning.

## Cloning

PCR products were isolated in $0.8 \%$ agarose and cleaned for ligation using Ultraclean 15 MOBIO DNA Purification Kit (BIO/CAN Scientific Inc. Mississauga, ON) and cloned using the TOPO TA Cloning PCR Version 2.1 (Invitrogen Corp., Carlsbad, CA) using $1 / 2$ volume. Clones were screened for presence of the insert in $10 \mu \mathrm{l}$ PCR reactions using Taq DNA Polymerase (Invitrogen Corp., Carlsbad, CA) with standard reagents and screening primers M13-20 and M13 Rev (conditions: $94^{\circ} \mathrm{C}$ for $2 \mathrm{~min}, 34$ cycles of $92{ }^{\circ} \mathrm{C}$ for $45 \mathrm{sec}, 55^{\circ} \mathrm{C}$ for 45 $\sec 72^{\circ} \mathrm{C}$ for 1 min 30 sec , followed by $72^{\circ} \mathrm{C}$ for 5 min ). Positive clones from master plates were grown in 3 ml of standard LB culture with 50 mM ampicillin by shaking at 220 rpm at 37 ${ }^{\circ} \mathrm{C}$ overnight. Plasmids were isolated for sequencing using the Rapid Plasmid Miniprep System (Gibco BRL, Gaithersburg, MD) following directions of the manufacturer.

## DNA sequencing

Sequencing was performed on the ABI PRISM 377 DNA automated sequencer using BigDye Terminator Version 3.1 fluorescent dye-labelled terminators with forward and reverse
primers and PCR conditions as recommended for the Taq terminators. Wherever possible, PCR products were sequenced in both directions and multiple PCR products and multiple clones were sequenced to check for Taq or sequencer errors.


#### Abstract

Alignment Ribosomal DNA and EF-1 $\alpha$ sequences from Loma species were easily aligned by eye using ESEE 3.2s (Eyeball SEquence Editor, Eric Cabot, 1998), except for parts of the ITS rDNA region, which were aligned with the help of Clustal W (version 1.74) allowing for frequent transitions between closely related taxa. Ribosomal DNA, EF-1 $\alpha$ and RPB1 sequences from Genbank that were used in phylogenetic analyses are listed in Table 4.13. These sequences from a variety of other microsporidians were aligned first by grouping them into clusters of taxa considered to be closely related from previous phylogenetic studies and then by aligning the more conserved regions by eye (ESEE 3.2s). RPB1 DNA sequences were then better aligned with an emphasis on amino acid sequence conservation. To reduce both alignment difficulty and phylogenetic analysis time, a preliminary phylogenetic analysis was performed to help choose taxa of interest and to eliminate taxa that were over-represented or fell into clades distant to the species of interest. Finally, less conserved regions of rDNA were aligned using various parameters in Clustal W, allowing frequent gaps near rDNA loop regions and emphasizing transitions over transversions for taxa that appear more closely related from conserved alignment regions.


## Flip algorithm

Sequencing directly from PCR products sometimes produced sequences that were double, having two different nucleotide signals at almost every site after a stretch of normal single nucleotide signals. Such stretches of double-nucleotide signal can arise from an indel difference between two copies (alleles) of a gene, where both copies have been sequenced simultaneously in the same sequencing reaction. Selection of the correct peak corresponding to each copy from the double peak region is difficult because peak heights typically vary more within each gene copy than between the two copies. Correct signals from the doubled portion of sequences were deduced using Flip Analyzer of REALEM Version 1.01 (developed this study), which swaps top and bottom peaks at each site, searching for a match between the sequences, and then identifies the likelihood of an indel of a user-specified size range. For more details, see Appendix 1.

## Polymorphic sites

Polymorphism (presence of several alleles in the population) can produce polymorphic sites in directly sequenced PCR products, in which two or more nucleotide signals appear at single site. Because these polymorphic sites may result from background contamination or sequencing artefacts, or alternately, may be valid polymorphisms in the data resulting from substitutional differences between two copies of the gene, the original sequence data was carefully re-examined for evidence of background signal (a low-read set of peaks out of alignment with the main signal) or sequencer software error (identified by examining chromatographs). In cases where these artefacts could not be distinguished from clear signals, sequencing was repeated. Where there was valid polymorphism in the data, it was counted as a partial substitutional difference between two sequences and given half the value of a full substitution in calculating percent sequence differences if it was of form "A/T vs. A", or the value of a full substitution if it was of form "A/T vs. G".

## Phylogenetic analysis

Phylogenetic analysis was performed in PAUP* Version 4.0b10 (Phylogenetic Analysis Using Parsimony; Swofford, 2001), using all three types of optimality criteria, maximum parsimony (MP), minimum evolution using distance (ME) and maximum likelihood (ML) with heuristic search, random stepwise sequence addition and TBR branch swapping with 10 repetitions, or only one repetition of NNI branch swapping for larger data sets in ML. For unweighted maximum parsimony and distance I used the logDet/paralinear model (which performs well with unequal base frequencies, which were found frequently in these data sets) to calculate pairwise distances for trees estimated under the minimum evolution (ME) criterion. Statistical support was assessed by bootstrap resampling with 1000 replicates reported on $50 \%$ majority rule trees. Maximum likelihood analysis was performed by first estimating the best-fit model of evolution for each input data set, as predicted under the Akaike Information Criterion (AIC) method (Akaike, 1974) using Modeltest Version 3.06 (Posada \& Crandall, 1998). Maximum likelihood heuristic searches were run using the best-fit model and parameters estimated from the data, first to generate a starting tree with neighbour-joining under maximum likelihood, then to re-estimate the parameters and heuristically search for maximum likelihood trees. (For further details see Appendices 2 to 4.) Bootstrap assessment consisted of 100 replicates either by ML faststep for smaller data sets or NJ searches for larger data sets. Outgroups were chosen based on preliminary analyses which suggested basal taxa, and were
applied for tree display purposes only. Outgroups were Loma sp. AUS or L. salmonae, or M. prosopium or $S$. cerevisiae.

Several different data sets were created for each alignment using the character exclusion feature of PAUP* to analyze the effects of missing and ambiguous character removal or gap removal (gap stripping). Additional data sets were created using Gap Matrix of REALEM Version 1.01 (explained in Appendix 1) to analyze the effects of adding a matrix containing gap information to either the full alignment or to the gap stripped input, as well as to examine a phylogeny estimated from gap information only.

A composite of rDNA data for all Loma species examined here was created for display purposes as shown in Figure 4.20. The tree of Loma species shown in Figure 4.20 was reconstructed from a single, composite reference sequence for each species, 1113 to 1846 bp long, and covering the full SSU, ITS and partial LSU rDNA regions. Nucleotide states of these sequences were estimated from the majority nucleotide state at each site taken from all available sequence fragments from all isolates of a species. Note that these composite sequences may combine data from different rDNA paralogs and different variants (or isolates) of a species, and should be interpreted with this in mind. Phylogenetic analysis was performed with heuristic ML search in PAUP*, as described above.

## Monophyly constraints and AU tests

Alternate hypothetical trees were created that place sequences ("taxa") together in monophyletic groups of the same Loma species, same genus, or same host family (see explanation in Appendix 5). Trees were viewed on slanted cladograms for easier recognition of topological differences. These trees were created by modifying parentheses in text versions of trees from TreeView Version 1.6.6 (Page, 1996) obtained from PAUP* analyses. The parenthetic trees were pasted into PAUP* commands as constraints prior to heuristic distance searches. Distance trees to be compared were inspected to be sure they had equal numbers of nodes (and therefore branches), and were topologically alike in clades outside the modified clades, to obey the assumptions of the statistical tests. Likelihood and parsimony trees created in this way often had different degrees of polytomy (multifurcating nodes), and therefore different node numbers, and so could not be easily used. Log likelihood values for individual sites (lscores) were calculated for the set of shortest trees, using the best-fit DNA substitution model and parameters estimated for the data (Modeltest). Likelihood scores for DNA sites (lscores) and overall tree likelihoods from these unconstrained and constrained trees were then
analyzed by statistics in the CONSEL version 0.1f (Shimodaira \& Hasegawa, 2001) software package. This package statistically compares trees using the Approximately Unbiased test (AU test), and other similar tests (see Goldman et al., 2000; Shimodaira \& Hasegawa, 2001; Shimodaira, 2002). For further details see Appendices 5 to 9. Further justification for this approach is given in Appendix 6.

## Tests for recombination

Recombination among diverged rDNA copies within a genome, or "paralogs", was examined using rDNA sequences using SplitsTree (Huson, 1998) and LARD Version 2.2 (Likelihood Analysis of Recombination in DNA; Holmes et al., 1999).

## RESULTS

## Spore size comparisons for Loma spp. in Atlantic cod and haddock

Spores from Atlantic cod $(\mathrm{n}=24)$ and haddock $(\mathrm{n}=35)$ were similar in size to those from Atlantic cod, being slightly longer (student's t-test p-value $=0.007$ ) and wider (width difference not statistically significant). Xenomas were not measured, but were observed, qualitatively, to be larger in haddock than in Atlantic cod.

Spore sizes were converted using shrinkage factors calculated as described in Materials \& Methods, and were compared to six out of eight studies where spore sizes were reported from L. branchialis and L. morhua, from various localities and hosts (Table 4.3). Both Atlantic cod and haddock Loma species' spores were smaller than frozen spores (type specimens of $L$. branchialis) from haddock in northeastern North America and/or Russia measured by Nemeczek (1911), and smaller than spores (type host, type locality of L. morhua) embedded in resin from both Atlantic cod and haddock in Halifax, Nova Scotia, Canada measured by Morrison \& Sprague (1981a; b) (Table 4.3). Morrison \& Sprague (1981a) used EPON resin, whereas Spurr's resin used in the present study may shrink differently. Both Atlantic cod and haddock derived spores in the present study were larger than fresh spores from haddock in New Brunswick, Canada, measured by Lom \& Laird (1976) (Table 4.3). These spores were about the same size as fresh spores from Atlantic cod in the Gulf of St. Lawrence, Canada measured by Fantham et al. (1941) and about the same size as frozen spores from Kildensis cod in the Barents Sea measured by Dogiel (1936) (Table 4.3). Because spores were not formalin-fixed in the present study, shrinkage-corrected sizes could not be compared for two other studies (Kabata, 1959 and Shulman \& Shulman-Albova, 1953); however, these were perhaps less relevant because spore sizes were in the middle-range of those listed above (see Table 4.3), and were from more distant localities, the Barents Sea and White Sea.

## Ribosomal DNA (SSU, ITS and LSU) sequence characteristics

Table 4.1 shows species and isolates sequenced, Appendix 12 shows a table of species name equivalents between chapters, and Appendix 13 shows a key to help read isolate labels used throughout the results. Amplifying long regions of rDNA (e.g. the whole region between the M5P and 580r primers, Table 4.20) was technically too difficult for most specimens. Therefore, rDNA was amplified in shorter segments. These segments could not be analyzed as though they were continuous (as contigs) because they might be from different, paralogous
cistrons. Thus, the rDNA region was divided into sections prior to phylogenetic analysis as shown in Figure 4.1. These rDNA sections were as follows: 5' half of SSU rDNA, positions 20 to $662=$ Region $1 ; 3^{\prime}$ half of SSU rDNA, positions 724 to $1351=$ Region 2; $3^{\prime}$ end of SSU rDNA, ITS, and 5' end of LSU, positions 1352 to $1830=$ Region 3.

The length of ribosomal DNA (rDNA) sequenced over the total cistron, or individual regions (SSU, ITS or LSU) varied among species (Table 4.4). Intraspecific sequence differences within PCR products sequenced directly or cloned were observed in all isolates of Loma species listed in Table 4.1 with the exception of Loma sp. AUS (an undescribed species from Australia), as summarized in Table 4.5. Intraspecific differences (Table 4.5) from all rDNA tended to be greatest in Loma species from gadid hosts (L. branchialis HAD, L. morhua ATL, and Loma sp. PAC and POL), and lowest in L. salmonae. Pairwise distances between species were estimated for reference sequences only (produced from the majority nucleotide state for all clones and isolates of each species), for the whole rDNA cistron (Appendix 14). While some of this intraspecific difference consisted of nucleotide substitutions, there were also indel differences (summarized in Table 4.6), and "polymorphic differences" or half differences, as defined in the Materials and Methods section, above (summarized in Table 4.7). Table 4.6 shows that all these Loma species other than Loma sp. AUS had at least one intraspecific indel difference, and that many indel differences were parsimony informative either among species or within species, where parsimony informative characters were defined as cases in which at least two taxa bear each state. This table also shows that the intra- and interspecific indel differences were distributed through all three rDNA regions (SSU, ITS, and LSU), and varied in size all regions. Often these intraspecific indels made sequencing directly from PCR products difficult, yielding signals that became double after an indel site. In such cases, FlipAnalyzer in REALEM (see Appendix 1) was used to extract sequences. In some cases, single xenomas produced such intra-PCR product indel differences. Table 4.7 shows only the polymorphic differences that were shared (at least two sequences bearing each character state) or were additive within a species, where additive sites are those defined as having all three states: two single characters and one double-signal with both single characters (a polymorphic difference). There was intraspecific additivity in all three rDNA regions, and in most species (L. morhua ATL, Loma sp. PAC, Loma sp. POL, Loma sp. TOM, Loma sp. LIN, L. salmonae and L. embiotocia) with L. morhua ATL having the most additive sites.

## Elongation factor-1 alpha (EF-1 $\alpha$ ) sequence characteristics

In total, 1007 nucleotide and 335 amino acid alignment positions of the EF-1 $\alpha$ gene were sequenced. The number of positions sequenced per Loma species varied as shown in Table 4.4. Intraspecific variation (Table 4.5) was highest in this gene in L. morhua ATL and Loma sp. PAC. While intraspecific variation was lowest in L. salmonae and Loma sp. TOM, this level was higher than that expected by Taq error (i.e. 1 in 1000). In total, sequence was obtained for 11 isolates of $L$. salmonae, three isolates of L. morhua ATL and Loma sp. POL, two isolates of Loma sp. TOM and Loma sp. PAC, and one isolate of Loma sp. BRO, L. embiotocia, and Loma sp. LIN. Within isolates, sequence variation tended to be low (usually zero), presumably because this is a single-copy gene, so multiple sequences from each isolate that were identical in the overlap were fused (for consensus regions) prior to phylogenetic analysis. Reference nucleotide and amino acid sequences (Appendices $15 \& 16$ ) that were taken from the majority character at each site (for details see Materials and Methods) showed that most substitutions were synonymous, and \% GC content was similar in all species of Loma ( 37 to $38 \%$ ).

## RNA polymerase largest subunit (RPB1) sequence characteristics

A region 836 nucleotides long ( 278 amino acids) from the 5 ' region of RNA polymerase largest subunit gene was sequenced in one isolate of $L$. salmonae (isolate BA9). This region corresponds to positions 241 to 1064 in Loma acerinae Genbank Accession AJ278951, beginning at amino acid 87 . The sequence (Appendix 17) had $36.1 \%$ GC content. This DNA sequence was 29.1 \% different from Glugea anomala Genbank Accession AJ278952 (20.1 \% amino acid difference), and $32.2 \%$ different from L. acerinae ( $22.7 \%$ amino acid difference). By comparison, G. anomala and L. acerinae were $26.6 \%$ different ( $13.7 \%$ amino acid) across the same alignment region, and were both higher in $\% \mathrm{GC}$ ( 38.4 and 42.8 , respectively).

## Phylogenetic relationships among Loma species for rDNA Region 1 (5' half of SSU)

Figure 4.2 combines results of distance bootstrap, maximum parsimony, and maximum likelihood analyses using the best-fit model from Modeltest (F81+G), for rDNA Region 1 (5' half of SSU), on a single distance tree. Some branches varied (largely at tips) such that there were 35 most parsimonious trees, 15 shortest distance trees, and one maximum likelihood tree. All three types of analysis consistently produced several branches (shown in bold, Fig. 4.2), which include those forming monophyletic clades for the individual species: L. salmonae, Loma sp. TOM, Loma sp. POL, Loma sp. BRO, and Loma sp. LIN. Bootstrap support from all three analysis methods was high for four of these species (L. salmonae, Loma sp. LIN, Loma sp.

TOM, and Loma sp. BRO). Two clades were observed consistently, clades "P" (including Loma sp. POL and Loma sp. PAC) and "A" (including clade P and Loma sp. ATL), although bootstrap support was generally low in this part of the tree, being at most $80 \%$ ML ( 100 replicates) for clade A. Addition of a gap matrix to the input data, consisting of seven additional characters, and adding six parsimony-informative characters to the analyses, produced no remarkable changes in topologies from the three methods, and only small changes (usually less than 5\%) in bootstrap support.

## Phylogenetic relationships among Loma species for rDNA Region 2 (3' half of SSU)

Figure 4.3 shows results from rDNA Region 2 ( $3^{\prime}$ half of SSU) by analyses identical to those presented in the previous section. The best-fit model from Modeltest was TVM + I. Analyses produced six shortest distance trees, 14 maximum likelihood trees, and a large number ( $>900$ ) of most parsimonious trees, with most of the topological differences in the tips of the tree within two clades: "A" (including all of L. morhua ATL, Loma sp. POL and Loma sp. PAC) and "B" (containing all of Loma sp. LIN and Loma sp. BLK). Several branches were consistently obtained by all three types of analysis (shown in bold, Fig. 4.3), including those forming monophyletic clades for the individual species: L. salmonae, L. embiotocia and Loma sp. BRO, all three of which had high ( $>80 \%$ ) bootstrap support from all three methods. There were three other consistently observed clades: clade "H" (including all but one sequence from $L$. branchialis HAD) and clade "NF" (consisting of most $L$. morhua ATL isolates from Newfoundland), and clade "G", consisting of all Loma species from gadid hosts (L. branchialis HAD, L. morhua ATL and Loma spp. PAC, POL and TOM), with the exception of a single isolate of L. branchialis (HA2) which usually, but not always, fell within this group. A clade placing Loma sp. TOM sequences together (clade "T") was observed in most, but not all, analyses. When a gap matrix of eight additional characters (including three additional parsimony-informative sites) was added to the analyses, topologies and bootstrap values were similar to those described above.

## Phylogenetic relationships among Loma species for rDNA Region 3 ( $3^{\prime}$ ' end of SSU, ITS and 5' end of LSU)

Figure 4.4 shows results from rDNA Region 3 ( $3^{\prime}$ end of SSU, ITS, and 5' end of LSU) analyzed and presented as in the previous sections. The best-fit model from Modeltest was TVM $+\mathrm{I}+\mathrm{G}$. Analyses produced one maximum likelihood tree, and a large number ( $>900$ ) of
most parsimonious trees and shortest distance trees, due to low resolution of tips of trees. This topology produced clades $\mathrm{A}, \mathrm{B}, \mathrm{H}, \mathrm{T}, \mathrm{NF}$, and G , which were similar to those observed for the previous data (3' SSU, compare to Fig. 4.3), and additional clade "L" including L. salmonae and Loma spp. LIN and BLK. As in the previous analyses, clades A and B were polyphyletic for Loma species: clade A being polyphyletic for L. morhua ATL, Loma spp. POL and PAC, and clade B for Loma spp. LIN and BLK. Consistent and well-supported branches placed several Loma species with conspecifics and close sister-species: L. salmonae and Loma sp. BRO, and with less bootstrap support and fewer best trees also L. branchialis HAD (clade H), and Loma sp. TOM (clade T). Loma species from gadid hosts: L. branchialis HAD, L. morhua ATL and Loma spp. PAC, POL and TOM grouped consistently together (clade G).

Because many parsimony-informative inter- and intraspecific gaps occurred in this rDNA region (see Table 4.4), gap data were analyzed both alone and in combination with substitutional nucleotide data. The gap matrix for this portion of the rDNA consisted of 27 characters and 21 additional parsimony-informative sites when combined with nucleotide substitutional data. Figure 4.5 shows a maximum parsimony tree resulting from analysis of this gap matrix data alone, without the addition of any nucleotide substitutional or polymorphic substitutional data. Although the resolution of the tree was low, the overall topology, and some clades (branches shown in gray Fig. 4.5) were similar to those observed from nucleotide substitutional data, where the default treatment of gaps is " N " = missing data (compare Fig. 4.5 with Fig. 4.4). Loma sp . BRO grouped together in this tree, as did members of clades B and G.

Figure 4.6 shows results from addition of this gap matrix to the substitutional nucleotide data (best-fit model GTR $+\mathrm{I}+\mathrm{G}$ ). Phylograms (trees with branch lengths drawn proportional to distance) for these data are shown in Appendix 18. The major clades obtained in this analysis, $G$ (with subclades $A, N F, H$ and $T$ ) and $L$ (with subclade $B$ ), were identical to those found previously without addition of a gap matrix. Bootstrap support was similar or slightly higher for several nodes. Members of clade G, comprising Loma species from gadid hosts, formed groups suggestive of paralogous relationships (Fig. 4.6 and 4.7) among divergent rDNA copies within isolates. Presumably, rDNA copies diverged at points indicated with stars on Figure 4.7. Paralogous relationships were observed in several isolates of each of Loma spp. TOM, PAC, and POL, L. branchialis HAD, and L. morhua ATL.

Phylogenetic analysis was also performed to include polymorphic characters, which were shown previously to be numerous across the rDNA, many occurring in the 3' end of the SSU, the ITS and 5' LSU (refer to Table 4.5). Many polymorphic characters were "additive"
sites as defined in Diekmann et al. (2001) - sites at which all three character-states (e.g. A, G, and $A+G$ ) occurred in a set of individuals or isolates population. To bypass the default feature in PAUP* that treats such characters as " N ", unknown or missing data, maximum parsimony was performed by encoding polymorphic data using the "AND" rather than "OR" definition of multi-state characters (i.e. $R=A$ and $G$, rather than $R=A$ or $G$ ), under the hypothesis that these polymorphic states are real in these species rather than ambiguity due to sequencing or Taq mistakes. Resulting topologies and bootstrap values were similar to those found without these polymorphic characters (Appendix 19). Only maximum parsimony analysis in PAUP* could be used with this feature.

## Phylogenetic relationships among Loma species for partial EF-1 $\alpha$

Figures 4.8 and 4.9 show phylogenetic relationships from 966 alignment positions of the partial EF-1 $\alpha$ gene sequence (best-fit model from Modeltest GTR+G). Figure 4.8 shows that the distance tree was slightly different from maximum parsimony and maximum likelihood trees in topology, as it placed one member of clade G (from rDNA analyses) outside the clade. Otherwise, EF-1 $\alpha$ results were similar to rDNA results, producing clades as follows: clade L ( $L$. salmonae with L. embiotocia and Loma sp. LIN) with high bootstrap support, clade G (Loma species from gadid hosts) with high support, and subclade A with little support. As in rDNA trees, Loma sp . TOM was basal to clade A. The EF-1 $\alpha$ trees from all three methods placed Loma sp. BRO outside the entire group, after Loma sp. AUS, in a basal position after $G$. plecoglossi, with high bootstrap support.

## Monophyly tests (AU-tests) for Loma species for rDNA Region 1 (5' half of

 SSU)Hypothetical trees were designed to constrain Loma species with conspecifics or in clades with hosts of the same family, and for each constraint, a heuristic search found the shortest distance tree shown in Figure 4.10. For this region of the rDNA, four such constraints were created without a gap matrix, and three constraint trees with a gap matrix. Log likelihood scores for individual sites for each tree were calculated using the parameters estimated in Modeltest for this data as input for analyses in CONSEL. The results of constrained and unconstrained tree comparison by the AU and other similar statistics are shown in Table 4.9. At the significance level $\alpha$ of 0.01 , these results suggested that both with or without addition of a gap matrix, a monophyly of Loma sp. POL could not be rejected, Loma sp. BRO as a sistergroup to $L$. salmonae could not be rejected, Loma sp. from gadids grouping together could not
be rejected, whereas the placement of Loma sp BRO internal to the $L$. salmonae clade could be rejected (e.g. trees \# 4 and 9, Fig. 4.10).

## Monophyly tests (AU-tests) for Loma species for rDNA Region 2 (3' half of SSU)

Ten constraints placing Loma species with conspecifics or in clades with hosts of the same family were generated for this rDNA region (Fig. 4.11) for each analysis, either with or without addition of a gap matrix. Log likelihood scores for individual sites for each tree were calculated using the parameters estimated in Modeltest for this data as input for CONSEL AU and other statistical test comparisons among trees. The results are shown in Table 4.9. At the significance level of $\alpha=0.01$, these results suggested that both with and without addition of a gap matrix, monophylies of Loma spp. POL or PAC or L. morhua ATL (Fig. 4.11, trees \#2-4 and 13-15) could not be rejected. However, the unconstrained tree in which these three species form a polyphyly (Fig. 4.11, trees \#1 and 12), could not be rejected. Placement of L. morhua ATL and L. branchialis HAD together as a single monophyletic clade could be rejected. Loma spp. LIN and BLK could not be rejected as separate clades (Fig. 4.11, trees \#6 and 17), although the polyphyly of these species (unconstrained trees \#1 and 12) was not rejected either. Loma sp. BRO was rejected as a sister-group or an internal group within the L. salmonae clade. Because one sequence of $L$. branchialis HAD sometimes grouped outside the " H " clade, two constraint trees were created to test placement of this sequence with L. salmonae (Fig. 4.11, trees \#9 and 20) and with Loma sp. TOM (Fig. 4.11, trees \#10 and 21), and the result for both data sets (with or without gaps) was that these trees could not be rejected. A polyphyly of Loma sp. TOM with L. branchialis HAD (Fig. 4.11, trees \#11 and 22) could not be rejected.

## Monophyly tests (AU-tests) for Loma species for rDNA Region 3 (3' end of SSU, ITS, and 5' end of LSU)

Placement of Loma species with conspecifics or in clades with hosts of the same family, or in other arrangements, was tested as before for each analysis, either with or without addition of a gap matrix. Ten such constraint trees were generated for this rDNA region (Fig. 4.12). The results are summarized in Table 4.10 for the significance level $\alpha$, of 0.01 . Several results were the same, regardless of whether a gap matrix was added. Three such examples follow. Loma sp. LIN and Loma sp. BLK were polyphyletic in an unconstrained tree (Fig. 4.12, trees \#1 and 12) but constraining each into monophyly could not be rejected (Fig. 4.12, trees \#6 and 17). Separation of the L. morhua ATL "NF" clade to make it polyphyletic (Fig. 4.12, trees \#8 and
19) was rejected. Separation of Loma sp. TOM clade to make it paraphyletic with L. branchialis and others could not be rejected.

Other results depended on whether a gap matrix was added. Without a gap matrix, separate monophylies for Loma sp. POL, Loma sp. PAC and L. morhua ATL could not be rejected (Fig. 4.12, trees \#2-4), whereas with addition of a gap matrix, these monophyletic groups were rejected (Fig. 4.12, trees \#13-15). Without a gap matrix, Loma sp. BRO could not be rejected as a sister to L. salmonae (Fig. 4.12, tree \#5) but could be rejected as an internal group in this clade (tree \#9), whereas with a gap matrix neither the separation of Loma sp. BRO from the $L$. salmonae clade (i.e. separation of the sister-clades from unconstrained tree \#12 apart, as shown in tree \#16, Fig. 4.12) nor the placement of Loma sp. BRO within the $L$. salmonae clade (tree \#20) could not be rejected.

Intraspecific indel differences affected tree topologies (Figs. 4.4-4.6) and were phylogenetically informative (Figs. 4.5). This suggests paralogous relationships exist in the "G" clade (Fig. 4.7). To examine whether paralogs (diverged gene copies in a single haploid genome) can be grouped into orthologs (gene copies derived by descent from a common ancestor) by the presence or absence of two common indels, these indels were used to constrain groups of isolates. Presence or absence of the insert (or indel) at positions 1381 (in the ITS) or 1622 (in the LSU) (see Table 4.6) were used to group taxa, as shown in trees \#10, 11, 21 and 22 in Fig. 4.12. Without a gap matrix, grouping taxa with and without inserts (indels) 1 and 2 into monophylies was rejected (Fig. 4.12, trees \#10 and 11). With addition of a gap matrix, monophylies of taxa with and without inserts 1 and 2 (Fig. 4.12, trees \#21 and 22) could not be rejected.

## Monophyly tests (AU-tests) for Loma species for partial EF-1 $\alpha$

Figure 4.13 shows six constraints, placing Loma species with conspecifics or in clades with hosts of the same family, generated for the partial EF-1 $\alpha$ gene. The statistical comparisons among trees are summarized in Table 4.11. That Loma sp. POL, Loma sp. TOM, and L. morhua ATL formed their own monophylies was rejected at the significance level $\alpha$ of 0.01 , but the monophyly of Loma sp. PAC could not be rejected. Placement of Loma spp. from gadid hosts (equivalent to clade $G$ ) or the grouping equivalent to clade $A$ also could not be rejected. However, the rejection of the unconstrained tree suggested the shortest distance tree was not the best tree when trees were evaluated using maximum likelihood (lscores) under the best-fit model from Modeltest.

## Recombination tests

Results from single-xenoma DNA extractions (represented by isolates "Gi" and "Ai" of Loma sp. PAC and L. morhua ATL, respectively, in Table 4.1) showed that at least two different rDNA sequences could be found in a single xenoma. These results suggest rDNA paralogs (diverged copies in a single haploid genome) exist within the genome, provided a single, uninucleate spore founds each xenoma. If paralogs recombine within a genome and within an intermixing (sexual) species, this recombination may be of interest in determination of speciesboundaries or species-hybridization. The results of standard tests for recombination (programs SplitsTree of Huson, 1998; and LARD Version 2.2, Holmes et al., 1999) (results not shown) were unable to show support for recombination in any of the sequences from single-xenoma extractions or clade $G$, which appeared to have numerous paralogs.

## Phylogenetic relationships among genera for SSU rDNA

To examine relationships among genera, a single sequence from each Loma species presented in previous analyses was used in phylogenetic analyses with all available species of Loma, Glugea and other species from fishes, crustaceans or myxosporean hosts from Genbank (see Table 4.13). Figure 4.14 shows results of maximum parsimony, distance and maximum likelihood analyses for these species from 810 bp of the small subunit ribosomal RNA (SSU rDNA) gene with alignment gaps, or indels removed. The best-fit model chosen by Modeltest was GTR $+\mathrm{I}+\mathrm{G}$. Heuristic searches produced four maximum likelihood trees, two shortest distance trees, and many ( $>700$ ) most parsimonious trees that differed mostly in the poorly resolved closely related tip taxa. Clades "Loma" and "Glugea" were named for the presence of the type species (L. morhua or G. anomala) in these clusters. The overall topologies and branch lengths were similar from the three methods, except that the clade including $L$. acerinae changed position in the tree, depending on the method (Fig. 4.14).

Figure 4.14 shows that all three analyses place the unnamed Loma sp. Nil from Enchelyopus cimbrius well within the Loma clade, sometimes associated with members of clade G (from gadid hosts). All analyses place the 'Unidentified microsporidium MYX1' from an unidentified myxosporean parasitizing a tiger puffer, Takifugu ruripes, as the first branch just outside Loma sp. AUS (from a surf bream in Australia), before the Ichthyosporidium and Pseudoloma clade. All analyses also placed L. acerinae (both sequences) well outside the Loma clade. However, the place of $L$. acerinae varied: it was sister to the Loma clade in parsimony, sister to the Glugea clade in distance, and basal to these two clades in maximum likelihood.

Branches in common among all three analysis methods (in bold) and bootstrap values $>50 \%$, plotted on a maximum parsimony tree in Figure 4.15, show several strongly supported relationships. The Glugea and Loma clades were strongly supported, with $\sim 100 \%$ and $\sim 80 \%$ bootstrap values, respectively. These two clades, plus some nearby relatives, were strongly supported ( $\sim 100 \%$ ) sisters, branching after "Group 4" (Spragueidae-Tetramicridae) defined in Lom \& Nilsen (2003). Groups 1 (Pleistophoridae) and 3 (no family name) from Lom \& Nilsen (2003) were strongly supported in these analyses. However, Group 4 (Glugeidae) was only sometimes found in these analyses. Sometimes the Glugeidae group included L. acerinae in these results. Figure 4.15 shows that Loma sp. AUS and Loma sp. BRO were basal with respect to other species in the Loma clade. All analyses placed these taxa basally. The remaining species of the Loma clade were arranged as shown in Figure 4.15, or with slight differences, depending on the method used. Resolution of some taxa in the Loma clade was somewhat lower than that observed previously (Figures 4.2 to 4.9 and in Lom \& Nilsen, 2003), presumably due to due to gap-stripping in these data, which removed the most variable SSU regions.

## Phylogenetic relationships among genera for partial RPB1

Figure 4.16 shows maximum parsimony, distance and maximum likelihood trees (best-fit model GTR $+\mathrm{I}+\mathrm{G}$ ) for 783 alignment positions from the partial RPB1 DNA gene for available species from Genbank (see Table 4.13), showing overall branch length differences, and variation in tree topology. Heuristic searches produced just one shortest, most parsimonious or most likely tree in each case. The placement of $L$. acerinae in relation to $L$. salmonae and $G$. anomala depended on the method used. As with SSU rDNA trees (Figs. 4.14 and 4.15), $L$. acerinae was basal only in maximum likelihood trees; however, bootstrap support in this part of the tree was low ( $\sim 50 \%$ ). Other relationships in these trees were poorly supported (Fig. 4.17), except that corresponding to Group 1 (Pleistophoridae) of Lom \& Nilsen (2003).

## Monophyly tests (AU-tests) for genera for SSU rDNA

Four hypothetical trees were created to constrain members of genus Loma together, or Loma species in clades with hosts of the same family (Fig. 4.18, Table 4.14). Statistical comparisons of trees were calculated in CONSEL using likelihood scores calculated using the parameters estimated from this data by Modeltest. At the significance level $\alpha$ of $0.01, L$. acerinae was rejected as a sister to the Loma or Glugea clades (Fig. 4.18, trees \#2 and 5), or a sister to L. branchialis (Fig. 4.18, tree \#3). A tree constraining Loma species from gadids (Loma sp. Nil, Loma spp. POL, PAC, TOM, L. branchialis and L. morhua) as a monophyly could not
be rejected. Conversely, the distance-based unconstrained tree in which Loma species from gadids were polyphyletic was rejected, presumably because the shortest distance tree was not the best tree when trees were evaluated using maximum likelihood Iscores created under the best-fit model in Modeltest.

## Monophyly tests (AU-tests) for genera for partial RPB1

Two constraint trees were created to test the place of $L$. salmonae, L. acerinae, and $G$. anomala with the RPB1 gene data (Fig. 4.19 and Table 4.14). Statistical comparisons, calculated in CONSEL showed at the significance level $\alpha$ of 0.01 , that none of the three possible topologies among these three species could be rejected, given the data available.

## DISCUSSION

## Overview of major results

This study presented analyses of DNA sequences from three rDNA regions, partial EF$1 \alpha$ and RPB1 genes and examined phylogenetic relationships of four nominal and seven undescribed species of Loma, five of which were described informally in Chapter 3 (see Appendix 12 for table of name equivalents). Several unique sequences were compared from each of several isolates across each species' range for most Loma species shown in Table 4.1. Phylogenetic results (Figs. 4.2-4.9) and statistical comparisons of trees (summarized in Table 4.12), as well as other data such as spore morphology (Table 4.3) and sequence similarity suggested several well-separated groups that are consistent with the working definition of species given in the introduction (and see Appendix 5). Sister-species separated from one another into well supported groups in all or most analyses were: L. branchialis from haddock and L. morhua from Atlantic cod, Loma sp. from brook trout and L. salmonae, L. embiotocia and L. salmonae, Loma sp. from Pacific tomcod and all others, and Loma sp. from an Australian surf bream and all others. Species-groups that were not well separated in these analyses were: $L$. morhua from Atlantic cod and two sister species in Pacific gadids, Loma sp. in Pacific cod and Loma sp. in walleye pollock, Loma sp. in lingcod and Loma sp. in sablefish. Despite the weak support for these latter species (L. morhua, Loma species from Pacific cod, walleye pollock, lingcod and sablefish), these species did not constitute "rejectable groups" when individuals of a species were forced into monophyletic clades (Figs. 4.10-4.13) and trees were compared by the AU test of Shimodaira \& Hasegawa (2001), except when emphasis was placed on unusual characters (indel differences in the ITS and LSU) that may represent relationships between paralogs (diverged gene copies in a single haploid genome) rather than orthologs (gene copies derived by descent from a common ancestor) (see Table 4.12; and Fig. 4.7). All available Loma species grouped together in a well-supported monophyletic clade in all analyses, with the exception of $L$. acerinae, which was rejected as a member of both genus Loma and genus Glugea in rDNA analyses, but could not be rejected from either genus in RPB1 analyses. Phylogenetic relationships among Loma species, which were largely consistent across gene regions and genetic loci, were consistent with host-parasite co-evolution for species from gadids (Fig. 4.20) and possibly Scorpaeniformes.

## Identity of spores with $L$. branchialis and L. morhua (Morrison \& Sprague, 1981)

Spores from Atlantic cod were significantly larger than those from haddock, while Morrison \& Sprague (1981a; b) found the opposite size relationship from the same hosts at the same location. That study did not provide a variance around the mean, so it is not clear whether the size difference from Morrison \& Sprague (1981a; b) was statistically valid. Also, spore sizes in the present study and those of Morrison \& Sprague (1981a; b) may differ only due to differences in the type of resin, EPON in Morison \& Sprague (1981a) vs. Spurr's resin in the present study. When spore shrinkage was taken into account, spores in this study were still smaller than those of Nemeczek (1911) for L. branchialis; however, Nemeczek (1911) did not indicate whether his measurements were a mean or just a single "typical" size. Thus, again it was difficult to evaluate similarity of spore size with any statistical confidence. As a result, it must be concluded that spores from haddock and Atlantic cod in the present study may be conspecific with L. morhua (and L. branchialis) as defined by Morrison \& Sprague (1981a; b), and that, if they are separate species, host is still a better character to identify material with that from these studies.

However, the present study did show that spore sizes fell into two statistically distinct groups in the two potentially sympatric host species. The observation of spore size differences in the two hosts is consistent with several other studies in Atlantic Canada. For example, when shrinkage was taken into account, spores from both Atlantic cod and haddock in the present study were larger than those from haddock at a nearby location measured by Lom \& Laird (1976) and similar in size to those from Atlantic cod in a nearby location (Fantham et al., 1941). Thus, the present study confirms previous studies that present the opposite result from that of Morrison \& Sprague (1981a; b), and suggest that Loma species from Atlantic cod in eastern Canada (here referred to as L. morhua) have larger spores than do sympatric Loma species from haddock (here referred to as L. branchialis).

In summary, spore-shrinkage calculations in this study (Table 4.3; and see "Spore size comparisons" in Results) suggested that fixation accounts for some, but not all, of the difference in spore sizes among Loma species from gadid hosts, as was predicted by Kabata (1959) and Morrison \& Sprague (1981a). Statistically significant differences in spore sizes from the two hosts, combined with previous studies (see Morrison \& Sprague, 1981b; and Table 4.3) and the present study, showing qualitative differences in xenoma size in these two potentially sympatric
hosts supports the idea that these are two distinct Loma species. DNA data (Figs. 4.3-4.7, 4.11, 4.12; see discussion below) strongly corroborate this morphological (and host) data, suggesting these are separate species. Therefore, according to the law of priority of the International Code for Zoological Nomenclature, I conservatively employed the names originally given for material from this locality and these hosts, using L. morhua Morrison \& Sprague, 1981, for the species in Atlantic cod, and L. branchialis for material in haddock. If L. morhua is hereafter recognized again as a valid species, it would be the type species as was originally proposed by Morrison \& Sprague (1981a), as opposed to L. branchialis, which was considered as the type only after Canning \& Lom (1986) chose not to recognize L. morhua.

## Phylogenetic support for separate species $L$. branchialis and $L$. morhua

## (Morrison \& Sprague, 1981)

Loma species from haddock (L. branchialis HAD) and Atlantic cod (L. morhua ATL) were consistently separated in rDNA analyses, and placement of sequences from species together as a monophyletic clade was always rejected (AU tests), suggesting they represent distinct species or at least distinct clusters of ancient paralogs (see further explanation of this in 'Reticulate speciation' section below). Loma branchialis formed a consistent single species monophyly in SSU/ITS/LSU analyses, and formed a strong clade with only one sequence (isolate "HA2") outside this clade in analyses of the 3' portion of the SSU. Placement of this stray sequence varied, depending on the analysis method. However, neither its placement with conspecifics nor with other species (Loma sp. TOM or L. salmonae) could be rejected, so it is not clear whether this represents a variant strain or species in haddock, or merely an rDNA paralog (discussed above) or polymerase error.

Loma morhua did not form a single-species monophyly, but was, instead mixed either paraphyletically or polyphyletically with Loma species in clade A (Figs. 4.3-4.6, 4.8, 4.9). Species in clade A were all from closely related gadid hosts, Atlantic cod, G. morhua, Pacific cod, G. macrocephalus, and walleye pollock, T. chalcogramma, the relationships of which, according to Carr et al. (1999) are difficult to resolve. However, a species monophyly for $L$. morhua could not be rejected except in the analysis of the SSU/ITS/LSU with the addition of a gap matrix. Evidence discussed earlier shows that these indel differences were due to rDNA paralogs, so the correct identification of orthologs may be necessary before $L$. morhua can be resolved into one or more monophyletic clades using rDNA sequence. Possible explanations for the lack of resolution of $L$. morhua into a single group are that divergent strains may exist or two
species of Loma may exist in Atlantic cod. Others have suggested gadids may be infected with multiple species of Loma (Kabata, 1959, Morrison \& Sprague, 1981a; b). In this study, Loma species from Atlantic cod from Newfoundland consistently formed a separate sub-clade (clade $N F$ ), which always included one sequence from an isolate outside Newfoundland (sequence \#1 from a Nova Scotian isolate named "Aa") (Figs. 4.3-4.6). Separation of the NF clade making it polyphyletic (Fig. 4.12, trees \#8, 19) was rejected, suggesting this clade could represent a distinct species, or possibly merely a divergent strain. Laboratory studies (Dr. D. Barker, personal communication) of these Newfoundland L. morhua isolates suggest they also possess developmental and virulence differences, suggesting there may be two strains in Newfoundland. Here there was evidence that Newfoundland and Nova Scotian strains may differ genetically, but further data are needed to examine whether these genetic differences correspond to morphological or virulence differences at these locations or across the geographic range of Loma species from Atlantic cod. Polyphyly of $L$. morhua was also suggested by results from the single-copy gene EF-1 $\alpha$, for which a single isolate (labeled "Aa" in Table 4.1) fell outside the other group of $L$. morhua sequences. Analyses from this gene suggested a single species monophyly could be rejected ( $p=0.0004$ in AU test). The isolate "Aa" also appeared to have divergent rDNA sequences, such that one rDNA sequence from this isolate grouped with the well-supported NF clade, which otherwise included only $L$. morhua isolates from Newfoundland. These two genetic loci suggest this "Aa" isolate of $L$. morhua may have had not only diverse rDNA paralogs, but also variant strains, or may be a distinct species.

Analyses of partial SSU rDNA for a variety of microsporidians from fishes placed the Loma species from haddock (L. branchialis HAD) in a position basal to the Loma species from Atlantic cod (L. morhua ATL) and its sisters in clade A. Analyses of rDNA from only Loma species also placed L. branchialis basally, suggesting L. branchialis is an older species that diverged earlier than others in clade A. L. morhua from Newfoundland (clade NF) often occurred in a basal position within clade A , and similarly, many other isolates of $L$. morhua were also placed in a deep position in clade A, suggesting that these may be ancestral to undescribed Loma species from Pacific cod and walleye pollock, although this pattern is weak and is perhaps obscured by rDNA paralogs.

## Phylogenetic support for seven undescribed Loma spp.

Analyses of partial rDNA from a range of microsporidia or from just Loma species always resulted in placement of the Loma species from the Australian surf bream
(Acanthopagrus australis) (Loma sp. AUS) as a distant, basal member of the Loma clade (Figs. $4.3,4.4,4.14,4.15)$. This result was not surprising, given the distant geographic location and the distant taxonomic status of the host (order Perciformes, family Sparidae), compared to other Loma species sequenced to date. Microsporidia may co-evolve with their hosts often (Baker et al., 1997), suggesting that two other species, L. diplodae and L. boopsi, from Sparidae hosts may also be distant and basal members of the main Loma clade. Similarly, under a co-evolutionary hypothesis, perhaps some or all of the other nine described species of Loma from hosts in the order Perciformes (listed in Table 1.1 of Chapter 1) would group at the base of the Loma clade.

Partial SSU rDNA analyses that included many species of Loma and related genera (Figs. $4.14,4.15$ ) suggested the Loma sp. from brook trout in Prince Edward Island, Canada (Loma sp. BRO) branched early, just after Loma sp. AUS. In most rDNA analyses of just Loma species, Loma sp. BRO fell outside the $L$. salmonae cluster, and not usually as an immediate sister-species (see Figs. 4.3, 4.4). This result was consistent with partial EF-1 $\alpha$ sequence analyses (Figs. 4.8, 4.9), suggesting these two Loma species are not conspecific. However, this result was unexpected, given that both Loma sp. BRO and L. salmonae have hosts in the Salmonidae. Sometimes Loma sp. BRO could not be rejected as a sister-species to the $L$. salmonae cluster, suggesting they may be close relatives. In only one case, for the 5 ' end of the SSU, ITS and partial LSU with addition of a gap matrix, Loma sp. BRO fell within the $L$. salmonae cluster; however, this relationship was never supported with bootstrap confidence $>50$ $\%$, and the separation of these species into 2 monophyletic clades (Fig. 4.12, tree \#16) in these analyses could not be rejected. This poorly supported relationship from the normally variable ITS/LSU region arose presumably because L. salmonae and Loma sp. BRO have unusually few substitutional differences in this region, one of which was a parsimony-informative indel. In contrast, the SSU rDNA and partial EF-1 $\alpha$ genes reveal a higher level of variation ( $\sim 1.4 \%$ and $2.3 \%$, respectively) between these species. These data suggest the indel shared by these species might have arisen in a common ancestor, prior to the divergence of these species and the accumulation of other mutations in the EF-1 $\alpha$ and SSU rDNA.

The unusually low interspecific difference among the ITS/LSU of species in salmonid hosts make it difficult to design species-specific probes. For example, the supposedly $L$. salmonae specific "LS-1" primer (Docker et al., 1997a; Brown \& Kent, 2002) will match Loma sp. BRO rDNA perfectly, while the "LS-2" reverse primer may also bind at lower annealing temperatures (Brown \& Kent, 2002). These results do bring to question the identity of Loma sp.

BRO with a morphologically distinct species, L. fontinalis Morrison \& Sprague, 1983 in the same host (brook trout, Salvelinus fontinalis) from a locality (Cobequid hatchery, Nova Scotia, Canada) near the source locality of the brook trout with the infection in this study. These species may be identical, but neither detailed morphological analysis, nor DNA sequence are available for L. fontinalis, although some inconsistent amplification was obtained from a paraffin-embedded sample of L. fontinalis using the LS-1 \& LS-2 primers of Docker et al. (1997a) (Brown \& Kent, 2002). It is important to note that this species, designated Loma sp. BRO here is identical with material named provisionally " $L$. salmonae SV" by others (Sánchez et al. 2001b; 2001a; Speare \& Daley, 2003). Speare et al. (1998a) and later others attempted to transmit L. salmonae from Oncorhynchus spp. (from British Columbia) to brook trout, with no success. They found infected brook trout following a later trial, and found spores from these brook trout were different in biology from all L. salmonae observed previously, in that they preferentially infected Salvelinus species over Oncorhynchus species. In this study, DNA sequences from infected gills of five of those brook trout suggest this infection was from a different species that was not present in British Columbian L. salmonae from the same laboratory that Speare et al. (1998a) and Sánchez et al. (2001a; b) and others obtained spores for their transmission experiments. This study also examined DNA from L. salmonae isolates in California, Colorado, Idaho, and Chile, with the result that all isolates of L. salmonae appear significantly diverged from Loma sp. BRO (= "L. salmonae SV"). I suggest that this result requires an explanation that brook trout used in the trial that yielded these infections carried a light, undetected infection with a separate Loma species (Loma sp. BRO). This infection could be L. fontinalis. Other studies confirm that Loma species can reside in tissues, undetectable by the naked eye or light microscopy (Shaw et al. 2000c).

Most analyses placed the five other undescribed Loma species into two major groups, one a monophyletic group consisting of all Loma spp. from hosts of family Gadidae (clade G, discussed above), and the other a paraphyletic group consisting of all species from non-gadid hosts with the gadid-parasitic monophyly nested within it (Figs. 4.3-4.6, 4.9, 4,15; ML \& MP but not ME of Fig. 4.8; and note Fig. 4.10 tree \#5 not rejected; Fig. 4.13, tree \#6 not rejected; Fig. 4.18 tree \#4 not rejected). This suggests seven Loma species (or eight, counting Loma sp. Nil) are grouped within closely related hosts. Similarly, within the non-gadid-host Loma species, two undescribed Loma species (Loma sp. LIN and Loma sp. BLK) from hosts of order Scorpaeniformes always fell into a separate clade together, clade B (Figs. 4.3-4.6, 4.15), often with $>50 \%$ bootstrap support (Figs. 4.4-4.6, 4.15). Together, these results suggest there is
generalized host specificity and the host may be an important indicator of species, as is suggested in Baker et al. (1997) and Baker et al. (1998). These data are also consistent with host-parasite co-evolution (co-speciation). If host-parasite co-evolution had not occurred in these Loma species, clusters might have been expected to group by geography. For example, Loma species from eastern North America (L. morhua, L. branchialis, Loma sp. BRO) could have grouped together while species from western North America (L. salmonae, L. embiotocia, Loma spp. PAC, POL, and TOM) could have formed a clade.

Partial rDNA sequences from the Loma sp. from Pacific tomcod (Loma sp. TOM) fell in a basal position in clade $G$ in analyses from a variety of fish-parasitic microsporidia (Figs. 4.14, 4.15), while in analyses of only Loma species, Loma sp. TOM formed either a monophyletic cluster (Figs. 4.2-4.4) or formed paraphyletic branches basal to clade G (Figs. 4.4, 4.6), and AU tests suggested neither monophyly nor paraphyly could be rejected (Tables 4.9, 4.10). Elongation factor- $1 \alpha$ results confirmed rDNA results, consistently placing Loma sp. TOM sequences together as a pair of paraphyletic basal branches in clade G (Fig. 4.9), whereas AU tests rejected their monophyly (Table 4.11). These results suggest Loma sp. TOM may be ancestral members of a larger group of Loma species in gadids, or they may suggest species have undergone hybridization (as discussed above) with sympatric congeners (e.g. L. salmonae). More sequences from a greater number of loci and more individuals of all species are needed to distinguish such hypotheses. Morphological data (Chapter 3) tend to suggest that Loma sp. TOM is a separate and distinct species from others in gadids, and these two results are consistent with what is known about the phylogeny of the hosts (discussed below).

Loma species from Pacific cod (Loma sp. PAC) and walleye pollock (Loma sp. POL) formed unresolved or polyphyletic clades with L. morhua in the sub-clade A, within clade G for all but one DNA region (Figs. 4.2-4.6, 4.9). Only in the analysis of the $5^{\prime}$ SSU region did the two sequences from Loma sp. POL fall together (with low bootstrap support, Fig. 4.2). Despite the low resolution and prevalence of polyphylies in these clades, AU tests often failed to reject the monophyly of each of these two species. There were two exceptions where monophylies were rejected for these two species. The first, when gaps were added to the SSU/ITS/LSU region can be explained by paralogs (discussed above), while the second case of species monophyly rejection, only for Loma sp. POL from partial EF-1 $\alpha$ analysis would require an alternate explanation. While overall divergence of the partial EF-1 $\alpha$ gene was low, the \% intraspecific divergence in this gene was higher in species from family Gadidae, suggesting that
the poly- or paraphylies in this clade could be due to multiple, divergent copies of this supposedly single-copy nuclear gene (L. morhua isolate A10, for example had two sequences), or these fish could be infected with multiple strains (or even species) of Loma. More sequences from a greater number of individuals, individual xenomas, and from a greater geographic area would be needed to confirm or refute such possibilities.

Clade B, containing Loma species from lingcod (Loma sp. LIN) and sablefish (Loma sp. BLK), was similarly unresolved for these two species in all analyses, perhaps due to paralogs or reticulation (as discussed above); however, the separation of these two species into species monophylies could not be rejected by AU tests, therefore suggesting there is insufficient evidence here to place these species either together into one species, or separately, into two species.

In order to resolve the question of species validity of Loma species from Pacific cod, walleye pollock, lingcod, and sablefish, one must consider the general subject of species definitions from the lineage-based species concept literature (see Quicke, 1993; and essays in Wheeler \& Meier, 2000). According to some operational species concepts (Templeton, 1989, Templeton, 1994; Sites \& Crandall, 1997), "good" species would be expected to form separate, monophyletic clades by molecular data, whereas in cases where data suggest polyphylies or polytomies among candidate species, there is not sufficient evidence that they are separate (Mallet, 1995). Hence, the question of their species status must be left open where species monophylies cannot be rejected, as was the case for Loma sp. LIN and Loma sp. BLK, and Loma sp. PAC and Loma sp. POL, until further data, such as morphological, geographic or host specificity data are obtained (see Chapter 3).

Morphological differences do, in fact, exist among the five undescribed Pacific Loma species (Chapter 3). These morphological differences include that spores of Loma sp. BLK were statistically significantly larger than spores of Loma sp. LIN and developmental stages differed between Loma sp . PAC and Loma sp . POL. These differences arguably support that there has been some degree of reproductive isolation in these Loma infections from different hosts, such that they may be valid, but recent species. Under this hypothesis, the low DNA sequence divergence could be caused by inadequate sequence information (length) compared to the mutations accumulated (time) since speciation. Another explanation for the lack of resolution is that paralogs older than the species may contribute to polyphyly in the phylogenetic analyses (see next section). The idea of a recent speciation is also supported by the observation that these morphological differences were greater than those observed between other close Loma
species pairs, like L. salmonae and L. embiotocia, which were not able to transmit to reciprocal hosts (Shaw \& Kent, 1999; Shaw et al., 2000b, c), and supports what is known about host phylogeography (discussed below). More careful transmission studies, together with quantitative molecular and morphological studies are needed for examining such close species pairs.

## Reticulate speciation in the " G " clade?

Results from rDNA analysis suggested taxa in clade G may have undergone reticulate speciation (i.e. non-treelike branching in which branches may reticulate or anastomose). Another possibility is that the rDNA genes of species in clade $G$ have undergone reticulation (i.e. crossing-over, gene conversion, or some other recombination), instead of diverging in a normal dichotomous pattern. Three kinds of evidence suggested this reticulation in clade $G$ (including Loma species from gadid hosts, with sub-clade A that includes L. morhua and two relatives, Loma spp. PAC \& POL). First, there was unexpectedly low resolution and paralogous relationships in clade $G$ when gaps were included in the analyses (Fig. 4.6). Second, there were more substitutional polymorphic sites in this group compared to other species of Loma (Table 4.7). Third, there were more intraspecific gaps (indel) polymorphisms in species of clade G compared to other species (Table 4.6).

Reticulate speciation occurs when populations of true species undergo a single or several bouts of hybridization, such that their true tree has an anastomosing rather than a dichotomizing appearance (Quicke, 1993, and references therein). Genes typically undergo reticulation when they occur in multiple copies (e.g. rDNA and other multi-gene families) and diverge, then later recombine in the genome, although this should only occur within an interbreeding group. Ribosomal genes would not show a reticulating pattern across reproductively isolated groups, whereas organisms that are largely isolated but have occasional bouts of hybridization may show reticulation, thus both genes and organisms may reticulate together in some cases (Sang et al., 1995; Hughes \& Petersen, 2001).

Evidence for these alternate possibilities will be discussed separately.

## Possibility 1: Reticulate evolution of genes.

Loma species were not resolved in subclade A, perhaps due to recent speciation and failure of rDNA paralogs to diverge into species-based clusters, rather than copy-based clusters (see Fig. 4.7). Direct evidence of rDNA paralogs was seen in isolates labeled "Gi" and "Ai" from Table 4.1, which were each derived from a single xenoma. Each of these xenomas
produced two unique rDNA sequences, differing by a 6 bp indel and at least 6 substitutions. If a xenoma were the product of vegetative divisions from a single infective spore (sporoplasm), these rDNA copies would be derived from a single genome (and therefore be proof of rDNA paralogs). This is because somatic mutations in dividing stages in xenomas would not be expected to accumulate quickly enough to produce these differences. If a xenoma is formed by more than one founder-cell (i.e. two sporoplasms enter a host cell) sexual exchange among nonrelatives may occur and the rDNA copies in single xenomas could be the result of two alleles at a locus. It is not known how often two sporoplasms can infect a single cell. However, different species of Loma would not be expected to have such similar sequences for both alleles (see Fig. 4.7) if the rDNA copies can recombine, so it is much more likely that single xenoma rDNA copies represent non-recombining paralogs that diverged before the divergence of species.

Lack of resolution in clade $G$ (and subclade A) may be solely because of recent speciation and the presence of paralogs; however, sequence divergence within members of this clade is high compared to that in related species outside this clade. This suggests sufficient time has passed that mutations should have accumulated that distinguish species in clade A . Moreover, if all the polymorphic variation in clade $G$ was due to rDNA paralogs, one would expect species to resolve into perfect paralog groups and also to be better resolved by EF-1 $\alpha$ (a single-copy gene). This was not the case. Hence, some of the lack of resolution of species may be due to strain divergence. For example, the high frequency of polymorphic sites in clade G could arise from multiple parasite strains present in these hosts (e.g. the cluster of L. morhua isolates comprising the NF clade may be a diverged strain). Furthermore, upon closer examination, the paralogs characterized by indels in the ITS and LSU appear to occur both deeper (in L. branchialis and Loma sp. TOM) and later in clade G, at the tips of branches in subclade A, as shown in Figure 4.6, and seen by comparing Figure 4.6 to tree results using gaps only (Fig. 4.8). This suggests gene duplication or loss of homogenization among dispersed rDNA occurred repeatedly in this clade.

The addition of gaps to analyses, and analysis of gaps alone suggested the presence of either recombining (reticulating) or non-recombining paralogs; however, analyses without gaps and gap stripping did little to improve the resolution of species in clade $G$ or sub-clade $A$. This suggests that paralogs differing by indels were not the only reason for non-treelike evolution in these clades, or that the species (organisms) may also have reticulated.

This possibility is examined below.

## Possibility 2: Reticulate evolution of organisms.

While organismal reticulation might have caused the low resolution of species in clades G and A, this need not exclude Possibility 1, and so the case of paralogs (above) may stand, regardless of possible peculiarities in the organismal speciation. The potential support for recombination across species rather than recombination only within a genome, caused by species hybridization and causing reticulate evolution (= anastomosing relationships), comes from the observation of both shared indels and shared substitutional polymorphisms across species in Tables 4.6 and 4.7. The relationship between indels and recombination was suggested by results of Cheynier et al. (2001), in a study of simian immunodeficiency virus (SIV), it appears recombination can frequently cause indels, which remain as markers of such events. While recombining rDNA copies may not accumulate indels from recombining at the same rate as retrovirus genes do this, presumably all recombination, which requires a break of the DNA backbone in both strands will sometimes result in indel-errors during the re-joining of strands, and the rate these errors occur may be affected by the built-in repair ability of the genome. Such repair apparatus may be poor, or reduced (like many processes and biochemical pathways in the microsporidian genome) in some or all microsporidians. While shared intraspecific and interspecific indels in Loma species could be due to common ancestry; alternatively, indels that are shared could arise from interspecific hybridization (recombination) events. Diekmann et al. (2001) suggested a high frequency of "additive-type" polymorphic sites, where three characterstates occur at a site in a set of individuals or isolates population, the "A only", "G only", and "A/G double-signal in the rDNA, which suggests species have hybridized in the past. Diekmann et al.'s (2001) data looks similar to that found here (in Table 4.7, 13 sites are additive), suggesting similar processes may be responsible.

Diekmann et al. (2001) argued that numerous additive sites in the ITS of coral species were hard to explain by insufficient lineage sorting at so many sites, given what was known about the morphological, phylogeographic, and overall genetic divergence among those coral species. The data from this study appear similar, particularly as 10 sites (positions 157, 751, $981,1367,1393,1394,1469,1649,1717$ and 1750) show additivity in the G clade, and four sites (1711, 1715, 1717 and 1785) show additivity only in one G clade member, Loma sp. TOM. Under the reticulation hypothesis, Loma sp. TOM would appear to have hybridized with ecological (geographic) sisters (e.g. L. salmonae and L. embiotocia) rather than phylogenetic sisters (other clade G members such as L. morhua ATL and L. branchialis HAD; Figs. 4.4-4.7) according to one of these sites (1717). While support for such organismal hybridization is weak, it is interesting that it has been postulated to occur in other marine species (e.g. Mytilus spp.

Hilbish, 1996; and Madracis spp. Diekmann et al., 2001) perhaps by a process of "surface circulation vicariance" (Vernon, 1995), where species can remain able to hybridize after separations of tens of thousands of years by intermittent re-mixing of dispersal stages in ocean waters. This long-distance ocean mixing might need to be invoked to explain possible specieshybridization in Loma spp., as members of clade G occur in opposite ocean basins of North America.

To examine this question of potential hybridization, some of the sequences in clade G (and sub-clade A) were analyzed by software designed to detect recombinants or recombination sites, including SplitsTree (Huson, 1998) and LARD Version 2.2 (Likelihood Analysis of Recombination in DNA; Holmes et al., 1999). However, the results (not shown) were unable to show support for recombination, perhaps because the individual sequences were short (the full rDNA region was sequenced in shorter products), or perhaps because there was little divergence and so little signal. These programs have limitations; most importantly, they cannot distinguish between recombination between taxa (organisms) and recombination among rDNA copies within a genome.

## Improving resolution in clade "G"

Given a better understanding of the underlying genetics of the species, coalescent theory might estimate gene genealogies more efficiently for Loma species of clade G. Coalescent theory does not assume a truly dichotomously branching tree of either the organisms or the genes (see Rosenberg \& Nordburg, 2002). Coalescent theory might be the only appropriate framework for examining such polymorphism data, because it is the only approach that models reticulating evolution, whether this is due to hybridization of rDNA copies within a genome, a population, among populations or among species. Coalescent theory could allow us to answer questions about population parameters such as the degree of hybridization, population size, immigration, mutation rate, etc. Use of the coalescent depends on knowledge of ploidy, sexual recombination system, number of genetic paralogs, or other parameters that are presently unavailable, and therefore could not be applied to these data at this time.

## Integrity of genus Loma

Despite the omission of some members of the genus, this study considerably broadened the phylogenetic picture for the genus by contributing an additional nine Loma species from four families and five orders of host from 3 distant geographic localities for rDNA analyses. Twelve species (Figs. 4.14, 4.15) clustered with the type species and remain valid members of the genus,
whereas $L$. acerinae appeared to be invalid, and data were unavailable for the remaining nine species of Loma (un-shaded in Table 1.1 of Chapter 1). Thus, this study was the first to demonstrate the integrity of Loma by showing that all but one species cluster together in a wellsupported monophyletic clade centered about its type species, L. morhua Morrison \& Sprague, 1981. This clear monophyly puts to rest questions raised by Lom \& Pekkarinen (1999) about whether several of these species of Loma are valid members of the genus.

Loma acerinae (formerly Glugea acerinae) was an outsider to the Loma and Glugea clades that could not be placed with support together with either Loma or Glugea in both rDNA and partial RPB1 analyses. This result confirms the conclusion made by Lom \& Nilsen (2003), that $L$. acerinae has been misplaced, and that despite its morphological similarity to Loma species, it is a distant relative, needing a new genus name. The agreement of RPB1 and rDNA for $L$. acerinae strengthens the case made by Cheney et al. (2001) that RPB1 may be a useful and reliable locus for microsporidia, which, unlike rDNA, is single-copy and so may be able to resolve close species for which rDNA paralogs are a problem. In this study RPB1 was not easily amplified, particularly for small quantities of DNA, presumably because, unlike multicopy rDNA, this single-copy gene is rare enough that good copies are destroyed during fixation or DNA isolation steps. For example, RPB1 could be sequenced from large amounts of concentrated, purified-spore DNA, but could not be amplified or sequenced from whole, ethanol-fixed or frozen infected gills. A more delicate protocol may be necessary to reliably sequence this gene from infected gills.

Other relationships among fish-parasitic microsporidians in this study resembled those observed previously (Lom \& Nilsen, 2003; Bell et al., 2001; Nilsen, 2000; Nilsen et al., 1998; Matthews et al., 2001), which raises questions about the peculiar sister taxa of the genus Loma and how to find an appropriate family for Loma. Historically, taxonomists have also had difficulty classifying Loma and relatives to suitable families based on morphology. Loma was placed in the family Glugeidae Thélohan, 1892 by Morrison \& Sprague (1981a) based on similarity of its members to Glugea species. However, this similarity has been questioned to such an extent that two other families were later suggested for Loma (see Sprague et al., 1992). Recently Lom \& Nilsen (2003) decided to remove Loma and relatives from any family because they form a paraphyletic assemblage of strange, morphologically mixed genera. Loma and these mixed genera formed a sub-clade within Lom \& Nilsen's (2003) "Group 3" (Figs. 4.14, 4.15), wherein Loma species form a subclade among close relatives with distinctly different features. For example, Ichthyosporidium is uniquely diplokaryotic and has no sporophorous vesicle, and
the close relative Pseudoloma forms unusual xenomas specialized for neural tissue of the hosts, whereas the "Unidentified microsporidian MYX1"is peculiar as hyperparasite in a fish-parasitic myxosporean. The present study did not present a new solution for the problem of a family for Loma and relatives, other than one like that suggested by Nilsen \& Lom (2003) in which a family might need to be erected based on phylogenetic similarity, regardless of morphological differences. For now, it is best to wait until more species are examined and patterns of morphological evolution of the microsporidia are better understood.

This study showed partial EF-1 $\alpha$ is not better than rDNA for resolving close species, since it seems to evolve slowly, but this slow rate may be ideal for resolving genus, family or higher-level taxonomic relationships in the microsporidia. Both EF-1 $\alpha$ and rDNA genes function in the process of translation, and may be under the same unusual evolutionary constraints (Inagaki et al., 2003), therefore this gene may not be as evolutionarily independent from rDNA as is the RPB1 gene. Non-independence of the genes may not be a problem for closer relationships since the mutations in question are at mostly selectively neutral sites, although it remains unclear whether this would pose a problem at the genus- and higher levels. Nevertheless, EF-1 $\alpha$ is a low-copy number gene that was more easily amplified than RPB1 in this study, and therefore could be suitable for resolving intra-generic level relationships for Loma and relatives when sequences become available for more outgroup taxa.

## Co-evolution and phylogeography of Loma spp. in gadids

Phylogenetic analyses of gadid evolution using mtDNA data (Carr et al., 1999) combined with results of this study shows close congruence between host-parasite phylogenies in both topology and relative branch lengths (see Fig. 4.20). This suggests Loma species have undergone co-migrations with their gadid hosts for at least 12 million years. Carr et al. (1999) analyzed mitochondrial cytochrome $b$ and cytochrome oxidase I genes from 14 gadid species from eight genera, including all five species of gadids from which Loma species were analyzed here. The topology of Loma species from gadids matched that of their five hosts, except that Carr et al.'s (1999) study could not resolve Pacific cod, Atlantic cod, and walleye pollock in some analyses. The parasites of these three fishes were also difficult to resolve.

Carr et al. (1999) were able to date these mtDNA divergences (using Bermingham et al., 1997, and geological data), and hypothesized from their trees that there must have been at least three independent Pacific basin invasions by gadid species. They suggested these invasions by Microgadus, Gadus, and Theragra species occurred during the re-opening of the Bering Strait at
about 3.0 to 3.5 million years ago (mya), and of the five hosts for these Loma species, that Microgadus (host of Loma sp. TOM) diverged earliest from other gadids in the Atlantic basin at about 12 mya, and later, perhaps 7-8 mya, Melanogrammus (host of $L$. branchialis HAD) diverged, followed by Gadus macrocephalus (host of Loma sp. PAC), 3.5 mya, and then Theragra (host of Loma sp. POL) diverged at about 3.0 mya.

These phylogeographic data help to explain why classification of the gadids has been difficult, at times calling for the recognition of several sub-species of Atlantic cod, and at other times calling for the elevation or transfer of subspecies to full species. The present study shows Loma species may have undergone a parallel phylogeographic history with their gadid hosts, and so may share some of their host's population (or species) substructuring (e.g. spores have been recognized to differ in size in different gadids from different localities, discussed above). If Loma species from gadids have undergone reticulate (hybridizing) evolution, as discussed above, perhaps they have done so with their hosts, thereby explaining why three fishes (Pacific cod, walleye pollock, and Atlantic cod) in Carr et al.'s (1999) study were diverged, but could not be properly resolved.

Using this data (see Fig. 4.20), the molecular evolutionary rate for gadid-parasitic Loma species can be estimated and used to calculate divergence times for other related species, under the assumption that rDNA evolves at a similar rate among very close relatives. Here, the rate can be estimated at about 1.958 base pairs change per 1344 base pairs of SSU rDNA ( $0.1458 \%$ ) per million years, as estimated from 23.5 bp differences between Loma sp. TOM and L. morhua, thought to have diverged 12 mya (see Figure 4.20). If rDNA evolves at a constant rate among close relatives and over time, then L. salmonae and Loma sp. BRO would have diverged about 8.2 mya, L. salmonae and L. morhua diverged about 9.2 mya, and L. morhua vs. Loma sp. AUS diverged about 39.3 mya.

Divergence time estimates can help estimate the time required for speciation in Loma. For example, close sister-species that are experimentally non-transmissible to reciprocal hosts (implying a reproductive boundary) can be compared to species that successfully transmit experimentally individuals of the same host species (implying no isolation). Loma salmonae, $L$. embiotocia and Loma sp. LIN (from lingcod) are sister-species that have been demonstrated to be non-transmissible to reciprocal hosts (Shaw et al., 2000c; S. W. Shaw personal communication). The closest genetic relationship among these was between $L$. salmonae and $L$. embiotocia ( $0.741 \% \mathrm{SSU}$ ), which represents a divergence of about 5.1 mya. The species most diverged intraspecifically among these was Loma sp. LIN ( $0.38 \%$ intraspecific SSU
divergence), which represents a divergence of about 2.6 mya, whereas for $L$. embiotocia and $L$. salmonae intraspecific divergences were about 1 mya and 100,000 years ago, respectively. These results suggest speciation can occur in 5.1 million years, or may not occur for up to 2.6 million years in Loma. This explains why the estimated 3 million year divergence between species in clade A (L. morhua from Atlantic cod, Loma sp. PAC from Pacific cod and Loma sp. POL from walleye pollock) may be too recent to be resolved phylogenetically in this study. These results also strengthen the evidence in support of validity of separate species of Loma in Atlantic gadids (Atlantic cod and haddock) that have been questioned in the past. For example, Loma sp. of Nilsen (2000) from the fourbeard rockling, Enchelyopus cimbrius, and L. morhua from Atlantic cod are about $0.752 \%$ ( 5.2 mya) diverged, and L. morhua from Atlantic cod and L. branchialis from haddock that are about $0.797 \%$ ( 5.46 mya) diverged.

The diversity in spore sizes and high intraspecific rDNA and EF-1 $\alpha$ sequence diversity of Loma species from gadid hosts examined thus far suggests additional Loma species might be found in gadids across their phylogenetic and geographic ranges. Further molecular data will be needed to resolve species boundaries and relationships in this group that appears to be quite successful in gadid hosts. Particular attention must be given to sequencing the ITS and LSU regions of the rDNA if this region is used, since it contains useful information that may answer population-genetic level and species-level questions, but it seems to contain potentially confounding indels and paralogs.

## CONCLUSIONS

This study presented DNA analyses of seven undescribed Loma species, helping to examine and characterize their divergence from one another and their relationships to six previously described Loma species. Ribosomal DNA and EF-1 $\alpha$ trees were largely in agreement, suggesting generalized host specialization and possible co-evolution with hosts. These data suggest that $L$. branchialis from haddock, $L$. morhua from Atlantic cod, Loma sp . from brook trout, Loma sp. from Pacific tomcod, and Loma sp. from an Australian surf bream may be valid, distinct species. However, the molecular data were inconclusive in separating undescribed species, including Loma spp. from lingcod and sablefish, and Loma spp. from Pacific cod, walleye pollock, and Atlantic cod. This result may suggest that either reproductive isolation is incomplete between these species or that the sequence length analyzed was
insufficient relative to the amount of mutation accumulation (i.e. they are recent species), because morphological, host, and in some cases geographic locality suggest these are separate species.

## FURTHER INVESTIGATION

Significant intraspecific variation observed in some of these Loma species would suggest that future studies attempting to examine species-boundaries should examine more individuals from a wider geographic range, to ensure that intraspecific polymorphism is distinguished from species-level divergence. This study also suggested there is great potential for development of population-genetic markers, and markers of strains or species using information from these rDNA sequences, if the rDNA paralogs can be better identified and characterized. While rDNA appears to be a valuable, highly variable marker that is easy to obtain from small or poorly fixed tissue samples, this study confirms results of others (Cheney et al. 2000; Cheney et al., 2001) suggesting that rDNA sometimes does not vary sufficiently between species, compared to the level of variation between paralogs; thus, good primers need to be developed for amplification of other independent, single-copy, nuclear loci. This study suggests such single-copy markers will be essential for species-boundaries questions in Loma and other microsporidia, helping to untangle problems of paralogs, multiple strains, or multiple species in the same host.

## ACKNOWLEDGEMENTS

This work was funded by the Natural Sciences and Engineering Research Council of Canada strategic grant 582073 to M. L. Adamson. I thank Dr. P. J. Keeling and his laboratory for their assistance with cloning. I am also grateful to Michael Coury for helping with software design. I am sincerely grateful to Dr. Ross Shaw, Dr. Jim Boutillier, and Sheila Dawe at the Pacific Biological Station and the crew of the research vessel W. E. Ricker for their assistance in identifying and collecting Pacific fishes. Robert Adlard kindly provided material from a species of Loma in Australia, and David Speare and J. Genaro Sánchez-Martinez generously provided infected gills from brook trout. I thank Dr. Stewart Johnson at the National Research Council in Halifax, for his assistance with collection of Atlantic cod and haddock heads, and for the use of his laboratory during the pungent work of examining them.

## LITERATURE CITED

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19:716-723.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Andreadis, T. G. 1998. Phylogeny of Amblyospora (Microsporida: Amblyosporidae) and related genera based on small subunit ribosomal DNA data: a possible example of host parasite cospeciation. Journal of Invertebrate Pathology 71:199-206.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Maddox, J. V. 1997. Phylogenetic position of Amblyospora Hazard \& Oldacre (Microspora: Amblyosporidae) based on small subunit rRNA data and its implication for the evolution of the microsporidia. Journal of Eukaryotic Microbiology 44(3):220-225.

Baker, M. D., Vossbrinck, C. R., Didier, E. S., Maddox, J. V. and Shadduck, J. A. 1995. Small subunit ribosomal DNA phylogeny of various microsporidia with emphasis on AIDS related forms. Journal of Eukaryotic Microbiology 42(5):564-570.

Bernardi, G., Holbrook, S. J., Schmitt, R. J., Crane, N. L. and DeMartini, E. 2002. Species boundaries, populations and colour morphs in the coral reef three-spot damselfish (Dasycyllus trimaculatus) species complex. Proceedings of the Royal Society of London Series B 269:599-605.

Bell, A. S., Aoki, T. and Yokoyama, H. 2001. Phylogenetic relationships among microsporidia based on rDNA sequence data, with particular reference to fish-infecting Microsporidium Balbiani 1884 species. Journal of Eukaryotic Microbiology 48(3):258-265.

Bradley, R. D. and Baker, R. J. 2001. A test of the genetic species concept: Cytochrome-b sequences and mammals. Journal of Mammalogy 82(4):960-973.

Brown, A. M. V. and Kent M. L. 2002. Molecular diagnostics for Loma salmonae and Nucleospora salmonis (microsporidia) In Molecular diagnostics of salmonid diseases. Cunningham, C. O. (ed.). Kluwer Academic Publishers, Dordrecht p. 267-283.

Canning, E. U. and Lom, J. 1986. The microsporidia of vertebrates. Academic Press, London. 289 pp.
Carr, S. M., Kivlichan, D. S., Pepin, P. and Crutcher, D. C. 1999. Molecular systematics of gadid fishes: implications for the biogeographic origins of Pacific species. Canadian Journal of Zoology 77(1):19-26.

Cheney, S. A., Lafranchi-Tristem, N. J., Bourges, D. and Canning, E. U. 2001. Relationships of microsporidian genera, with emphasis on the polysporous genera, revealed by sequences of the largest subunit of RNA polymerase II (RPB1). Journal of Eukaryotic Microbiology 48(1):111-117.

Cheney, S. A., Lafranchi-Tristem, N. J. and Canning, E. U. 2000. Phylogenetic relationships of Pleistophora-like microsporidia based on small subunit ribosomal DNA sequences and implications for the source of Trachipleistophora hominis infections. Journal of Eukaryotic Microbiology 47:280-287.

Chen, C. A., Chen, C.-P., Fan, T.-Y., Yu, T.-K. and Hseih, H.-L. 2002. Nucleotide sequences of ribosomal internal transcribed spacers and their utility in distinguishing closely related Perinereis polychaetes (Annelida; Polychaeta; Nereidae). Marine Biotechnology 4:17-29.

Cheynier, R., Kils-Hütten, L., Meyerhans, A. and Wain-Hobson, S. 2001. Insertion/deletion frequencies match those of point mutations in the hypervariable regions of the simian immunodeficiency virus surface envelope gene. Journal of General Virology 82:1613-1619.

Cho, S., Mitchell, A., Regier, J. C., Mitter, C., Poole, R. W., Friedlander, T. P. and Zhao, S. 1995. A highly conserved nuclear gene for low-level phylogenetics: elongation factor-l $\alpha$ recovers morphology-based tree for heliothine moths. Molecular Biology and Evolution 12(4):650-656.

Crandall, K. A. and Templeton, A. R. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134:959-969.

Diekmann, O. E., Bak, R. P. M., Stam, W. T. and Olsen, J. L. 2001. Molecular genetic evidence for probable reticulate speciation in the coral genus Madracis from a Caribbean fringing reef slope. Marine Biology 139:221-233.

Docker, M. F., Devlin, R. H., Richard, J., Khattra, J. and Kent, M. L. 1997a. Sensitive and specific polymerase chain reaction assay for detection of Loma salmonae (Microsporea). Diseases of Aquatic Organisms 29(1):41-48.

Dogiel, V. A. 1936. Parasites of cod from the relic lake Mogilny. Annals of Leningrad University no. 7 Biological Series. Problems of ecological parasitology 123-133. [RUSSIAN]

Fantham, H. B., Porter, A. and Richardson, L. R. 1941. Some microsporidia found in certain fishes and insects in eastern Canada. Parasitology 33:186-208.

Fomena, A., Coste, F. and Bouix, G. 1992. Loma camerounensis new species (Protozoa: Microsporida) a parasite of Oreochromis niloticus Linnaeus 1757 Teleost Cichlidae in fish-rearing ponds in Melen Yaounde Cameroon. Parasitology Research 78(3):201-208.

Goldman, N., Anderson, J. P. and Rodrigo, A. G. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49(4):652-670.

Hauck, A. K. 1984. Mortality and associated tissue reactions of chinook salmon, Oncorhynchus tshawytscha (Walbaum), caused by the microsporidian Loma sp. Journal of Fish Diseases 7:217-229.

Hilbish, T. J. 1996. Population genetics of marine species: the interaction of natural selection and historically differentiated populations. Journal of Experimental Marine Biology 200:67-83.

Holmes, E. C., Worobey, M. and Rambaut, A. 1999. Phylogenetic evidence for recombination in dengue virus. Molecular Biology and Evolution 16(3):405-409.

Hughes, K. W. and Peterson, R. H. 2001. Apparent recombination or gene conversion in the ribosomal ITS region of a Flammulina (Fungi, Agaricales) hybrid. Molecular Biology and Evolution 18:94-96.

Huson, D. H. 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14(1):68-73.
Inagaki, Y., Blouin, C., Susko, E. and Roger, A. J. 2003. Assessing functional divergence in EF-1 $\alpha$ and its paralogs in eukaryotes and archaebacteria. Nucleic Acids Research 31(14):4227-4237.

Johnson, G. D. 1993. Percomorph phylogeny: Progress and Problems. Bulletin of Marine Science 52(1):3-28.
Johnson, G. D. and Patterson, C. 1993. Percomorph phylogeny: A survey of Acanthomorphs and a new proposal. Bulletin of Marine Science 52(1):554-626.

Kabata, Z. 1959. On two little-known microsporidia of marine fishes. Parasitology 49:309-315.
Kamaishi, T., Hashimoto, T., Nadamura, Y., Nakamura, F., Murata, S., Okada, N., Okamoto, K., Shimizu, M. and Hasegawa, M. 1996. Protein phylogeny of translation elongation factor EF-1 $\alpha$ suggests microsporidians are extremely ancient eukaryotes. Journal of Molecular Evolution 42:257-263.

Kent, M. L., Docker, M., Khattra, J., Vossbrinck, C.R., Speare, D. J. and Devlin, R. H. 1999. A new Microsporidium sp. (Microsporidia) from the musculature of the Mountain Whitefish Prosopium williamsoni from British Columbia: morphology and phylogeny. Journal of Parasitology 85(6):1114-1119.

Kent, M. L., Traxler, G. S., Kieser, D., Richard, J., Dawe, S. C., Shaw, R. W., Prosperi-Porta, G., Ketcheson, J. and Evelyn, T. P. T. 1998. Survey of salmonid pathogens in ocean-caught fishes in British Columbia, Canada. Journal of Aquatic Animal Health 10:211-219.

Lom, J. 2002. A catalogue of described genera and species of microsporidians parasitic in fish. Systematic Parasitology 53:81-99.

Lom, J. and Nilsen, F. 2003. Fish microsporidia: fine structural diversity and phylogeny. International Journal for Parasitology 33:107-127.

Lom, J. and Laird, M. 1976. Parasitic protozoa from marine and euryhaline fish of Newfoundland and New Brunswick. II. Microsporida. Transactions of the American Microscopy Society 95(4):569-580.

Lom, J. and Pekkarinen, M. 1999. Ultrastructural observations on Loma acerinae (Jirovec, 1930) comb. nov. (Phylum Microsporidia). Acta Protozoologica 38:61-74.

Lopez, J. V., Peterson, C. L., Willoughby, R., Wright, A. E., Enright, E., Zoladz, S., Reed, J. K. and Pomponi, S. A. 2002. Characterization of genetic markers for in vitro cell line identification of the marine sponge Axinella corrugata. The Journal of Heredity 93(1):27-36.

Lydeard, C. and Roe, K. J. 1997. The phylogenetic utility of the mitochondrial cytochrome b gene for inferring relationships among Actinopterygian fishes. In Molecular systematics of fishes. Kocker, T. D. and Stepian, C. A. (eds.). Academic Press, San Diego. 314 pp .

Mallet, J. 1995. A species definition for the modern synthesis. Trends in Ecology and Evolution 10:294-299.
Matthews, J. L., Brown, A. M. V., Larison, K., Bishop-Stewart, J. K. and Kent, M. L. 2001. Pseudoloma neurophilia, n. gen., n. sp., a new microsporidium from the central nervous system of the zebrafish. Journal of Eukaryotic Microbiology 48:227-233.

Moreira, D., Le Guyader, H. and Philippe, H. 1999. Unusually high evolutionary rate of the elongation factor $1 \alpha$ genes from the ciliophora and its impact on the phylogeny of eukaryotes. Molecular Biology and Evolution 16(2):234-245.

Morrison, C. M. and Marryatt, V. 1986. Further observations on Loma morhua Morrison \& Sprague, 1981. Journal of Fish Diseases 9:63-67.

Morrison, C. M. and Sprague, V. 1981a. Electron microscopical study of a new genus and new species of microsporida in the gills of Atlantic cod Gadus morhua L. Journal of Fish Diseases 4:15-32.

Morrison, C. M. and Sprague, V. 1981b. Light and electron microscope study of microsporida in the gill of haddock, Melanogrammus aeglefinus (L.). Journal of Fish Diseases 4:179-184.

Morrison, C. M. and Sprague, V. 1981c. Microsporidian parasites in the gills of salmonid fishes. Journal of Fish Diseases 4:371-386.

Morrison, C. M. and Sprague, V. 1983. Loma salmonae (Putz, Hoffman and Dunbar, 1965) in the rainbow trout, Salmo gairdneri Richardson, and L. fontinalis sp. nov. (Microsporidia) in the brook trout, Salvelinus fontinalis (Mitchill). Journal of Fish Diseases 6:345-353.

Nelson, J. S. 1994. Fishes of the world. New York, J. Wiley. 600 pp.
Nilsen, F. 2000. Small subunit ribosomal DNA phylogeny of microsporidia with particular reference to genera that infect fish. Journal of Parasitology 86(1):128-133.

Nilsen, F., Endresen, C. and Hordvick, I. 1998. Molecular phylogeny of microsporidians with particular reference to species that infect the muscles of fish. Journal of Eukaryotic Microbiology 45:535-543.

Pomport-Castillon, C., De Jonkheere, J. F. and Romestand, B. 2000. Ribosomal DNA sequences of Glugea anomala, G. stephani, G. americanus and Spraguea lophii (Microsporidia): phylogenetic reconstruction. Diseases of Aquatic Organisms 40:125-129.

Posada, D. and Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14(9):817-818.

Poynton, S. L. 1986. Distribution of the flagellate Hexamita salmonis Moore, 1922 and the microsporidian Loma salmonae Putz, Hoffman and Dunbar, 1965 in brown trout Salmo trutta L., and rainbow trout, Salmo gairdneri Richardson, in the River Itchen (U.K.) and three of its fish farms. Journal of Fish Biology 29:417429.

Puorto, G., da Graça Salomão, M., Theakston, R. D. G., Thorpe, R. S., Warrell, D. A. and Wüster, W. 2001. Combining mitochondrial DNA sequences and morphological data to infer species boundaries: phylogeography of lanceheaded pitvipers in the Brazilian Atlantic forest, and the status of Bothrops pradoi (Squamata: Serpentes: Viperidea). Journal of Evolutionary Biology 14:527-538.

Quicke, D. L. J. 1993. Principles and techniques of contemporary taxonomy. Blackie Academic \& Professional. London. 311 pp .

Rodriquez-Robles, J. A. and de Jesus-Escobar, J. M. 2000. Molecular systematics of new world gopher, bull, and pinesnakes (Pituophis: Colubridae), a transcontinental species complex. Molecular Phylogenetics \& Evolution 14(1):35-50.

Rosenberg, N. A. and Nordburg, M. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nature Reviews Genetics 3:380-390.

Sánchez, J. G., Speare, D. J., Markham, R. J. F. and Jones, S. R. M. 2001 a. Experimental vaccination of rainbow trout against Loma salmonae using a live low-virulence variant of $L$. salmonae. Journal of Fish Biology 59:442-448.

Sánchez, J. G., Speare, D. J., Markham, R. J. F. and Jones, S. R. M. 2001b. Isolation of a Loma salmonae variant: biological characteristics and host range. Journal of Fish Biology 59:427-441.

Sang, T., Crawford, D. J. and Stuessy, T. F. 1995. Documentation of reticulate evolution in peonies (Paeonia) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proceedings of the National Academy of Sciences USA 92:6813-6817.

Shaw, R. W. and Kent, M. L. 1999. Fish microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 418-446.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000a. Viability of Loma salmonae (Microsporidia) under laboratory conditions. Parasitology Research 86:978-981.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000b. Innate susceptibility differences in chinook salmon Oncorhynchus tshawytscha to Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 43:49-53.

Shaw, R. W., Kent, M. L., Brown, A. M. V., Whipps, C. M. and Adamson, M. L. 2000c. Experimental and natural host specificity of Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 40:131-136.

Shaw, R. W., Kent, M. L., Docker, M. F., Brown, A. M. V., Devlin, R. H. and Adamson M. L. 1997. A new species of Loma (Microsporea) in shiner perch (Cymatogaster aggregata). Journal of Parasitology 83(2):296-301.

Shimodaira, H. and Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17(12):1246-1247.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology 51(3):492508.

Shulman, S. S. and Shulman-Albova, R. E. 1953. Fish parasites of the White Sea. Izdatel'stvo AN SSSR. MoskvaLeningrad, Russia, 192 p.

Sites, J. W. and Crandall, K. A. 1997. Testing species boundaries in biodiversity studies. Conservation Biology 11(6):1289-1297.

Speare, D. J., Beaman, H. J., Jones, S. R. M., Markham, R. J. F. and Arsenault, G. J. 1998a. Induced resistance in rainbow trout, Oncorhynchus mykiss (Walbaum), to gill disease associated with the microsporidian gill parasite Loma salmonae. Journal of Fish Diseases 21(2):93-100.

Speare, D. J., Daley, J., Markham, R. J. F., Sheppard, J., Beaman, H. J. and Sánchez, G. J. 1998b. Loma salmonaeassociated growth rate suppression in rainbow trout, Oncorhynchus mykiss (Walbaum), occurs during early onset xenoma dissolution as determined by in situ hybridization and immunohistochemistry. Journal of Fish Diseases 21(5):345-354.

Speare, D. J. and Daley, J. 2003. Failure of vaccination in brook trout Salvelinus fontinalis against Loma salmonae (Microspora). Fish Pathology 38(1):27-28.

Sprague, V. 1977. Annotated list of species of Microsporidia. In Comparative Pathology. 2. Systematics of the Microsporidia. Plenum Press, New York, New York, 333 p.

Sprague, V., Becnel, J. J. and Hazard, E. I. 1992. Taxonomy of phylum Microspora. Critical Reviews in Microbiology 18(5/6):285-395.

Swofford, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Templeton, A. R. 1989. The meaning of species and speciation: a genetic perspective. pgs. 3-27 in Otte, D. and Endler, J. A. (eds) Speciation and its consequences. Sinauer Associates, Sunderland, Massachusetts.

Templeton, A. R. 1994. The role of molecular genetics in speciation studies. pgs. 455-477 in Schierwater, B., Streit, B., Wagner, G. P. and DeSalle, R. (eds) Molecular ecology and evolution: approaches and applications. Bierkauser Verlag, Basel, Switzerland.

Van Oppen, M. J. H., Willis, B. L., van Vugt, H. W. J. A. and Miller, D. J. 2000. Examination of species boundaries in the Acropa cervicornis group (Scleractinia, Cnidaria) using nuclear DNA sequence analyses. Molecular Ecology 9:1363-1373.

Van Oppen, M. J. H., Willis, B. L., Van Rheede, T. and Miller, D. J. 2002. Spawning times, reproductive compatibilities and genetic structuring in the Acropora aspera group: evidence for natural hybridization and semi-permeable species boundaries in corals. Molecular Ecology 11:1363-1376.

Vernon, J. E. N. 1995. Corals of Australia and the Indo-Pacific. University of Hawaii Press, Honolulu.
Weiss, L. M. and Vossbrinck, C. R. 1999. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 129-171.

Wheeler, Q. D. and Meier, R. 2000. Species concepts and phylogenetic theory. Columbia University Press. New York, NY. 230 pp

Wu, C. I. 1991. Inferences of species phylogeny in relation to segregation of ancient polymorphisms. Genetics 127:429-435.

Table 4.1: Species of Loma, their hosts, isolates for PCR (separate host individuals) and collection data for ribosomal DNA (rDNA), elongation factor-1 alpha (EF-1 $\alpha$ ) and RNA polymerase largest subunit (RPB1) gene sequencing. Isolate labels correspond to those in phylogenetic analyses throughout this study. $\mathrm{T}=$ trawl number. Collectors: $\mathrm{S}=$ R.W. Shaw, $\mathrm{B}=$ A.M.V. Brown, $\mathrm{K}=$ M.L. Kent, $\mathrm{Sz}=\mathrm{G}$. Sánchez, $\mathrm{D}=\mathrm{S} . \mathrm{C}$. Dawe, $\mathrm{Bk}=\mathrm{D}$. Barker, $\mathrm{A}=\mathrm{R}$. Adlard. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=L$. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=L$. salmonae SV of Sánchez et al. (2001a).

| Loma species \& Host species | Isolate label | Locality | Date, collector \& notes |
| :---: | :---: | :---: | :---: |
| Loma salmonae (Putz, Hoffman \& Dunbar, 1965) |  |  |  |
|  | SPO | Laboratory, Nanaimo, BC | 5.97 B |
|  | N-14 | " | 7.97 S |
|  | C-24 | " | 9.97 S |
|  | BA2 | " | 11.97 B |
|  | BA3 | " | 4.99 B |
|  | BA4 | " | 6.99 B |
|  | BA9 | " | 4.00 B |
|  | I-27 | Indian Bay, BC | 10.97 S farm |
|  | IB-27 | " | " |
|  | KC-17 | Kuncchin, BC | 11.97 S hatchery |
|  | COR | Seashelt, BC | 2.98 S farm |
|  | L6 | California, USA | 6.96 K farm |
| salmon and trout Oncorhynchus spp. | CAL | " | " |
|  | Li | Idaho, USA | 3.96 K hatchery |
|  | Coll | Colorado, USA | 3.98 K hatchery |
|  | Col2 | " | " |
|  | Lsc2 | Chile | 12.00 K farm |
|  | S52 | Texada Island, BC | 9.97 D |
|  | C237 | Barkley Sound, BC | 5.97 B |
|  | S21 | Cape Mudge, BC | 9.97 D |
|  | S41 | Ballenas Island, BC | " |
|  | LOSO | Great Central L., BC | 11.96 S |
|  | SP-24 | Sproat River, BC | 7.97 S |
|  | ST-24 | Stamp River, BC | " |
|  | S71 | Boundary Bay, BC | 9.97 D |
|  | SF-25 | Fulton River, BC | " S |


| Loma embiotocia (Shaw, Kent, Docker, Brown, Devlin \& Adamson, 1997) |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  | SP3 | Eagle Bay, BC | 9.95 | S |
|  | SPN5 | Nanaimo, BC | 10.95 | S |
| shiner perch | PD | Maple Bay, Duncan, BC | 7.96 | S |
| Cymatogaster aggregata | SHI | West coast Van. Island, BC, T82* |  | B gonads |
| Gibbons 1854 | S103 | " T45 | 5.99 | B |

Loma sp. BRO (L. salmonae SV var. Sánchez et al. 2001)
LF1 Atlantic Vet. College, PEI 4.99 Sz
brook trout LF5 "

| Salvelinus fontinalis | LF7 | $"$ |
| :--- | :--- | :--- |
| (Mitchill, 1814) | LF8 | $"$ |

Loma sp. LIN (Loma sp. Kent et al. 1998, undescribed)

| lingcod | L10 | West coast Van. Island, BC, T55 | 5.97 | B |
| :--- | :--- | :--- | ---: | :--- |
| Ophiodon elongatus | L16 |  | T122 | $"$ |



Table 4.2: Primers used to amplify and sequence part of the rDNA cistron and partial EF-1 $\alpha$ and RPB1 genes for Loma species, showing source of primer. Primers are listed 5' to 3', and reference sequences used for primer positions are: for rDNA Loma salmonae reference sequence; for EF-1 $\alpha$ Glugea plecoglossi (Genbank D84253); and for RPB1 Loma acerinae (Genbank AJ278951). Primers designed in this study were created from conserved regions in alignments of a wide range of microsporidian and other taxa. * Docker, Kent et al. 1997 called it 18 eMIC in that paper.

| Primer | Sequence | Position | Source |
| :---: | :---: | :---: | :---: |
| rDNA forward |  |  |  |
| M5P | CAC CAG GTT GAT TCT GCC | -18-0 | Docker, Kent et al. 1997* |
| Seqlf | CGT TGT AGT TCT AGC AGT | 701-718 | provided by M. F. Docker |
| L7f | ATT AGT GAG ACC TCR GCC | 983-1000 | Docker, Devlin et al. 1997 |
| 580f | GAT AYA AGT CGT AAC AAG | 1299-1316 | this study |
| L1 | CTG GAT CAG ACC GAT TTA TAT | 1339-1359 | Docker, Devlin et al., 1997 |
| rDNA reverse |  |  |  |
| SeqR | AAC AGG GAC KYA TTC ATC | 1198-1215 | this study |
| 1492r | GGT TAC CTT GTT ACG ACT T | 1304-1450 | Baker et al., 1995 |
| L3R | CGA CTC CTG CAC ATT TCG | 1590-1610 | this study |
| L2 | ATG ACA TCT CAC ATA ATT GTG | 1590-1610 | Docker, Devlin et al. 1997 |
| 580r | GGT CCG TGT TTC AAG ACG G | 1830-1848 | Vossbrinck et al. 1987 |
| EF-1 $\alpha$ forward |  |  |  |
| EFZ | TTG CTT CAT TGG NCA CGT MGA | 32-52 | this study |
| EFE | AGA AAG AGG TAG WGG TWC | 143-160 | this study |
| EFV | GTA CAT ATC GTG gTa tTa C | 198-217 | this study |
| EF-1 $\alpha$ reverse |  |  |  |
| EFD | TGC ACC TGT ACT ACY CTN CCN GT | 806-828 | this study |
| EFW | AAG TCA CAT TTT CAC CTT T | 1203-1221 | this study |
| EFG | AGT TTC CAT KAC RAC TTG | 1241-1259 | this study |
| RPB1 forward |  |  |  |
| RPY | TTG TSC WGG KCA TTT TGG | -16-2 | this study |
| RPB1 reverse |  |  |  |
| RPF | GAG CCA TCA TGC TCA TTT | 1079-1096 | this study |

Table 4.3: Spore sizes of $L$. branchialis (Nemeczek, 1911) Morrison \& Sprague, 1981 and L. morhua Morrison \& Sprague, 1981 (synonym of $L$. branchialis in Canning \& Lom, 1986) from this and other studies, showing sizes converted using estimated shrinkage factors for each fixation method. Converted spore size range was calculated using $95 \%$ confidence intervals around the converted mean. Host abbreviations: Had = haddock, Melanogrammus aeglefinus; Atl = Atlantic cod, Gadus morhua; Kild = Kildin cod, G. morhua kildinensis; Green = Greenland cod, Gadus ogac. Fixation methods: $\mathrm{Fz}=\mathrm{frozen} ; \mathrm{Fsh}=\mathrm{fresh} ; \mathrm{Fm}=$ formalin; $\mathrm{R}=$ resin; $\mathrm{Et}=$ ethanol. Spore length in $\mu \mathrm{m}$. Xenoma size in $\mathrm{mm} . *(\mathrm{Atl})=$ authors reported haddock were more often infected. $\bullet$ originally named $L$. morhua.

| Author | Host | Locality | Fixation | Xenoma Size | Spore Length | Converted Spore Length (this study) |  | Spore Width | Converted Spore Width (this study) |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  | L. morhua <br> (in Atlantic <br> cod) | L. branchialis (in Haddock) |  | L. morhua (in Atlantic cod) | L. branchialis (in Haddock) |
| Nemeczek, 1911 | Had | America/ <br> Russia | Fz | 0.2-0.5, 1 | $6: 3$ | 5.16-5.24 | 4.85-5.09 | $3.5$ | 2.89-3.11 | 2.83-2.97 |
| Bazikalova, 1932 | Had \& $(\mathrm{Atl})^{*}$ | Barents Sea | - | - |  | - | - |  | - | - |
| Dogiel, 1936 | Kild | Barents Sea | Fz | - | 54-6\% | 5.16-5.24 | 4.85-5.09 | EI. | - | - |
| Fantham et al. $1941$ | Atl | Gulf of St. <br> Lawrence | Fsh | - | $57.66$ | 5.83-6.13 | 5.54-5.86 | $3.5-42$ | 3.13-3.35 | 3.06-3.22 |
| Shulman \& Shulman-Albova, 1953 | Green | White Sea | Fm | - | $43 \cdot 5.1$ | n/a | n/a | $2-2.3$ | n/a | n/a |
| Kabata, 1959 | Atl | Southern Iceland | Fm | 0.5-1.2 | $4.9$ | n/a | n/a | $2.4$ | n/a | n/a |
| Lom \& Laird, $1976$ | Had | New <br> Brunswick | Fsh | 1.2 | $4.8$ | $5.83-6.13$ | 5.54-5.86 | $2.3$ | 3.13-3.35 | 3.06-3.22 |
|  <br> Sprague, 1981a | Atl | Halifax, NS | R | 0.05-0.15 | $42$ | 3.36-3.40 | 3.15-3.31 | $20$ | 1.65-1.77 | 1.61-1.69 |
| Morrison \& Sprague, 1981b | Had | Halifax, NS | R | - | $44$ | 3.36-3.40 | 3.15-3.31 | $2.1$ | 1.65-1.77 | 1.61-1.69 |
| this study | Atl | Halifax, NS | Et | - | $432$ | n/a | n/a | $243$ | n/a | n/a |
| this study | Had | Halifax, NS | Et | - | 4.13 | n/a | n/a | 2.35 \% | n/a | $\mathrm{n} / \mathrm{a}$ |

Table 4.4: Length of DNA sequenced (number of alignment positions) for 5 gene regions in Loma species, including ribosomal DNA (rDNA), small subunit (SSU), internal transcribed spacer (ITS), and partial large subunit (LSU), elongation factor-1 alpha (EF-1 $\alpha$ ), and RNA polymerase largest subunit (RPB1): Chapters $2 \& 3$ name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=L$. pacificodae, Loma sp. $\mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \operatorname{LIN}=$ L. lingcodae, Loma sp. BLK $=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

| Length of region sequenced <br> (number of alignment positions) |  |  |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: |
| Loma species | Total <br> rDNA <br> sequenced | SSU | ITS | LSU | EF-1 $\alpha$ | RPB1 |
| Loma sp. AUS | 1157 | 677 | 37 | 443 | - | - |
| L. salmonae | 1830 | 1342 | 37 | 451 | 1132 | 836 |
| Loma sp. POL | 1842 | 1341 | 41 | 460 | 1021 | - |
| Loma sp. BRO | 1830 | 1341 | 37 | 452 | 1030 | - |
| Loma sp. PAC | 1843 | 1338 | 45 | 460 | 1016 | - |
| L. morhua ATL | 1846 | 1341 | 45 | 460 | 656 | - |
| L. embiotocia | 1832 | 1341 | 45 | 446 | 951 | - |
| Loma sp. TOM | 1842 | 1343 | 39 | 460 | 1051 | - |
| Loma sp. BLK | 1113 | 624 | 37 | 452 | - | - |
| Lom $a$ sp. LIN | 1835 | 1346 | 37 | 452 | 965 | - |
| L. branchialis HAD | 1128 | 625 | 45 | 458 | - | - |

Table 4.5: Intraspecific differences within Loma species for 4 gene regions: ribosomal DNA small subunit (SSU), internal transcribed spacer (ITS), and large subunit (LSU) 5' region, and elongation factor-1alpha (EF-1 $\alpha$ ), calculated by adding substitutional and indel differences for each species compared to a single reference sequence* for that species, and dividing this value by the number of nucleotides sequenced for that species across all isolates, clones or PCR products. *The single reference sequence for each Loma species was created using an alignment of all sequence fragments by taking the most common character (nucleotide or indel) at each position, among all sequence fragments. Chapters $2 \& 3$ name equivalents: Loma sp . $\mathrm{PAC}=L$. pacificodae, Loma sp. POL $=$ L. wallae, Loma sp. $\mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \operatorname{LIN}=L$. lingcodae, Loma sp. $\mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

|  | Intraspecific difference <br> (\% difference from reference <br> sequences) |  |  |  | Number of nucleotides sequenced <br> (all isolates, clones or PCR <br> products) |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| Loma species | SSU |  |  | ITS | LSU | EF-1 $\boldsymbol{\alpha}$ | SSU | ITS |
| LSU | EF-1 $\boldsymbol{\alpha}$ |  |  |  |  |  |  |  |
| Loma sp. AUS | 0 | 0 | 0 | - | 1638 | 111 | 437 | - |
| L. salmonae | 0.02 | 0 | 0.11 | 0.03 | 35320 | 2351 | 21228 | 27374 |
| Loma sp. POL | 0.06 | 4.05 | 0.46 | 0.18 | 10400 | 593 | 4216 | 3653 |
| Loma sp. BRO | 0.09 | 0.21 | 0.14 | 0 | 16776 | 481 | 4281 | 1189 |
| Loma sp. PAC | 0.09 | 2.80 | 0.89 | 0.46 | 9392 | 543 | 3201 | 2276 |
| L. morhua ATL | 0.14 | 3.50 | 0.36 | 0.34 | 16605 | 1429 | 11280 | 1916 |
| L. embiotocia | 0.15 | 0.20 | 0 | 0.16 | 8433 | 495 | 2502 | 1233 |
| Loma sp. TOM | 0.17 | 1.07 | 0.90 | 0.06 | 7237 | 468 | 3388 | 3407 |
| Loma sp. BLK | 0.29 | 0.41 | 0.50 | - | 4445 | 370 | 3015 | - |
| Loma sp. LIN | 0.38 | 0.66 | 0.45 | 0.22 | 12436 | 610 | 4767 | 1125 |
| L. branchialis HAD | 0.40 | 3.93 | 0.31 | - | 5671 | 585 | 4504 | - |

Table 4.6: Inter- and intraspecific insertions or deletions (indels) in ribosomal DNA sequence of species of Loma. Position numbers refer to L. salmonae reference sequence. Sites with parsimony-informative indels for interspecific comparisons are shaded lightly, and species with intraspecific indels (mostly parsimony-informative at intraspecific level) are shaded darkly. Indels present in some but not all of the isolates of a species are indicated on separate lines. For example, in L. morhua, the second line which shows dashes at positions 1381-1387 indicates an indel 6 sites long that was present in some of the isolates and only 4 sites long in other isolates. Chapters $2 \& 3$ name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma sp. $\mathrm{POL}=L$. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=L$. salmonae SV of Sánchez et al. (2001a).

| Position in: | SSU |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ITS |  |  |  |  |  | LSU |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Species: |  |  |  |  | $10 \mid$ | $\begin{array}{l\|l\|l\|l\|l\|l\|l\|} 5 & 5 \\ 3 & 3 \\ 2 \end{array}$ | $\begin{array}{\|l\|l} 5 & 6 \\ 3 & 6 \\ 7 & 1 \\ \hline \end{array}$ | $\begin{array}{l\|l\|} 6 & 7 \\ 6 & 5 \\ 1 & 2 \end{array}$ | $\begin{array}{llll} 8 & 8 & 8 & 8 \\ 8 & 9 & 9 & 9 \\ 9 & 0 & 1 & 2 \end{array}$ | 9 <br> 6 <br> 5 |  | 1 2 1 | $\left[\begin{array}{l} 1 \\ 2 \\ 8 \\ 8 \\ 0 \end{array}\right]$ | $\left[\begin{array}{ll} 1 \\ 2 & 1 \\ 9 & 3 \\ 9 \end{array}\right]$ | $\left[\begin{array}{l} 1 \\ 3 \\ 4 \\ 9 \\ 9 \end{array}\right]$ | $\begin{aligned} & 1 \\ & 6 \\ & 6 \end{aligned}$ | 3 | 3 | 71414141 3.3 3 | $\left.\begin{array}{\|c} 1 \\ 3 \\ 9 \\ 1 \end{array}\right]$ | \|lllll |  | $\begin{array}{\|ll\|} \hline 1 & 1 \\ 4 & 4 \\ 6 & 6 \\ 5 & 6 \\ \hline \end{array}$ | $\begin{array}{\|l\|} \hline 145 \\ 51 \\ 23 \\ \hline 1 \end{array}$ | $\begin{array}{\|lll} 1 & 1 & 1 \\ 5 & 5 & 5 \\ 7 & 7 & 7 \\ 7 & 8 & 9 \end{array}$ | $\begin{aligned} & 146 \\ & 662 \\ & 23 \\ & 23 \end{aligned}$ | $111111 \% 111$ 66666666 22222233 4567.8901 | $\left.\begin{array}{\|ll} 1 & 1 \\ 6 & 6 \\ 3 & 3 \\ 2 & 3 \end{array} \right\rvert\,$ | 187  <br> 7 1 <br> 0 7 <br> 0 7 <br> 5.65  |
| L. morhua ATL | $\frac{x}{4}$ |  |  |  | - |  | c - | -- | ---- |  |  |  |  |  |  |  |  |  | GYATATG |  | A T |  |  |  |  |  | TAATTAATA |  | --A |
| Loma sp. PAC | T |  |  |  |  |  | $C C$ |  |  |  |  |  |  | G |  |  |  |  |  |  | A T |  | - - |  | G T A |  | TAATTAA | A T | A |
| Loma sp. POL | T |  | G |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | A 7 |  | - - |  |  | Tin | TAATTA | A T | A |
| L. branchialis HAD |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ${ }^{\text {A }}$ | - | - - |  | G T A |  | TAATTAATA | A T | A |
| Loma sp. TOM |  |  | $\frac{10}{4} \mathrm{G}$ |  |  |  | C- | - | - | T |  | C |  |  |  |  | A |  | -- - - ${ }^{\text {G }}$ |  | A T |  | - - |  | G T A |  | FATTIAAT |  | A |
| Loma sp. LIN |  |  | $-6$ |  |  | c | C- | -- | ARci | $5$ |  | c |  | G |  | T |  | - | ------ |  | A 7 | - | - - |  | ¢Ta |  | $\cdots$ | A A | A |
| Loma sp. BLK |  |  |  |  |  |  |  |  |  | 4 |  | c |  |  |  |  |  | - | - |  | A T | - | - - | G T | $\mathrm{GNA}^{\mathrm{A}}$ |  | --- | A A | A |
| 2. salmonae |  |  | $G$ |  |  |  | C- |  |  |  |  |  |  |  |  |  |  |  | - |  | A T | - | - - |  | G T A |  | - - - - - - |  | - ${ }^{\text {a }}$ |
| Loma sp. BRO | T |  |  |  |  |  | C - | - - |  | T |  | c |  |  | 4 |  |  | - | - - - - - - | - | $2$ | - | - |  | G T A |  | ------ | G T | A T A |
| L. embiotocia | T |  | -G |  |  |  | C- |  | --- | T |  | c |  |  |  |  |  |  | G TATA - T |  |  |  | -- |  | G T A |  | - | A A | --A |
| Loma sp. AUS |  |  |  |  |  |  |  |  | ----1 | T | G. | c |  | G |  |  |  |  | ----- |  | A $T$ |  | G T | -- | A T A | - - | ---...--- | -- | - - A |

Table 4.7: Inter- and intraspecific polymorphic sites in rDNA sequence, showing only those sites with both a polymorphism and the 2 parental or alternate states (= "additive sites"), e.g. a $\mathrm{T}, \mathrm{A}$, and W together at one site, or where the additivity is shared by at least 2 isolates per species (for intraspecific additivity) or 2 species (for intraspecific additivity). Sites shaded in the lightest gray show additivity across species only (for just the species that are shaded), sites shaded in darker gray show intraspecific additivity, and positions with both inter- and intraspecific additivity are shaded in black. Letters representing 2 states follow the universal degenerate code: $\mathrm{R}=(\mathrm{A}$ and G$) \mathrm{Y}=(\mathrm{C}$ and T$) \mathrm{S}=(\mathrm{C}$ and G$) \mathrm{W}=(\mathrm{A}$ and T$) \mathrm{M}=(\mathrm{C}$ and A$) \mathrm{K}=(\mathrm{G}$ and T , and the order of letters from top to bottom, for a given species, represent the most to the least common character state observed, respectively. Chapters 2 \& 3 name equivalents: Loma sp. $\mathrm{PAC}=L$. pacificodae, Loma sp. $\mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=L$. salmonae SV of Sánchez et al. (2001a).


## Ribosomal DNA sequenced:



Figure 4.1: Diagram of ribosomal DNA small subunit (SSU), internal transcribed spacer (ITS) and partial large subunit (LSU) gene regions sequenced, and regions of the alignment that were used in phylogenetic analyses, showing nucleotide positions relative to the master alignment of Loma species for gene and alignment regions.


Figure 4.2: Consensus 50\% majority-rule distance tree from 643 alignment positions from rDNA Region 1, showing in bold the branches also shared in maximum parsimony (MP) and maximum likelihood (ML) analyses, with distance (ME), MP and ML bootstrap values $>50 \%$ shown on branches (from 1000 replicates with heuristic search in MP and ME, and 500 replicates in ML). Note weakly supported clades $A$ and $P$ grouping several species from gadid hosts. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=L$. pacificodae, Loma sp. $\mathrm{POL}=L$. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=L$. kenti, Loma $\mathrm{sp} . \operatorname{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=L$. salmonae SV of Sánchez et al. (2001a).


Figure 4.3: Consensus 50\% majority-rule distance tree from 628 alignment positions from rDNA Region 2, showing in bold the branches also shared in maximum parsimony (MP) and maximum likelihood (ML) analyses, with distance (ME), MP and ML bootstrap values >50\% shown on branches (from 1000 replicates with heuristic search in MP and ME, and 500 replicates in ML). Note weakly supported clades G from gadid hosts (with sub-clades A, NF, H, and T) and clade B. * indicates one sequence that variably sits inside or outside clade H . Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=L$. kenti, Loma $\mathrm{sp} . \operatorname{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma sp. $\mathrm{BRO}=L$. salmonae SV of Sánchez et al. (2001a).


Figure 4.4: Consensus $50 \%$ majority-rule distance tree from 479 alignment positions from rDNA Region 3, showing in bold the branches also shared in maximum parsimony (MP) and maximum likelihood (ML) analyses, with distance (ME), MP and ML bootstrap values $>50 \%$ shown on branches ( 1000 replicates with heuristic search in MP and ME, 100 replicates in ML), except on terminal nodes where " $\bullet$ " indicates $>50 \%$ support by at least one method. Clades G (subclades A, NF, H and T) and L (subclade B ) are illustrated. Loma sp . $\mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=L$. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=L$. lingcodae, Loma sp. BLK $=$ L. richardi, Loma sp. BRO = L. salmonae SV of Sánchez et al. (2001a).


Figure 4.5: Consensus 50\% majority-rule tree (maximum parsimony) generated from 27 gap characters (encoded gap ends only, no nucleotide substitutional data) generated using the GapMatrix program on 479 alignment positions of rDNA Region 3. Bootstrap values ( 1000 replicates with heuristic search) are shown on branches. Clade B and a clade containing all species from gadid hosts correspond to those in previous analyses. Clades A, P, H, and T (previous analyses) appear unresolved. Chapters $2 \& 3$ name equivalents: Loma sp.PAC $=L$. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma sp.BLK = L. richardi, Loma sp.BRO = L. salmonae SV of Sánchez et al. (2001a).


Figure 4.6: Consensus $50 \%$ majority-rule tree (maximum parsimony) from 479 alignment positions plus a gap matrix of 27 additional characters, from rDNA Region 3, showing 5 paralogous relationships with parentheses (see further details in Fig. 4.7), and bootstrap values $>50 \%$ ( 1000 replicates) or " $\bullet$ " on branches. Clades G (subclades A, NF, H and T) and L (subclade B) are identical to those in Fig. 4.4 (not labeled here for clarity). Loma sp.PAC $=L$. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\operatorname{sp} . \mathrm{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).


Figure 4.7: Details from Figure 4.6 showing 5 examples of paralogous relationships among isolates of Loma species. Boxes enclose Loma isolate and sequence numbers for which paralogous rDNA copies diverged at branches indicated by stars. For example, isolate "Aa" of L. morhua (see isolate labels in Table 4.1) is represented 4 times in this figure, with sequences 4 and 5 in paralogous relationship 1 and sequences 6 and 7 in paralogous relationship 2. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=L$. wallae, Loma sp . $\mathrm{TOM}=L$. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=L$. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=L$. salmonae SV of Sánchez et al. (2001a).


Figure 4.8: Maximum parsimony, distance and maximum likelihood trees for 966 bp of elongation factor $1-\alpha(E F-1 \alpha)$ for species of Loma, with Glugea plecoglossi as outgroup, showing overall branch length differences, and variation in tree topology, depending on method used. Branche to G. plecoglossi shortened here for display purposes. Loma sp.PAC $=L$. pacificodae, Loma sp.POL = L. wallae, Loma sp. $\cdot \mathrm{TOM}=$ L. kenti, Loma $\operatorname{sp} . \mathrm{LIN}=L$. lingcodae, Loma sp.BLK $=L$. richardi, Loma sp.BRO $=L$. salmonae SV of Sánchez et al. (2001a).


Figure 4.9: Consensus 50\% majority-rule tree (maximum parsimony) generated from 966 alignment positions from the partial EF-1 $\alpha$ gene (nucleotide sequence), showing in bold the branches also shared in maximum parsimony (MP) and maximum likelihood (ML) analyses, with bootstrap values $>50 \%$ shown on branches (from 1000 replicates with heuristic search in MP and ME, and 500 replicates in ML). Clades L, G, and A correspond to those presented in rDNA trees. Chapters $2 \& 3$ name equivalents: Loma sp . PAC $=$ L. pacificodae, Loma sp . POL $=$ L. wallae, Loma sp. $\mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma sp. $\mathrm{BLK}=L$. richardi, Loma sp. BRO = L. salmonae SV of Sánchez et al. (2001a).

Figure 4.10: Tree topologies used for Loma species boundaries tests from rDNA Region 1. Monophyletic groups (shown in boxes), corresponding to various hypotheses, were created as constraints prior to heuristic tree searches.. From these topologies likelihood scores (site likelihood scores in PAUP* 4.0-10b Swofford, 2000) were calculated under the best-fit substitution model using parameters estimated from the data in Modeltest V. 3.06 (Posada \& Crandall, 1998) for statistical tests of best trees using the CONSEL V.0.1.f software package (Shimodaira \& Hasegawa, 2001) (see Table 4.8). Chapters 2 \& 3 name equivalents: Loma sp. $\mathrm{PAC}=$ L. pacificodae, Loma sp. $\mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).


Table 4.8: Comparison of unconstrained and monophyly-constrained trees from rDNA Region 1 for Loma species, both without gap matrix and with addition of a gap matrix, using the Approximately Unbiased or "AU", weighted and unweighted Kishino-Hasegawa or "KH" and Shimodaira-Hasegawa or "SH" tests in the CONSEL software package. Tree numbers correspond to trees in Figure 4.10. Conclusion, based on AU-test, with rejection at the p-value of $\alpha=0.01$. Chapters $2 \& 3$ name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}$ $=L$. wallae, Loma sp. $\mathrm{TOM}=L$. kenti, Loma $\mathrm{sp} . \operatorname{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

|  |  | Test results (p-values) |  |  |  |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Tree <br> $\#$ | Constraint: | AU | KH | SH | WKH | WSH | Conclusion |
| without gap matrix |  |  |  |  |  |  |  |
| 1 | unconstrained | 0.522 | 0.395 | 0.796 | 0.395 | 0.793 | do not reject |
| 2 | Loma sp. POL together | 0.645 | 0.605 | 0.787 | 0.605 | 0.802 | do not reject |
| $\mathbf{3}$ | Loma sp. BRO sister to L. salmonae | 0.040 | 0.235 | 0.576 | 0.154 | 0.469 | do not reject |
| 4 | Loma sp. BRO internal to $L$. salmonae | $2 \mathrm{e}-074$ | 0.004 | 0.005 | 0.001 | 0.003 | REJECT |
| $\mathbf{5}$ | Loma sp. from gadids together | 0.298 | 0.288 | 0.546 | 0.288 | 0.596 | do not reject |
| gap matrix added |  |  |  |  |  |  |  |
| 6 | unconstrained | 0.424 | 0.345 | 0.762 | 0.345 | 0.701 | do not reject |
| 7 | Loma sp. POL together | 0.409 | 0.413 | 0.608 | 0.413 | 0.597 | do not reject |
| 8 | Loma sp. BRO sister to L. salmonae | 0.599 | 0.655 | 0.913 | 0.587 | 0.917 | do not reject |
| 9 | Loma sp. BRO internal to $L$. salmonae | $1 \mathrm{e}-010$ | 0.001 | 0.007 | 0.001 | 0.001 | REJECT |

Figure 4.11: Tree topologies used for Loma species boundaries tests on rDNA Region 2. Monophyletic groups corresponding to various hypotheses created as constraints prior to heuristic tree searches are shown in boxes. Likelihood scores (using a best-fit substitution model from Modeltest) from these topologies were used to statistically test best trees (see Table 4.9) using the CONSEL software package. Chapters $2 \& 3$ name equivalents: Loma sp. PAC $=L$. pacificodae, Loma sp. $\mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma sp. LIN $=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).







Table 4.9: Comparison of unconstrained and monophyly-constrained trees from rDNA Region 2 for Loma species, both without gap matrix and with addition of a gap matrix, using the Approximately Unbiased or "AU", weighted and unweighted Kishino-Hasegawa or "KH" and Shimodaira-Hasegawa or "SH" tests in the CONSEL software package. Tree numbers correspond to trees in Figure 4.11. Chapters $2 \& 3$ name equivalents: Loma sp. PAC $=L$. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=L$. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

| Tree <br> $\#$ | Constraint: | AU | KH | SH | WKH | WSH | Conclusion |
| :---: | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| without gap matrix <br> 1 | unconstrained |  |  |  |  |  |  |
| 2 | Loma sp. POL together | 0.821 | 0.511 | 0.967 | 0.511 | 0.967 | do not reject |
| 3 | Loma sp. PAC together | 0.736 | 0.381 | 0.966 | 0.381 | 0.966 | do not reject |
| 4 | L. morhua ATL together | 0.751 | 0.489 | 0.981 | 0.489 | 0.981 | do not reject |
| 5 | L. branchialis with $L$. morhua | 0.068 | 0.153 | 0.836 | 0.153 | 0.545 | do not reject |
| 6 | Loma sp. LIN separate from BLK | 0.005 | 0.034 | 0.169 | 0.034 | 0.078 | REJECT |
| 7 | Loma sp. BRO sister to L. salmonae | 0.006 | 0.076 | 0.229 | 0.035 | 0.083 | REJECT |
| 8 | Loma sp. BRO within $L$. salmonae | $1 \mathrm{e}-005$ | 0.010 | 0.055 | 0.010 | 0.020 | REJECT |
| 9 | L. branchialis "HA2" with L.salmonae | 0.387 | 0.376 | 0.869 | 0.376 | 0.867 | do not reject |
| 10 | L. branchialis "HA2" with TOM | 0.735 | 0.381 | 0.966 | 0.381 | 0.966 | do not reject |
| 11 | Loma sp. TOM polyphyletic with HAD | 0.821 | 0.511 | 0.967 | 0.511 | 0.967 | do not reject |
| with gap matrix |  |  |  |  |  |  |  |
| 12 | unconstrained | 0.836 | 0 | 0.995 | 0 | 0.995 | do not reject |
| 13 | Loma sp. POL together | 0.836 | 0.304 | 0.989 | 0.304 | 0.989 | do not reject |
| 14 | Loma sp. PAC together | 0.244 | 0.216 | 0.886 | 0.216 | 0.695 | do not reject |
| 15 | L. morhua ATL together | 0.040 | 0.103 | 0.626 | 0.103 | 0.383 | do not reject |
| 16 | L. branchialis with L. morhua | 0.004 | 0.056 | 0.181 | 0.056 | 0.128 | REJECT |
| 17 | Loma sp. LIN separate from BLK | 0.069 | 0.104 | 0.529 | 0.104 | 0.247 | do not reject |
| 18 | Loma sp. BRO sister to L. salmonae | 0.010 | 0.076 | 0.267 | 0.043 | 0.106 | REJECT |
| 19 | Loma sp. BRO within $L$. salmonae | $7 \mathrm{e}-005$ | 0.012 | 0.064 | 0.012 | 0.025 | REJECT |
| 20 | L. branchialis "HA2" with L.salmonae | 0.425 | 0.376 | 0.843 | 0.376 | 0.830 | do not reject |
| 21 | L. branchialis "HA2" with TOM | 0.836 | 0 | 0.995 | 0 | 0.995 | do not reject |
| 22 | Loma sp. TOM polyphyletc with HAD | 0.107 | 0.147 | 0.685 | 0.147 | 0.410 | do not reject |

Figure 4.12: Tree topologies used for Loma species boundaries tests on rDNA Region 3. Boxes show monophyletic groups, corresponding to various hypotheses created as constraints prior to heuristic tree searches. Likelihood scores from these topologies were calculated using a best-fit substitution model from Modeltest and parameters estimated from the data to statistically test best trees (see Table 4.10) using the CONSEL software package. Chapters $2 \& 3$ name equivalents: Loma sp. PAC $=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=L$. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=L$. kenti, Loma sp. LIN $=$ L. lingcodae, Loma sp. BLK $=L$. richardi, Loma sp. BRO $=L$. salmonae SV of Sánchez et al. (2001a).













Table 4.10: Comparison of unconstrained and monophyly-constrained trees from rDNA Region 3 for Loma species, both without gap matrix and with addition of a gap matrix, using the Approximately Unbiased or "AU", weighted and unweighted Kishino-Hasegawa or "KH" and Shimodaira-Hasegawa or "SH" tests in the CONSEL software package. Tree numbers correspond to trees in Figure 4.12. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=L$. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=L$. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

| $\begin{gathered} \hline \text { Tree } \\ \# \end{gathered}$ | Constraint: | AU | KH | SH | WKH | WSH | Conclusion |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| without gap matrix |  |  |  |  |  |  |  |
| 1 | unconstrained | 0.039 | 0.110 | 0.813 | 0.110 | 0.556 | do not reject |
| 2 | Loma sp. POL together | 0.761 | 0.612 | 0.956 | 0.612 | 0.959 | do not reject |
| 3 | Loma sp. PAC together | 0.542 | 0.388 | 0.935 | 0.388 | 0.919 | do not reject |
| 4 | L. morhua ATL together | 0.112 | 0.088 | 0.824 | 0.088 | 0.404 | do not reject |
| 5 | Loma sp. BRO sister to L. salmonae | 0.168 | 0.147 | 0.738 | 0.147 | 0.588 | do not reject |
| 6 | Loma sp. BLK and LIN separated | 0.039 | 0.110 | 0.813 | 0.110 | 0.556 | do not reject |
| 7 | Loma sp. TOM paraphyletic | 0.187 | 0.170 | 0.783 | 0.170 | 0.620 | do not reject |
| 8 | L. morhua ATL "NF" polyphyletic | 0.003 | 0.011 | 0.039 | 0.011 | 0.032 | REJECT |
| 9 | Loma sp. BRO internal to L. salmonae | 0.003 | 0.042 | 0.305 | 0.042 | 0.136 | REJECT |
| 10 | Loma spp. grouped by indel 1 (ITS) | 0.007 | 0.023 | 0.168 | 0.023 | 0.056 | REJECT |
| 11 | Loma spp. grouped by indel 2 (LSU) | $7 \mathrm{e}-005$ | 0.010 | 0.021 | 0.004 | 0.008 | REJECT |
| gap matrix added |  |  |  |  |  |  |  |
| 12 | unconstrained | 0.510 | 0.244 | 0.941 | 0.244 | 0.890 | do not reject |
| 13 | Loma sp. POL together | 0.004 | 0.013 | 0.086 | 0.013 | 0.071 | REJECT |
| 14 | Loma sp. PAC together | $7 \mathrm{e}-012$ | 0.012 | 0.050 | 0.002 | 0.011 | REJECT |
| 15 | L. morhua ATL together | 0.010 | 0.018 | 0.088 | 0.018 | 0.087 | REJECT |
| 16 | Loma sp. BRO and L.sal. separated | 0.016 | 0.017 | 0.248 | 0.017 | 0.072 | do not reject |
| 17 | Loma sp. BLK and LIN separated | 0.374 | 0.310 | 0.657 | 0.310 | 0.678 | do not reject |
| 18 | Loma sp. TOM paraphyletic | 0.796 | 0.756 | 0.980 | 0.690 | 0.985 | do not reject |
| 19 | L. morhua ATL "NF" polyphyletic | $3 \mathrm{e}-092$ | $3 \mathrm{e}-004$ | 0.004 | 3e-004 | 0.001 | REJECT |
| 20 | Loma sp. BRO internal to L. salmonae | 0.076 | 0.101 | 0.661 | 0.101 | 0.441 | do not reject |
| 21 | Loma spp. grouped by indel 1 (ITS) | 0.113 | 0.080 | 0.450 | 0.080 | 0.317 | do not reject |
| 22 | Loma spp. grouped by indel 2 (LSU) | 0.018 | 0.019 | 0.159 | 0.019 | 0.071 | do not reject |



Figure 4.13: Tree topologies used for Loma species boundaries tests on the partial EF-1 $\alpha$ gene. Boxes show monophyletic groups, corresponding to various hypotheses created as constraints prior to heuristic tree searches. Likelihood scores from these topologies were calculated using a best-fit substitution model from Modeltest and parameters estimated from the data to statistically test best trees (see Table 4.11) using the CONSEL software package. Loma $\operatorname{sp} . \mathrm{PAC}=L$. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma sp.LIN $=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

Table 4.11: Comparison of unconstrained and monophyly-constrained trees from the partial EF$1 \alpha$ gene for Loma species, using the Approximately Unbiased or "AU", weighted and unweighted Kishino-Hasegawa or "KH" and Shimodaira-Hasegawa or "SH" tests in the CONSEL software package. Tree numbers correspond to those in Figure 4.13. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}$ $=L$. kenti, Loma sp. $\mathrm{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=L$. salmonae SV of Sánchez et al. (2001a).

| Tree <br> $\#$ | Constraint: | AU | KH | SH | WKH | WSH | Conclusion |
| :---: | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 1 | unconstrained | 0.005 | 0.049 | 0.087 | 0.049 | 0.094 | REJECT |
| 2 | L. morhua ATL together | $4 \mathrm{e}-004$ | 0.060 | 0.087 | 0.052 | 0.134 | REJECT |
| 3 | Loma sp. POL together | 0.005 | 0.081 | 0.424 | 0.081 | 0.199 | REJECT |
| 4 | Loma sp. PAC together | 0.103 | 0.221 | 0.517 | 0.230 | 0.527 | do not reject |
| 5 | Loma sp. TOM together | 0.005 | 0.081 | 0.424 | 0.081 | 0.199 | REJECT |
| 6 | Loma spp. from gadids together | 0.720 | 0.668 | 0.906 | 0.668 | 0.918 | do not reject |
| 7 | L. morhua with Loma sp. PAC \& POL | 0.470 | 0.332 | 0.687 | 0.332 | 0.713 | do not reject |

Table 4.12: Summary of monophyly tests for all gene regions, and for rDNA without and with addition of a gap matrix ( $\mathrm{w} / \mathrm{o}=$ without gap matrix; + gap $=$ with gap matrix) for Loma species. Cases in which all isolates of a species formed a monophyletic clade in all trees are indicated with " Y ". Analyses in which isolates of a species did not form a monophyletic clade and where the forced monophyly was rejected (at 0.01 significance level) are indicated with " $R$ ". Cases in which isolates of a species did not always form a monophyletic clade but in which monophyly could not be rejected are indicated by "dnr". In some cases, indicated by "-", the region was not sequenced or only 1 isolate was available for a species. Chapters $2 \& 3$ name equivalents: Loma sp. $\mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \operatorname{LIN}$ $=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

| Species | rDNA |  |  |  |  |  | EF-1 $\alpha$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Region 1 |  | Region 2 |  | Region 3 |  |  |
|  | w/o | +gap | w/o | +gap | w/o | +gap |  |
| Undescribed |  |  |  |  |  |  |  |
| Loma sp. PAC | - | - | dnr | dnr | dnr | R | dnr |
| Loma sp. POL | Y | Y | dnr | dnr | dnr | R | R |
| Loma sp. TOM | Y | Y | dnr | dnr | dnr | R | dnr |
| Loma sp. LIN | Y | Y | dnr | dnr | dnr | dnr | - |
| Loma sp. BLK | - | - | dnr | dnr | dnr | dnr | - |
| Loma sp. BRO | Y | Y | Y | Y | Y | Y | - |
| Loma sp. AUS | - | - | - | - | - | - | - |
| Described |  |  |  |  |  |  |  |
| L. branchialis | - | - | dnr | dnr | Y | Y | - |
| L. morhua | - | - | dnr | dnr | dnr | R | R |
| L. salmonae | Y | Y | Y | Y | Y | Y | dnr |
| L. embiotocia | - | - | Y | Y | - | - | - |

Table 4.13: Microsporidian SSU rDNA, EF-1 $\alpha$ and RPB1 gene sequences from Genbank that were included in phylogenetic analyses showing Genbank accession numbers, current family designation, and host name. $\mathrm{F}=$ fish, $\mathrm{I}=$ insect, $\mathrm{C}=$ crustacean. *this species was a hyperparasite in the fish, found in an unidentified myxosporean in the gut epithelium of the fish. ${ }^{1}$ Pleistophora sp. PA has recently been transferred to Tuzetia weidneri (Lom \& Nilsen, 2003)

| Gene region and Species | Accession | Family | Host |
| :---: | :---: | :---: | :---: |
| SSU rDNA |  |  |  |
| Glugea americanus | AF056014 | Glugeidae | F: Lophius americanus |
| Glugea anomala Nil | AF044391 | Glugeidae | F: Gasterosteus aculeatus |
| Glugea anomala Pom | AF056016 | Glugeidae | F: Gasterosteus aculeatus |
| Glugea atherinae | U15987 | Glugeidae | F : Atherina boyeri |
| Glugea plecoglossi | AJ295326 | Glugeidae | F: Plecoglossus altivelis |
| Glugea sp. GS1 | AJ295325 | Glugeidae | F: Gasterosteus aculeatus |
| Glugea sp. Wu | AY090038 | Glugeidae | F: Epinephelus awoara |
| Glugea stephani | AF056015 | Glugeidae | F: Platichthys flesus |
| Heterosporis anguillarum | AF387331 | Pleistophoridae | F: Anguilla japonica |
| Ichthyosporidium sp. BAK | L39110 | Nosematidae | F: Leiostomus xanthurus |
| Loma acerinae Che | AJ252951 | Glugeidae | F: Gymnocephalus cernuиs |
| Loma acerinae Nil | AF356224 | Glugeidae | F: Gymnocephalus cernuus |
| Loma sp. Nil | AF104081 | Glugeidae | F: Enchelyopus cimbrius |
| Microgemma caulleryi | AY033054 | Unikaryonidae | F: Hyperoplus lanceolatus |
| Microsporidium cypselurus | AJ300706 | not placed | F: Cypselurus pinnatibarbatus japonicus |
| Microsporidium prosopium | AF151529 | not placed | F: Prosopium williamsoni |
| Microsporidium sp. exGamm | AF397404 | not placed | C: Gammarus duebeni duebeni |
| Pleistophora finisterrensis | AF044393 | Pleistophoridae | F: Micromesistius poutassou |
| Pleistophora mirandellae | AJ252954 | Pleistophoridae | F : Rutilus rutilus |
| Pleistophora sp. $\mathrm{PA}^{1}$ | AJ252958 | Pleistophoridae | C: Penaeus aztecus |
| Pleistophora typicalis | AJ252956 | Pleistophoridae | F: Myoxocephalus scorpius |
| Pseudoloma neurophilia | AF322654 | not placed | F: Danio rerio |
| Spraguea lophii | AF056013 | Spragueidae | F: Lophius piscatorius. |
| Vavraia oncoperae | X74112 | Pleistophoridae | I: Wiseana spp. |
| Unidentified microsporidium MYX1 | AJ295329 | not placed | F*:Takifugu ruripes |
| EF-1 $\alpha$ gene |  |  |  |
| Glugea plecoglossi | D84253 | Glugeidae | F: Plecoglossus altivelis |
| RPB1 gene |  |  |  |
| Cystosporogenes operophterae | AJ278949 | Glugeidae | I: Operophtera brumata |
| Glugea anomala | AJ278952 | Glugeidae | F: Gasterosteus aculeatus |
| Loma acerinae | AJ278951 | Glugeidae | F: Gymnocephalus cernuus |
| Nosema locustae | AF061288 | Nosematidae | I: Orthoptera spp. |
| Nosema tyriae | AJ278948 | Nosematidae | I: Tyria jacobaeae |
| Pleistophora hippoglossoideos | AJ278950 | Pleistophoridae | F: Hippoglossoideos platessoides |
| Pleistophora typicalis | AJ278946 | Pleistophoridae | F: Myoxocephalus scorpious |
| Pleistophora sp. LS | AJ278947 | Pleistophoridae | F: Litopenaeus setiferus |
| Saccharomyces cerevisiae | X03128 | (FUNGI) | n/a |
| Vairimorpha necatrix | AF060234 | Burenellidae | I: Lepidoptera spp. |
| Vavraia culicis | AJ278956 | Pleistophoridae | I: Aedes albopictus |

Figure 4.14: Maximum parsimony, distance and maximum likelihood trees for 810 bp of the small subunit ribosomal DNA (SSU rDNA) for all available species of Loma, Glugea and several members of other clades parasitic in fish and crustaceans, showing overall branch length differences, and variation in tree topology, particularly the placement of L. acerinae, depending on method used. Clades "Loma" and "Glugea" were designated based on the presence of the type species for these genera (L. morhua and G. anomala) in these clades. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=L$. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=L$. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=L$. kenti, Loma sp. LIN $=$ L. lingcodae, Loma sp. BLK $=L$. richardi, Loma sp. BRO $=L$. salmonae SV of Sánchez et al. (2001a).



Figure 4.15: Consensus $50 \%$ majority-rule maximum parsimony tree from 810 alignment positions of SSU rDNA for species of Loma, Glugea and others parasitic in fishes, crustaceans, or myxosporeans, showing branches obtained by all 3 analysis methods in bold (maximum parsimony MP, distance ME, and maximum likelihood ML). Bootstrap values $>50 \%$ (MP 1000 replicates, ME 1000 replicates, ML 100 replicates) shown on branches. Clades "Loma" and "Glugea" (black) contain the type species (L. morhua and G. anomala) of these genera. Clades numbered 1 to 4 (in gray) refer to groups in Lom \& Nilsen (2003). Loma $\operatorname{sp} . \mathrm{PAC}=L$. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma $\operatorname{sp} . \mathrm{BLK}=L$. richardi, Loma $\mathrm{sp} . \mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).


Figure 4.16: Maximum parsimony, distance and maximum likelihood trees for 783 alignment positions from the partial RPB1 DNA gene for available species, showing overall branch length differences, and variation in tree topology, particularly the variable placement of $L$. acerinae in relation to L. salmonae and G. anomala depending on method used.


Figure 4.17: Consensus $50 \%$ majority-rule tree (distance) generated from 783 alignment positions from the partial RPB1 DNA sequences for species available, showing in bold the branches also shared in maximum parsimony (MP) and maximum likelihood (ML) analyses, with distance (ME), MP and ML bootstrap values $>50 \%$ shown on branches (from 1000 replicates with heuristic search in MP and ME, and 100 replicates in ML). Clades in gray, numbers 1, 3 and 4 refer to groups (families or other designation) as defined in Lom and Nilsen (2003).

Figure 4.18: Tree topologies used for Loma genus boundaries tests from the partial SSU rDNA gene. Boxes show monophyletic groups, corresponding to various hypotheses created as constraints prior to heuristic tree searches. Likelihood scores from these topologies were calculated using a best-fit substitution model from Modeltest and parameters estimated from the data to statistically test best trees (see Table 4.14) using the CONSEL software package. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=$ L. kenti, Loma $\mathrm{sp} . \operatorname{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma sp. BRO = L. salmonae SV of Sánchez et al. (2001a).




Figure 4.19: Tree topologies used for Loma genus boundaries tests from the partial RPB1 gene. Boxes show monophyletic groups, corresponding to various hypotheses created as constraints prior to heuristic tree searches. Likelihood scores from these topologies were calculated using a best-fit substitution model from Modeltest and parameters estimated from the data to statistically test best trees (see Table 4.14) using the CONSEL software package.

Table 4.14: Comparison of unconstrained and monophyly-constrained trees from the partial SSU rDNA and RPB1 gene sequences for Loma species, using the Approximately Unbiased or "AU", weighted and unweighted Kishino-Hasegawa or "KH" and Shimodaira-Hasegawa or "SH" tests in the CONSEL software package. Tree numbers correspond to those in Figure 4.18 (SSU rDNA) and Figure 4.19 (RPB1).

| Tree <br> $\#$ | Constraint: | AU | KH | SH | WKH | WSH | Conclusion |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| partial SSU rDNA |  |  |  |  |  |  |  |
| 1 | unconstrained | 0.003 | 0.046 | 0.230 | 0.046 | 0.110 | REJECT |
| 2 | L. acerinae sister to Loma clade | $4 \mathrm{e}-032$ | 0 | 0 | 0 | 0 | REJECT |
| 3 | L. acerinae with L. branchialis | $5 \mathrm{e}-054$ | 0 | 0 | 0 | 0 | REJECT |
| 4 | Loma spp. from gadids together | 0.999 | 0.954 | 0.984 | 0.954 | 0.996 | do not reject |
| 5 | L. acerinae sister to Glugea clade | 0.002 | 0.023 | 0.089 | 0.023 | 0.047 | REJECT |
|  |  |  |  |  |  |  |  |
| partial RPB1 gene |  |  |  |  |  |  |  |
| 1 | unconstrained | 0.585 | 0.515 | 0.666 | 0.515 | 0.666 | do not reject |
| 2 | L. acerinae with L. salmonae | 0.525 | 0.485 | 0.661 | 0.485 | 0.662 | do not reject |
| 3 | L. salmonae with G. anomala | 0.298 | 0.327 | 0.482 | 0.327 | 0.490 | do not reject |



Figure 4.20: Phylogenetic trees of Loma species and their hosts, showing possible coevolution (cospeciation) among Loma species from gadids (thick black branches) and Scorpaeniformes (gray branches). Loma species tree was reconstructed from a single, reference sequence (1113 to 1846 bp long) for each species using the majority nucleotide state among all isolates, for the SSU, ITS and LSU rDNA regions (heuristic ML search in PAUP*, showing bootstrap values > $50 \%, 100$ replicates, on branches). Branch to Loma sp. AUS was significantly shortened for clarity. Dotted lines connect parasites with their hosts and show two possible cases of "switching". Timeline (top right) applies only to gadid part of host tree. Gadid relationships and timeline re-drawn from Carr et al.'s (1999) mtDNA-based phylogeny calibrated using Bermingham et al. (1997). Remaining host tree adapted from Nelson (1994), and is not to scale with gadid branches. Order Perciformes may be polyphyletic (Nelson, 1994) and not as shown. Chapters 2 \& 3 name equivalents: Loma $\mathrm{sp} . \mathrm{PAC}=$ L. pacificodae, Loma $\mathrm{sp} . \mathrm{POL}=$ L. wallae, Loma $\mathrm{sp} . \mathrm{TOM}=L$. kenti, Loma $\mathrm{sp} . \mathrm{LIN}=$ L. lingcodae, Loma $\mathrm{sp} . \mathrm{BLK}=$ L. richardi, Loma sp. $\mathrm{BRO}=$ L. salmonae SV of Sánchez et al. (2001a).

# Chapter 5: Phylogenetic distance of Thelohania butleri Johnston, Vernick \& Sprague, 1978 (Microsporidia; Thelohaniidae) from the smooth pink shrimp Pandalus jordani Rathbun, 1902 from other Thelohania species: implications for genus Thelohania Henneguy, 1892 and family Thelohaniidae Hazard \& Oldacre 1975. 

## INTRODUCTION

Thelohania butleri Johnston, Vernick \& Sprague, 1978 (syn. Thelohania sp. Vernick, Sprague \& Krause, 1977), like other microsporidia, is a tiny, single-celled, spore-forming, eukaryotic, intracellular parasite. Thelohania butleri causes cotton shrimp or "cooked shrimp" disease in the smooth pink shrimp Pandalus jordani Rathbun, 1902 (Decapoda: Pandalidae) in Pacific waters of British Columbia, Canada and on the west coast of North America to California (Vernick et al., 1977; Johnston et al., 1978; Butler, 1980; Olson \& Lannan, 1984). In infected shrimp, the abdominal musculature becomes largely replaced by vegetative dividing stages and then by infective spores, giving the shrimp an opaque, white appearance, and presumably leading to mortality in most cases (Olson \& Lannan, 1984). Ongoing surveys of British Columbian stocks of smooth pink shrimp have shown that T. butleri persist at about 1-2 \% prevalence (Olson \& Lannan, 1984; J. Boutillier, Pacific Biological Station, Nanaimo, Canada, personal communication), but has been as high as $5 \%$, at which point it caused economic losses due to the unsightly appearance of shrimp (Olson \& Lannan, 1984). Thelohania butleri is therefore of interest as a persistent low-level shrimp pathogen, and is also of interest more widely to those who study microsporidia as it is one of just a few species of

Thelohania from the host group Decapoda from which the first named Thelohania species was found.

The genus Thelohania Henneguy, 1892 in Henneguy \& Thélohan (1892) is one of the oldest and most problematic in the Microsporidia. Virtually every author writing a revision of microsporidian groups has noted that this species-rich genus has been too broadly defined (Hazard \& Oldacre, 1975; Sprague, 1977; Canning \& Lom, 1986; Sprague et al., 1992; Larsson, 1999). Gurley (1893) designated one of three Thelohania species in Henneguy \& Thélohan (1892), T. giardi Henneguy, 1892 in Henneguy \& Thélohan (1892) as the type species. Thelohania giardi was described from the common European shrimp Crangon crangon Linnaeus, 1758 (= Crangon vulgaris Fabricius, 1798) (Decapoda: Crangonidae) in Boulogne, France. Despite the fact that T. giardi has been examined in over 22 publications (see Hazard \& Oldacre, 1975), there has never been transmission electron microscopy (TEM) or DNA sequence study of $T$. giardi, so authors have had to rely largely on two of the most comprehensive and early light microscopy studies of T. giardi by Henneguy \& Thélohan (1892) and Mercier (1909) for definition of this genus (Hazard \& Oldacre, 1975; Sprague et al., 1992; Larsson, 1999; Moodie et al, 2003). Sprague (1977) defined the genus more inclusively, e.g. "sporogony by endogenous budding...," with " 8 sporoblasts produced within a subpersistant pansporoblastic membrane", while acknowledging that this definition is "fragmentary and tentative". This definition, which includes species producing sporogonial plasmodia, has since been widely used, whereas the original type description was more precise and limited. For example, Johnston et al. (1978) point out that the Henneguy \& Thélohan (1892) and Mercier (1909) gave explicit descriptions and illustrations of sporogony in T. giardi by a series of three binary fissions without production of a plasmodium (e.g. "pas de masse plasmique proprement dite" Henneguy \& Thélohan, 1892). Johnston et al.'s (1978) study of T. butleri, which they presumed to be a very close relative of $T$. giardi from another marine shrimp, confirmed using TEM these three binary fissions and numerous other developmental details found in the original and Mercier's (1909) descriptions.

Hazard \& Oldacre (1975) and Larsson (1999), in their reviews of both older descriptions of T. giardi and of a number of newer Thelohania species pointed out that many species have been placed in this genus which bear little resemblance to one another and perhaps are only very distantly related. Thus, genus Thelohania, like genus Nosema has acted somewhat like an unofficial "holding group" for species until further data can place them elsewhere, whereas there is an official holding group, genus Microsporidium Balbiani, 1882, which might be more
appropriate, given the fact that having eight spores per sporophorous envelope is such a common feature. Therefore, these authors suggest limiting the genus to those species from decapod crustaceans (Hazard \& Oldacre, 1975; Larsson, 1999). Ribosomal DNA data suggest microsporidians frequently co-evolve with their host clades (Baker et al., 1997), so this hostbased restriction of the genus might be reasonable. However, more recent rDNA studies suggest microsporidia from crustaceans are dispersed throughout the tree (Refardt et al., 2002) suggesting perhaps the co-evolution "rule" does not apply well to species in crustaceans.

Regardless of how the genus is defined, it has been suggested and shown with rDNA (Lom et al., 2001; Moodie et al., 2003) to be polyphyletic. This is not surprising considering the broad definition (above) and its diversity of hosts and geographic locations. For example, to date, more than 80 species have been placed in this genus, about six in vertebrates, about 20 from taxonomically diverse crustaceans, and most others from terrestrial insects. About 14 named Thelohania species occur in decapod crustaceans, with perhaps a handful of these in the Infraorder Caridea comprising families Crangonidae and Pandalidae, in which are placed hosts of $T$. giardi and $T$. butleri, respectively. Phylogenetic analyses of rDNA sequences previously available for Thelohania species have only been able to show that species from fire ants in North America and from freshwater crayfish in Australia do not belong together, a not so surprising conclusion (Moodie et al., 2003). However, these studies have not been able to indicate which species are "true" Thelohania members, as sequence from T. giardi, or even a presumably close relative from a marine shrimp, are not available. While T. butleri from British Columbia is distant geographically from the type locality of T. giardi (France), the host taxonomy and developmental morphology (detailed and compared in Johnston et al., 1978) appear to be more similar than many others in this genus. Therefore, while the present study of T. butleri may not be able answer the question "how to define and recognize true Thelohania species" it may help expand our understanding of the evolution of the polyphyletic "Thelohania-like" species, and help to understand the diversity of such microsporidians in marine shrimp.

The first goal of this study was to compare the microsporidian collected from smooth pink shrimp P. jordani on the west coast of Vancouver Island, the type host and locality of $T$. butleri, to data of Johnston et al. (1978), using light and electron microscopy, to confirm its identity. The second goal was to sequence partial small and large subunit and internal transcribed spacer rDNA (SSU, LSU, and ITS, respectively) to find the relationship of this species to other microsporidians. The third goal of this study was to test alternative relationships between Thelohania and other species according to various hypotheses, such as
monophyly of Thelohania species, or monophyly of species from similar host groups. Such hypothesis testing was facilitated using statistical bootstrap tests including the AU and related tests as described by Shimodaira \& Hasegawa (2001), and justified in related references (Kishino \& Hasegawa, 1989; Goldman et al., 2000; Shimodaira \& Hasegawa, 2001; Shimodaira, 2002) as well as in Chapter 4 of this thesis. The final goal of this study was to characterize the rDNA sequence of the "cotton shrimp" pathogen, to aid in the development of molecular tools for diagnosis or detection of low-level infection, and to help distinguish between populations of this commercial shrimp-fisheries pathogen.

## MATERIALS AND METHODS

## Specimen collection

Smooth pink shrimp Pandalus jordani Rathbun, 1902 (Decapoda: Panaeiidae) were collected during a Fisheries and Oceans Canada shrimp survey between May 1997 and May 2002. Shrimp were caught by bottom trawling off the south, west and north coasts of Vancouver Island, British Columbia from 3 to 48 km distance from shore, 65 to 150 m depth, and particularly from around Juan de Fuca Strait lat/long 48.15 N 124.00 W, Barkley sound lat/long 48.50 N 125.20 W and Queen Charlotte Sound lat/long 51.20 N 129.00 W. Several shrimp with signs of cotton shrimp disease (opaque, whitish musculature) were examined fresh for spores, using light microscopy. Identification of Thelohania-like infections was fairly easy, as most cases of cotton shrimp disease (78.7\%) are reportedly caused by T. butleri, (Olson \& Lannan, 1984), and the other three more rare microsporidians could be distinguished by having hundreds of spores per sporophorous vesicle (SV) as does P. crangoni, or no SVs, whereas Thelohania species possess a characteristic eight spores per SV. Potential confusion due to coinfections with combinations of these species has been shown to be unlikely (Olson \& Lannan, 1984). Infected shrimp were divided such that some tissue was fixed for DNA study (in $95 \%$ ethanol) and some was fixed for transmission electron microscopy (described below).

## Light and transmission electron microscopy (TEM)

Tissue for transmission electron microscopy was placed in $4 \%$ glutaraldehyde for 12 to 24 hours, soaked in either Millonig's solution or Sörensen's solution ( 0.1 M ) for 24 hours and post fixed in $1 \%$ osmium tetraoxide for 1 hour. Tissue was embedded in Spurr's resin. Ultra thin sections were lifted onto copper grids and stained with $2 \%$ aqueous uranyl acetate followed by Reynold's lead citrate. Ultrastructural features observed by transmission electron microscope (TEM) were photographed on a Zeiss 10C Transmission Electron Microscope and negatives were scanned at high resolution ( 1200 dpi ) and opened in Adobe Photoshop 6.0 for closer examination and measurement.

Spores from semi-thin resin-embedded sections and ethanol-fixed material in wet mount were photographed under light microscopy at 1000X magnification. Photographic negatives were scanned at high resolution ( 2000 dpi ), opened in Adobe Photoshop 6.0 and the zoom feature was used to enlarge spores to about $15-20 \mathrm{~cm}$ for measurement on screen. Fresh and frozen spores were measured at 1000 X magnification using and ocular micrometer that had been
calibrated using a slide micrometer to an estimated accuracy of $\pm 0.05 \mu \mathrm{~m}$ or by drawing with a drawing tube.

## DNA isolation

For DNA study, tissues were fixed in at least 10X volume of $95 \%$ ethanol. Ethanol was removed by soaking approximately 50 mg of ethanol-fixed tissue for 15 minutes in lysis buffer ( 10 mM Tris, 1 mM EDTA, $10 \mathrm{mM} \mathrm{NaCl}, 1 \% \mathrm{SDS}$ ). DNA isolation from this material required bead beating to break open the resistant spores. Bead beating was performed following the procedure of Docker et al. (1997a), by shaking tissue with 0.5 mm silica beads in a MiniBeadbeater (Biospec Products, Bartlesville, OK, USA) with TE buffer ( 10 mM Tris, 1 mM EDTA) for 1 to 3 min . Tissue was then digested in lysis buffer with $0.5 \mathrm{mg} / \mathrm{ml}$ proteinase K for $4-6$ hours at $37^{\circ} \mathrm{C}$ in a rotating incubator. PCR amplification from this DNA template required removal of inhibitors using CTAB (10\% w/v Cetyl Trimethyl Ammonium Bromide in 0.7 M NaCl ) and a protocol involving heating at $65^{\circ} \mathrm{C}$ in $1 / 8$ vol. $\mathrm{CTAB}, 1 / 6$ vol. of 5 M NaCl , for $20-$ 35 min, mixing, extracting in 1/2 vol. phenol:chloroform:isoamyl alcohol (25:24:1), and repeating about five times. DNA was then precipitated in 2.5 x vol. of cold $95 \%$ ethanol, and washed four times with $3 \times$ vol. $70 \%$ ethanol to remove salt. DNA was then vacuum dried, resuspended in $40 \mu \mathrm{l}$ distilled water and stored for use at $-20^{\circ} \mathrm{C}$.

## Polymerase chain reaction (PCR) for cloning and sequencing

PCR was performed in a Perkin Elmer Cetus DNA Thermal Cycler 480 in $25 \mu \mathrm{l}$ reactions with standard PCR buffer $2.5 \mathrm{mM} \mathrm{MgCl} 2,0.2 \mathrm{mM} \mathrm{dNTP}, 15 \mathrm{pmol}$ of each primer, and 1-3 units of Taq DNA polymerase (Gibco BRL, Gaithersburg, MD), with conditions of $95^{\circ} \mathrm{C}$ for $2 \mathrm{~min}, 3$ cycles of $95^{\circ} \mathrm{C}$ for $50 \mathrm{sec}, 50^{\circ} \mathrm{C}$ for $40 \mathrm{sec}, 72^{\circ} \mathrm{C}$ for 2 min , then 35 cycles of 95 ${ }^{\circ} \mathrm{C}$ for $50 \mathrm{sec}, 54^{\circ} \mathrm{C}$ for $30 \mathrm{sec}, 72^{\circ} \mathrm{C}$ for 90 sec , and a final extension of $72^{\circ} \mathrm{C}$ for 5 min . Primers amplifying the small subunit (SSU), internal transcribed spacer (ITS) and large subunit (LSU) ribosomal DNA (rDNA) were: forward M5P - CAC CAG GTT GAT TCT GCC pos. 118 at 5' end of the SSU (18eMIC in Docker, Kent et al. 1997); Seqlf - CGT TGT AGT TCT AGC AGT pos. 719-736 in the SSU (provided by M. F. Docker); and reverse SeqR - AAC AGG GAC KYA TTC ATC pos. 1218-1235 in the SSU (this study); 580R GGT CCG TGT TTC AAG ACG G pos. 1847-1865 in the LSU (Vossbrinck et al. 1987). PCR products were visualized in ethidium bromide stained 1.5-2 \% agarose gels. Products were either prepared for cloning as described below or excised from gels and freeze-thaw extracted from agarose to be sequenced directly.

## Cloning

PCR products were run in $0.8 \%$ agarose to isolate the desired product, and then slices were cleaned for ligation using Ultraclean 15 MOBIO DNA Purification Kit (BIO/CAN Scientific Inc. Mississauga, ON). Products were ligated and transformed using the TOPO TA Cloning PCR Version 2.1 kit (Invitrogen Corp., Carlsbad, CA) and cells were grown on standard $\mathrm{LB}+$ ampicillin plates, and screened for the presence of the insert in $10 \mu \mathrm{PCR}$ reactions using Taq DNA Polymerase (Invitrogen Corp., Carlsbad, CA), standard reagents with screening primers M13-20 and M13 Rev ( $94^{\circ} \mathrm{C}$ for $2 \mathrm{~min}, 34$ cycles of $92{ }^{\circ} \mathrm{C}$ for $45 \mathrm{sec}, 55^{\circ} \mathrm{C}$ for 45 sec $72^{\circ} \mathrm{C}$ for 1 min 30 sec , followed by $72^{\circ} \mathrm{C}$ for 5 min ). Clones were re-grown from master plates in 3 ml of standard LB culture with 50 mM amp , shaking overnight at 220 rpm at $37^{\circ} \mathrm{C}$. Cells were pelleted at 1300 rpm 30 sec , rinsed and plasmids were isolated from cells using the Rapid Plasmid Miniprep System (Gibco BRL, Gaithersburg, MD).

## DNA sequencing

Sequencing was performed on the ABI PRISM 377 DNA automated sequencer using BigDye Terminator Version 3.1, fluorescent dye-labelled terminators with forward and reverse primers and PCR conditions as recommended by the manufacturer. Multiple PCR products and multiple clones were sequenced in both directions to check for Taq or sequencer errors.

## Flip algorithm

Sequencing directly from PCR products, without cloning, sometimes produced sequences that were double, having two different nucleotide signals, at almost every site after a stretch of normal single nucleotide signals. The correct signals from the doubled portion of sequences were extracted using Flip Analyzer of REALEM Version 1.01 (developed this study). For more details, see Appendix 1.

## Polymorphic sites

Polymorphisms (two or more alleles in the population) produced polymorphic sites, in which two or more nucleotide signals appeared at single site in a sequence, when PCR products were sequenced directly. Polymorphic sites are distinguished from double sequences (see Flip algorithm above) by being rare, rather than occurring at the majority of sites. Because of uncertainty as to whether these polymorphic sites were errors resulting from background contamination or sequencing artefacts, or whether they reflect valid polymorphism in the data resulting from substitutional differences between two copies of the gene, the original sequence data were carefully re-examined for background signal or sequencer software error. In cases
where there appeared to be uncertainty, sequencing was repeated. Where there was valid polymorphism in the data, the standard degenerate nucleotide code was used. Whereas a full substitutional difference between two separately sequenced products (e.g. A versus $T$ ) is reported as a transversion or transition, a partial substitutional difference between two sequences (e.g. A versus $\mathrm{A} / \mathrm{T}$ polymorphism) will hereafter be referred to as a "polymorphic difference" and will be counted as half the value of a full substitution in calculating percent sequence differences.

Percent sequence difference was calculated by adding the number of full substitutions, the number of indels (regardless of size) and half the number of polymorphic differences, divided by the total number of alignment positions for which both sequences are known.


#### Abstract

Alignment Sequences were aligned with a large set of rDNA sequences from microsporidia, fungi, and other outgroups from Genbank by first grouping them into clusters of taxa known to be related from previous phylogenetic studies, then by aligning more conserved regions by eye using ESEE 3.2s (Eyeball SEquence Editor, Eric Cabot, 1998). Less conserved regions of the rDNA were aligned with emphasis on conservation of secondary structure using models from other microsporidians in the rDNA structural database (Van de Peer et al., 2000) and also using Clustal W (version 1.74) (Thompson et al., 1994) under a variety of parameters emphasizing transitions over transversions for more closely related taxa or allowing frequent gaps near rDNA loop regions. A preliminary phylogenetic analysis was performed to help choose taxa that could be removed to speed analyses. Taxa chosen for phylogenetic analysis (Table 5.1) included several representatives from clades distant to clades containing Thelohania species, any taxa similar to T. butleri, all Thelohania species and any taxa found to be similar to these in Blast (NCBI) searches.


## Phylogenetic analysis

Phylogenetic analyses were performed using maximum parsimony (MP), distance (ME) and maximum likelihood (ML) in PAUP*4.0b10 (Swofford, 2001), with heuristic searches, random stepwise sequence addition and TBR branch swapping with 10 repetitions, or only one repetition for ML. Maximum parsimony (unweighted) was followed by 1000 replicates of bootstrap resampling (heuristic search), and the results summarized as $50 \%$ majority rule trees. A distance matrix was calculated using the logDet/paralinear model with pairwise distances calculated using the minimum evolution (ME) objective function. Distance bootstrap with
heuristic search was replicated 1000 times and reported as $50 \%$ majority rule consensus trees. Initial trees for maximum likelihood were generated using the HKY85+G model with transition/transversion ratio estimated from the data, a substitution rate matrix estimated using a four category discrete gamma approximation for among-site rate variation, and base frequencies calculated empirically. Further trees for maximum likelihood were generated using the best model predicted for this data using Modeltest Version 3.06 (Posada \& Crandall, 1998). (For further details, see Appendices 2 to 4). Maximum likelihood heuristic searches were run using the estimated parameters, and were followed by 100 replicates of bootstrap resampling (faststep search), and results reported as $50 \%$ majority-rule consensus trees.

Various different input data sets were created using the character exclusion feature of PAUP* to analyze the effects on tree topology and bootstrap support of adding or removing from the alignment: regions with uncertain alignment, regions with missing and ambiguous characters, or positions with gaps. Several additional input data sets were created using Gap Matrix of REALEM Version 1.01 (developed this study) to analyze the effects of adding a matrix containing gap information (location of gap ends), as explained in Appendix 1.

## Monophyly constraints and approximately unbiased (AU) tests

Monophyletic clades including species of the same host group, genus or family designation were created by opening trees from TreeView Version 1.6.6 (Page, 1996) as text files, removing tree lengths, and moving taxa into appropriate monophyletic clades. The resulting hypothetical trees were pasted into PAUP* and the monophyly constraint feature of the program was applied to create monophyly-constrained trees then produce log likelihood branch length scores (lscores) for both unconstrained and monophyly-constrained trees. Lscores were pasted into CONSEL Version 0.1f (Shimodaira \& Hasegawa, 2001) to compare tree topologies using the Approximately Unbiased test (AU test), and other similar tests. For further details and justification for use of this test see Appendices 5 through 9. Output AU test probabilities less than 0.01 suggest evidence sufficient to reject the null hypothesis of trees being equally likely. Trees included in these comparisons were always inspected to be sure they had equal numbers of nodes (and therefore branches).

## RESULTS

## Morphological features

Abdominal skeletal muscles of infected smooth pink shrimp Pandalus jordani Rathbun, 1902 were largely replaced by spores, densely packed and always in groups of eight spores per sporophorous vesicle (SV) (Fig. 5.1). Sporophorous vesicles were easily visible at 400X magnification in fresh and preserved material, and were resistant to mechanical damage during examination.

Spores were ovoid and slightly narrower at the anterior end with a round posterior vacuole. Spores fixed in ethanol were $4.5 \pm 0.22 \mu \mathrm{~m}$ (range $3.95-5.56 \mu \mathrm{~m}$ ) long x $3.2 \pm 0.15$ $\mu \mathrm{m}$ (range 2.92-3.95 $\mu \mathrm{m}$ ) wide $(\mathrm{n}=15)$. Spores fixed in glutaraldehyde and embedded in resin were $3.52 \pm 0.09 \mu \mathrm{~m}$ (range 3.25-3.82 $\mu \mathrm{m}$ ) long $\times 2.25 \pm 0.08 \mu \mathrm{~m}$ (range 1.93-2.50 $\mu \mathrm{m}$ ) wide $(n=14)$.

Under transmission electron microscopy (TEM), uninucleate stages about $4 \mu \mathrm{~m}$ wide (presumably sporonts) were observed within sporophorous vesicles in skeletal muscle cells (Figs. 5.2 - 5.5). In these stages the parasite cytoplasm was highly vacuolated and contained up to three conspicuous, dark, round, amorphous granules about $0.5 \mu \mathrm{~m}$ wide (Figs. 5.2, 5.3). Slightly larger (4.5-8 $\mu \mathrm{m}$ ), highly vacuolated diplokaryotic sporogonial stages were observed in closely pressed pairs inside sporophorous vesicles filled with finely fibrous material (Fig. 5.2). These stages contained one to three dark granules $\sim 0.5 \mu \mathrm{~m}$ wide in the cytoplasm. Sporogonial plasmodia were never seen. Golgi-like vesicles were sometimes observed in these stages (Fig. 5.2). Uninucleate sporonts were observed in groups of four (presumably in the 4-cell stage of sporogony) cells closely pressed together (Figs. 5.2, 5.3, 5.6). These 4-cell stages were always associated with dark, amorphous material in the episporontal space in the region between and partly surrounding the cells (Figs. 5.2 - 5.4, 5.6). Frequently there were dark vesicular inclusions about $1 \mu \mathrm{~m}$ wide, sometimes with a perforated appearance (Figs. 5.5,5.6) within the episporontal space. Host skeletal muscle cells adjacent to these stages contained lysosome-like vesicles. Early sporoblasts without signs of endospore formation were smaller (4-5 $\mu \mathrm{m}$ ), less vacuolated, contained polar filaments in a fairly advanced stage of development, and the finely granular material in the episporontal space was partly changed to tubule-like dense granules. Mature spores had crescent-shaped, single, unpaired nuclei and a centrally located polar capsule (Figs. 5.6-5.8), and typical lamellar and vesicular polaroplasts. Exospore had fine undulations
or striations on the surface. Polar filaments were isofilar with always 13 turns in two rows, of five and six turns, with the last two turns in a single row $(\mathrm{n}=6$ ) (Figs. 5.7, 5.8).

## Ribosomal DNA sequence characteristics

In total, 1625 nucleotide positions were sequenced, 1114 from the small subunit (SSU), 35 from the internal transcribed spacer (ITS), and 476 from the large subunit (LSU) rDNA. A low level of intraspecific difference ( $0.24 \%$ ) was observed in four very similar sequences obtained from separate clones or directly sequenced PCR products. Most intraspecific nucleotide differences were found in the ITS and LSU regions. PCR products that were sequenced directly had a 1 bp indel difference (see Flip algorithm in Materials and Methods) at position 1482 in the LSU region. Two clones shared just one difference. Because intraspecific difference was low, a single representative sequence from cloned PCR products (fused clones) will be used in phylogenetic analyses, and referred to hereafter as T. butleri. Percent GC content in the SSU region of this sequence was $46.4 \%$. This T. butleri sequence was most similar to an undescribed microsporidian, Microsporidium JES2002G (Genbank AJ438962) from the sand shrimp Gammarus chevreuxi Sexton, 1913 (Amphipoda) in River Avon, United Kingdom, having $13.87 \%$ difference in the SSU region.

## Phylogenetic relationships

Thelohania butleri fell within a large clade that included all fish-parasitic microsporidians (except Nucleospora salmonis) and many microsporidians from crustaceans, in all analyses of SSU rDNA using parsimony (MP), distance (ME), and maximum likelihood (ML) optimality criteria (best-fit model chosen from Modeltest was GTR $+\mathrm{I}+\mathrm{G}$ ) (Figs. 5.9 5.12). This fish/crustacean-parasite clade will hereafter be referred to as the " $\mathrm{F} / \mathrm{C}$ " clade. In all analyses, other Thelohania species fell in two separate, well-supported clades, one including Thelohania species from crayfish (clade T), and the other with Thelohania species from ants (clade V, after Visvesvaria species). Clades T and V always fell outside the F/C clade.

Analyses using more or less conservative characters in the input data sets (Table 5.2) always produced the $\mathrm{F} / \mathrm{C}$ clade ( 70 to $90 \%$ bootstrap support) and always placed $T$. butleri in this clade. The input data set having presumably the most conserved characters with uncertain alignment regions, missing or ambiguous data, and positions with gaps removed, consisting of 524 alignment positions (Table 5.2), was used for all further trees shown in this study.

One species, Ameson michaelis, had an exceptionally long branch length (Fig. 5.9) and fell inside the F/C clade in MP and ML analyses, but outside this clade in ME trees. This
unstable placement of $A$. michaelis in ME trees (inside and outside $\mathrm{F} / \mathrm{C}$ ) is associated with a lower bootstrap support for the F/C clade in ME analyses (Fig.10).

The placement of $T$. butleri among sub-clades within the large $\mathrm{F} / \mathrm{C}$ group was variable, and such relationships had low support ( $<50 \%$ bootstrap value, asterisks in Figs. 5.10-5.12). However, T. butleri and its closest sister-species, Microsporidium sp. JES2002G, often appeared (ME and ML analyses, Fig. 5.9) as a sister clade to microsporidian parasites of marine shrimp (Perezia nelsoni and Unidentified microsporidium S1), and less often appeared completely buried within fish-parasitic microsporidian clades (MP analyses), or sometimes appeared as a sister to a clade containing microsporidians from brackish-water amphipods (Microsporidium sp. JES2002B \& C, and Microsporidium sp. ex Gammarus) (Figs. 5.10, 5.12).

## Monophyly constraints and AU-tests

Because bootstrap support for deeper branches in the tree were low, the placement of $T$. butleri in various alternate clades was tested by using AU, SH, KH, and weighted SH and KH tests. These tests were performed on the ML log likelihood branch scores (using the GTR $+\mathrm{I}+\mathrm{G}$ substitution model) for the original ML tree versus multiple alternative hypothetical trees, listed in Table 5.3 and illustrated in Fig. 5.13. Test results shown in Table 5.3 strongly suggest rejection of all alternate placements for T. butleri except as a sister to Ameson michaelis, or in SH and WSH tests only, placement with species from decapod crustaceans or Thelohania species from crayfish were accepted (SH) or weakly accepted (WSH).

## DISCUSSION

## Morphological identity with T. butleri Johnston, Vernick \& Sprague, 1978

Molecular studies must ensure specimens are identified accurately to avoid serious confusion later from mismatched records. This microsporidian was identical to T. butleri Johnston, Vernick and Sprague, 1978 in gross appearance, prevalence, and most morphological details, and occurred in the type host (smooth pink shrimp P. jordani), near the type locality (Queen Charlotte Sound) of T. butleri. Spore shape and dimensions matched those of T. butleri, these spores within the range ( $4.5-5.0 \times 3.0-3.5 \mu \mathrm{~m}$ ) reported by Johnston et al. (1978) when estimated shrinkage due to fixation (about 1.2 to 1.5 x calculated in Chapter 3 of this thesis) was taken into account. The development (sporogony) and ultrastructure of stages in the host tissues were exactly identical to those described for $T$. butleri, including the presence of cytoplasmic granules in early stages, a series of three binary fissions, with the early production of polar filaments in sporoblasts, and always containing eight spores per subpersistant sporophorous vesicle (equivalent to the "pansporoblastic membrane" in that paper). The only difference observed from T. butleri was that these spores had polar filaments with exactly 13 turns in two rows, with an outer row of six turns and inner row of five turns, with two more turns in a single row at the posterior end, whereas Johnston et al. (1978) reported "about 10 turns" in two rows. However, they show a clear photograph of a spore with 10 turns (five double coils) in the plane of section which is obviously beyond the mid-axis such that even the posterior vacuole cannot be seen, suggesting that spores (at least the one shown in Fig. 16 of Johnston et al., 1978) look identical to those observed here, and have a few more turns than 10 , when in median longitudinal sections. Spores from off-median sections in the present material also appeared to have fewer than 13 polar filament turns. For this reason, these specimens were considered conspecific with $T$. butleri. Further sectioning or examination of more spores in the deposited specimens of $T$. butleri held in the International Protozoan Type Slide Collection, Smithsonian Institution, Washington, D. C., and elsewhere (see Johnston et al., 1978), should confirm the presence of more polar filament turns.

## Phylogenetic placement of $\boldsymbol{T}$. butleri

Phylogenetic results were consistent with numerous studies of microsporidian rDNA, placing all fish-parasitic microsporidians (except for N. salmonis) together consistently (Nilsen et al., 1998, Nilsen, 2000, Bell et al., 2001, Lom \& Nilsen, 2003) with various species from
small crustaceans interspersed in this clade (clade F/C) (Lom \& Nilsen, 2003, Moodie et al., 2003). The consistent placement of $T$. butleri in this clade was unexpected for taxonomic reasons (discussed below), but not so surprising given the newly recognized dispersal of crustacean-parasitic species in the tree (Refardt et al., 2002). While most microsporidians have co-evolved within major host groups (e.g. homothermous or poikilothermous vertebrates, protists, insects) (Baker et al., 1995; Baker et al., 1997; Weiss \& Vossbrinck, 1999), crustaceanparasitism in the microsporidia has followed a different pattern. Crustacean hosts have more diverse microsporidia than any other group (see Larsson, 1996; 1999; Refardt et al., 2002; Chapter 6). The crustacean-parasitism pattern cannot be confidently interpreted as being due to an older ancestry of crustacean-parasitism in the microsporidia versus greater opportunity or selection for host switching to or from crustacean hosts. Part of the difficulty in understanding this pattern is that SSU rDNA data do not resolve the deeper branches in the tree with high confidence. This lack of resolution (Figs. 5.10-5.12) may be due to the unequal rates of evolution among groups, as suggested by Nilsen \& Chen (2001) and Lom \& Nilsen (2003) or the poor representation of many intermediate species in the data. Further data, particularly from independent loci, like RNA polymerase largest subunit II (RPB1), or other genes may help resolve the phylogeny (Cheney et al., 2001).

Phylogenetic results (Figs. 5.9-5.12) were also consistent with past studies in placing Thelohania species from ants in a basal place relative to all species from fish-parasitic groups and as a sister-group to other species from insects (Nilsen \& Chen, 2001; Lom \& Nilsen, 2003; Moodie et al., 2003), as would be expected if the general host group is an important highertaxonomic character. Similarly, this study showed Thelohania species from crayfish grouped separately from Thelohania species from ants and as a sister-group to the largely fish-parasitic group (clade F/C) (Fig. 5.9), as was found by Moodie et al. (2003) who suggested host habitat (aquatic vs. terrestrial) is often a significant indicator of relatedness. Moreover, these results combined with the placement of T. butleri separately from other Thelohania species (Figs. 5.9 to 5.12), suggests characters used for genus Thelohania are adaptive and are not appropriate for higher-level taxonomy, whereas general host group is a better higher-level taxonomic character, as was suggested by Baker et al. (1995) and Baker et al. (1997). These results conform with the growing number of studies suggesting species placed in genera based on older morphological diagnostic features need to be transferred to appropriate groups such that taxonomy reflects phylogeny (Nilsen \& Chen, 2001; Bell et al., 2001; Lom \& Nilsen, 2003; Moodie et al., 2003; and arguments in Quicke, 1993).

It was unexpected that $T$. butleri would group consistently and with high ( $>75 \%$ ) bootstrap support with a microsporidian (Microsporidium sp. JES2002G Genbank AJ438962) from a brackish-water sand shrimp G. chevreuxi in River Avon, UK, given the distant geographic locality and diverged taxonomic relationships of the hosts. Under a strict coevolution hypothesis Microsporidium sp. JES2002G should have grouped with other species from Gammarus hosts, while T. butleri should have grouped with species in decapod crustaceans; however, these results suggested species from crustaceans do not strictly co-evolve. These species differed at $13.87 \%$ of the sites in SSU rDNA, suggesting they may not be close sister species and may not even belong in the same genus. Molecular analyses suggested they were at least related at the family or higher level; however, morphological data has not yet been published for Microsporidium sp. JES2002G (Dunn et al., 2001).

Inferences about the possible morphology of this important sister-species of T. butleri can be made from records of microsporidia from its host, G. chevreuxi. Microsporidium sp . JES2002G could be conspecific with a species Thelohania muelleri (Pfeiffer, 1894) Stempell, 1902. Gammarus chevreuxi was reported to carry T. muelleri, according to Goodrich (1928) in Sprague (1977), whereas others reported T. muelleri primarily infects other European Gammarus species (Friedrich et al., 1995; Ovcharenko, 1995; and see Sprague, 1977). Some features of $T$. muelleri seem similar to those of $T$. butleri; for example, it undergoes a series of three binary divisions during sporulation. However, such features of T. muelleri were compiled from references from other Gammarus hosts (Sprague, 1977), and morphological diversity in $T$. muelleri from different hosts was high (Friedrich et al., 1995). Consequently, Microsporidium sp. JES2002G from G. chevreuxi may possess T. muelleri-like features, but some T. muelleri populations might not possess T. butleri-like features. A recent molecular and TEM study of $T$. muelleri from G. duebeni celticus (a freshwater amphipod) by Terry et al. (2003) showed this isolate of T. muelleri was not a Thelohania species at all, but fell within the genus Pleistophora, so the authors renamed this species $P$. muelleri. The morphological features of that isolate differ substantially from those summarized in Sprague (1977); therefore, it remains likely that $T$. muelleri-like species from Gammarus spp. may be a polyphyletic cluster of distinct species, in which at least one species from G. chevreuxi falls together with microsporidians from marine decapod crustaceans.

Thelohania butleri and the undescribed species from G. chevreuxi were sometimes clustered with Perezia nelsoni (Sprague, 1950) Vivarès and Sprague, 1979 and Unidentified microsporidium S1 from marine shrimp of the suborder Penaeoida, family Peneidae
(Litopenaeus setiferus and Metapenaeus joineri), particularly in maximum likelihood (ML) analyses. Although this placement in ML analyses only received $35 \%$ bootstrap support, it might be a significantly better estimate of the phylogeny than that produced by parsimony (MP) or distance (ME). The significant branch length differences (especially for Ameson michaelis and Nosema bombycis in Fig. 5.9) might have caused long branch attraction that would be particularly likely to bias the MP results (Swofford, 2001, and references therein). The base frequency inequality detected by Modeltest was statistically significant and under these conditions ML was likely to estimate correct relationships more frequently than ME. ML results suggested that T. butleri and "Microsporidium sp. JES2002G" were sister to a third microsporidian from a marine decapod crustacean, A. michaelis (Sprague, 1970) Sprague, 1977, in the blue crab Callinectes sapidus (Fig. 5.12; Fig. 5.13 tree \#1). Placement of $T$. butleri plus "Microsporidium sp. JES2002G" and A. michaelis together could not be rejected by AU and other bootstrap resampling tests (Fig. 5.13 tree \#1; Table 5.3). Its proximity to Ameson (family Pereziidae) suggests $T$. butleri may be closely associated with, if not within family Pereziidae Loubès, Maurand, Comps \& Campillo, 1977. However, the family Pereziidae is defined as having "interfacial envelope absent, sporogonial plasmodium moniliform" (in Sprague et al., 1992). This definition clearly does not fit T. butleri. Therefore, for the present, T. butleri must remain in the polyphyletic genus Thelohania and family Thelohaniidae, until the phylogenetic position of the type species T. giardi is known.

## Implications for genus Thelohania Henneguy, 1892

The placement of $T$. butleri with other species from marine decapod crustaceans strengthens a pre-existing argument by Hazard \& Oldacre (1975) and Johnston et al. (1978) that species should only be classified in Thelohania if they occur in marine decapods like type species T. giardi. However, previous observations (Sprague et al., 1992; Moodie et al., 2003) and the present study suggest, that the host is not a sufficiently narrow diagnostic character. For example, T. butleri and other Thelohania species are sufficiently different in morphology and phylogenetic placement from species in marine decapods, A. michaelis and Perezia spp., that there is no evidence to transfer the latter species from their present genera. Furthermore, AU tests showed no support for either a "decapod host" clade, or a "crayfish/pink shrimp" clade, or any other crustacean-host based clade that included $T$. butleri, confirming that host taxonomy is an imperfect predictor of relationships.

Hence, developmental morphology and host are still problematic diagnostic characters for Thelohania, at least for species discussed in this study, but what about habitat? For some microsporidians, host habitat may be more important than host species, such that parasites may switch hosts more readily than they switch host habitat (e.g. marine vs. freshwater) over evolutionary time (Vossbrinck et al., 2004). Habitat seemed, at first, to be surprisingly important for nominal Thelohania species, given that all three species from crayfish (freshwater decapods) T. contejeani Henneguy, 1892, T. parastaci, Moodie, 2003, and T. montirivulorum Moodie, Le Jambre \& Katz, 2003 (Moodie et al., 2003) fell together in a separate clade, far from T. butleri and others in marine decapod hosts (see Fig. 5.12). However, habitats (and hosts) also differed for closely related species $T$. butleri (from a marine decapod) and "Microsporidium sp . JES2002G" (from a brackish water amphipod), suggesting habitat is also a poor character in Thelohania.

Despite the lack knowledge of T. giardi's place in the tree, it seems from all data available, including Henneguy \& Thélohan (1892), Mercier (1909), Johnston et al. (1978), Lom et al. (2001), Moodie et al. (2003) and this study, that T. giardi more closely resembles T. butleri than T. solenopsae or Thelohania spp. in crayfish. For now, it is hard to justify maintenance of the name Thelohania for species as distantly related as T. butleri and T. solenopsae, or T. butleri and T. contejeani. Before any of these species is transferred to a new genus, it would help to reexamine T. giardi carefully to refine the genus diagnostic characters. In particular, future researchers of T. giardi will have to look for the presence of T. butleri's developmental pattern, characterized by three binary fissions into 2-, 4- and 8-cell stages without production of a plasmodium, versus a pattern found in other Thelohania species, characterized by a sporogonial plasmodium that undergoes budding or "rosette" division. Study of T. giardi is needed to confirm the assertion by Johnston et al. (1978) that "true" Thelohania species (and T. giardi) must have the former type of sporogony. Molecular placement of T. giardi among other Thelohania species will, of course, be the most critical data to have in future.

## Implications for family Thelohaniidae Hazard \& Oldacre, 1975

Hazard \& Oldacre's (1975) family Thelohaniidae includes some 18 genera (Sprague et al., 1992), which are not close relatives, as shown by numerous molecular studies (Baker et al., 1997; Baker et al., 1998; Franzen \& Müller, 1999; Weiss \& Vossbrinck, 1999; Nilsen \& Chen, 2001; Andreadis \& Vossbrinck, 2002; Lom \& Nilsen, 2003; Moodie et al., 2003), and AU tests in this study. Revisions by Sprague (1977) and Sprague et al. (1992) had the effect of making
the family more broad, as both of these omit words in the first sentence of Hazard \& Oldacre's (1975) proposed family: "...species having sporonts that, after three sporogonic divisions, give rise to eight uninucleate microspores...". Results from this study bring to light the potential importance of the form of sporogonic divisions in the relationships of Thelohania-like species, such that it may be useful to look for the presence of three divisions, a primary feature of Hazard \& Oldacre's (1975) original group. If T. giardi is re-examined and data confirm that this species has three binary fissions without production of a plasmodial stage, as Henneguy \& Thélohan (1892) and Mercier (1909) described, this character might be "re-instated" as a feature of a new family revision, along with the molecular, host, and morphological relationships of species discussed above.

## FURTHER INVESTIGATION

Morphological or genetic variation across the geographic range of T. butleri, from California to British Columbia, is not known (Vernick et al., 1977; Johnston et al., 1978; Olson \& Lannan, 1984) and was not examined in this study. So, questions about species boundaries or population structuring of the parasite cannot be clarified at this time. Towards that end, this study provides a starting place for the development of rDNA probes that could be highly sensitive and specific for T. butleri (for example, like those developed for $L$. salmonae by Docker et al., 1997a). The results here suggest the ITS and LSU regions might have some intraspecific variation that could be further examined and used to characterize population differences of the parasite. These rDNA sequences could also be used to create PCR or in situ probes (see Brown \& Kent, 2002) to detect cryptic infections or elucidate the life cycle or early pathogenesis in the host (Vossbrinck et al., 1998; Brown \& Kent, 2002). Future studies on nomenclature for this group hinge upon the relationships of T. giardi, and so effort should be made to obtain data from this species. If general host group were important at higher-levels (as suggested by Hazard \& Oldacre, 1975; Baker et al., 1995; Weiss \& Vossbrinck, 1999; Moodie et al., 2003), we would predict that $T$. butleri would be a closer relative of the type species, $T$. giardi, than are Thelohania spp. from ants or crayfish, because of similarities in host.

## ACKNOWLEDGEMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada strategic grant 582073 to M. L. Adamson. I thank Jim Boutillier at the Pacific Biological Station for assistance in collection, assistance with initial literature, and for first suggesting the idea of looking at these "cottage-cheese" shrimp during a separate study of microsporidians in fishes. I am grateful to Elaine Humphrey and technicians at the BioImaging Facility, UBC, for their help with TEM and Michael Coury for his help designing software. I thank James Becnel for helpful criticism of this research when preliminary results were presented in Brazil.

## LITERATURE CITED

Andreadis, T. G. and Vossbrinck, C. R. 2002. Life cycle, ultrastructure and molecular phylogeny of Hyalinocysta chapmani (Microsporidia: Thelohaniidae), a parasite of Culiseta melanura (Diptera: Culicidae) and Orthocyclops modestus (Copepoda: Cyclopidae). Journal of Eukaryotic Microbiology 49(4):350-64.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Andreadis, T. G. 1998. Phylogeny of Amblyospora (Microsporida: Amblyosporidae) and related genera based on small subunit ribosomal DNA data: a possible example of host parasite cospeciation. Journal of Invertebrate Pathology 71:199-206.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Maddox, J. V. 1997. Phylogenetic position of Amblyospora Hazard \& Oldacre (Microspora: Amblyosporidae) based on small subunit rRNA data and its implication for the evolution of the microsporidia. Journal of Eukaryotic Microbiology 44(3):220-225.

Baker, M. D., Vossbrinck, C. R., Didier, E. S., Maddox, J. V. and Shadduck, J. A. 1995. Small subunit ribosomal DNA phylogeny of various microsporidia with emphasis on AIDS related forms. Journal of Eukaryotic Microbiology 42(5):564-570.

Bell, A. S., Aoki, T. and Yokoyama, H. 2001. Phylogenetic relationships among microsporidia based on rDNA sequence data, with particular reference to fish-infecting Microsporidium Balbiani 1884 species. Journal of Eukaryotic Microbiology 48(3):258-65.

Brown, A. M. V. and Kent, M. L. 2002. Molecular diagnostics for Loma salmonae and Nucleospora salmonis (microsporidia) In Molecular diagnostics of salmonid diseases. Cunningham, C. O. (ed.). Kluwer Academic Publishers, Dordrecht p. 267-283.

Butler, T. H. 1980. Shrimps of the Pacific coast of Canada. In Canadian Bulletin of Fisheries and Aquatic Sciences 202.280 pp .

Canning, E. U. and Lom, J. 1986. The microsporidia of vertebrates. Academic Press, London. 289 pp.
Cheney, S. A., Lafranchi-Tristem, N. J., Bourges, D. and Canning, E. U. 2001. Relationships of microsporidian genera, with emphasis on the polysporous genera, revealed by sequences of the largest subunit of RNA polymerase II (RPB1). Journal of Eukaryotic Microbiology 48(1):111-117.

Docker, M. F., Devlin, R. H., Richard, J., Khattra, J. and Kent, M. L. 1997a. Sensitive and specific polymerase chain reaction assay for detection of Loma salmonae (Microsporea). Diseases of Aquatic Organisms 29(1):41-48.

Docker, M. F., Kent, M. L., Hervio, D. M. L., Khattra, J. S., Weiss, L. M., Cali, A. and Devlin, R. H. 1997b. Ribosomal DNA sequence of Nucleospora salmonis Hedrick, Groff and Baxa, 1991 (Microsporea: Enterocytozoonidae): implications for phylogeny and nomenclature. Journal of Eukaryotic Microbiology 44(1):55-60.

Franzen, C. and Müller, A. 1999. Molecular techniques for detection, species differentiation, and phylogenetic analysis of microsporidia. Clinical Microbiology Reviews 12(2):243-285.

Friedrich, C., Kepka, O. and Lorenz, E. 1995. Distribution, host spectrum and ultrastructure of Thelohania muelleri Pfeiffer, 1895. Archiv fuer Protistenkunde 146(2):201-205.

Gurley, R. 1893. On the classification of the Myxosporidia, a group of protozoan parasites infesting fishes. Bulletin of the U. S. Fisheries Commission for 1891 11:407-420.

Hazard, E. I. and Oldacre, S. W. 1975. Revision of microsporida (Protozoa) close to Thelohania, with descriptions of one new family, eight new genera, and thirteen new species. U. S. Department of Agriculture Technical Bulletin 1530. U. S. Department of Agriculture, Washington, D. C.

Henneguy, F. and Thélohan, P. 1892. Myxosporidies parasites des muscles chez quelques crustacés décapodes. Annales des Micrographie 4:617-641.

Johnston, L. B., Vernick, S. H. and Sprague, V. 1978. Light and electron microscope study of a new species of Thelohania (Microsporida) in the shrimp Pandalus jordani. Journal of Invertebrate Pathology 32:278-290.

Kishino, H. and Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29:170-179.

Larsson, J. I. R. 1996. Microsporidia in microcrustaceans. Scandinavian Section Society of Protozoologists 18th Annual Meeting. 202.

Larsson, J. I. R. 1999. Identification of microsporidia. Acta Protozoologica 38:161-197.
Lom, J. and Nilsen, F. 2003. Fish microsporidia: fine structural diversity and phylogeny. International Journal for Parasitology 33(2):107-27.

Lom, J., Nilsen, F. and Dyková I. 2001. Thelohania contejeani Henneguy, 1892: dimorphic life cycle and taxonomic affinities, as indicated by ultrastructural and molecular study. Parasitology Research 87:860-872.

Mercier, L. 1909. Contribution à l'étude de la sexualité chez les myxosporidies et chez les microsporidies. Mémoires de L'Academie Royale de Belgique, Classe des Sciences 2(2):1-50.

Moodie, E. G., Le Jambre, L. F. and Katz, M. E. 2003. Thelohania montirivulorum sp. nov. (Microspora: Thelohaniidae), a parasite of the Australian freshwater crayfish, Cherax destructor (Decapoda: Parastacidae): fine ultrastructure, molecular characteristics and phylogenetic relationships. Parasitology Research 91:215-228.

Nilsen, F. 2000. Small subunit ribosomal DNA phylogeny of microsporidia with particular reference to genera that infect fish. Journal of Parasitology 86(1):128-133.

Nilsen, F. and Chen, W. J. 2001. rDNA phylogeny of Intrapredatorus barri (Microsporida: Amblyosporidae) parasitic to Culex fuscanus Wiedemann (Diptera: Culicidae). Parasitology 122:617-623.

Nilsen, F., Endresen, C. and Hordvik, I. 1998. Molecular phylogeny of microsporidians with particular reference to species that infect the muscles of fish. Journal of Eukaryotic Microbiology 45(5):535-43.

Olson, R. E. and Lannan, C. N. 1984. Prevalence of microsporidian infection in commercially caught pink shrimp, Pandalus jordani. Journal of Invertebrate Pathology 43:407-413.

Ovcharenko, N. A. and Koval'chuk, A. V. 1995. About the effect of the environment on intracellular parasites of hydrobionts. Gidrobiologicheskii Zhurnal 31(2):65-73.

Page, R. D. 1996. TreeView: an application to display phylogenetic trees on personal computers. Computer Applications in Biosciences 12:357-358.

Posada, D. and Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14(9):817-818.

Quicke, D. L. J. 1993. Principles and techniques of contemporary taxonomy. Blackie Academic \& Professional. London. 311 pp .

Refardt, D., Canning, E. U., Mathis, A., Cheney, S. A., Lafranchi-Tristem, N. J. and Ebert, D. 2002. Small subunit ribosomal DNA phylogeny of microsporidia that infect Daphnia (Crustacea: Cladocera). Parasitology 124:381-389.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology. 51(3):492-508.

Shimodaira, H. and Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17(12):1246-7.

Sprague, V. 1977. Annotated list of species of Microsporidia. In Comparative Pathology. 2. Systematics of the Microsporidia. Plenum Press, New York, New York, 333 p.

Sprague, V., Becnel, J. J. and Hazard, E. I. 1992. Taxonomy of phylum Microspora. Critical Reviews in Microbiology 18(5/6):285-395.

Swofford, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Terry, R. S., MacNeil, C., Dick, J. T. A., Smith, J. E. and Dunn, A. M. 2003. Resolution of a taxonomic conundrum: An ultrastructural and molecular description of the life cycle of Pleistophora mulleri (Pfeiffer 1895; Georgevitch 1929). Journal of Eukaryotic Microbiology 50(4):266-27.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.

Van de Peer, Y., De Rijik, R., Wuyts, J. Winkelmans, T. and De Wachter, R. 2000. The European small subunit ribosomal RNA database. Nucleic Acids Research 28:175-176.

Vernick, S. H., Sprague, V. and Krause, D. 1977. Some ultrastructural and functional aspects of the golgi apparatus of Thelohania sp. (Microsporida) in the shrimp Pandalus jordani Rahbun. Journal of Protozoology 24(1):9499.

Vivarès, C. P. and Sprague, V. 1979. The fine structure of Ameson pulvis (Microspora, Microsporidia) and its implications regarding classification and chromosome cycle. Journal of Invertebrate Pathology 33(1):40-52.

Vossbrinck, C. R., Andreadis, T. G., and Debrunner-Vossbrinck, B. A. 1998. Verification of intermediate hosts in the life cycles of microsporidia by small subunit rDNA sequencing. Journal of Eukaryotic Microbiology 45:290-292.

Vossbrinck, C. R., Andreadis, T. G., Vávra, J. and Becnel, J. J. 2004. Molecular phylogeny and evolution of mosquito parasitic microsporidia (Microsporidia: Amblyosporidae). Journal of Eukaryotic Microbiology 51(1):88-95.

Weiss, L. M. and Vossbrinck, C. R. 1999. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 129-171.

Table 5.1: Microsporidian rDNA sequences included in phylogenetic analysis showing Genbank accession numbers, current family designation, and host name. $\mathrm{F}=$ fish, $\mathrm{I}=$ insect, $\mathrm{C}=$ Crustacea, $\mathrm{B}=$ bryozoans, $\mathrm{O}=$ oligochaete.

| Species | Accession | Family | Host |
| :---: | :---: | :---: | :---: |
| Amblyospora californica | U68473 | Amblyosporidae | I: Culex tarsalis |
| Amblyospora connecticus | AF025685 | Amblyosporidae | I: Aedes cantator |
| Ameson michaelis | L15741 | Pereziidae | C: Callinectes sapidus |
| Antonospora scoticae | AF024655 | not placed | I: Antonospora scoticae |
| Bacillidium sp. | AF104087 | Mrazekiidae | O: Lumbriculus sp. |
| Bryonosema plumatellae | AF484692 | Pseudonosematidae | B: Plumatella nitens |
| Glugea americanus | AF056014 | Glugeidae | F: Lophius americanus |
| Glugea anomala Nil | AF044391 | Glugeidae | F: Gasterosteus aculeatus |
| Gurleya vavrai | AF394526 | Gurleyidae | C: Cladocera |
| Heterosporis sp. PF | AF356225 | Pleistophoridae | F: Perca flavescens |
| Hyalinocysta champmani | AF483837 | Thelohaniidae | I: Culiseta melanura |
| Kabatana takedai | AF356222 | not placed | F: Oncorhynchus masu |
| Larssonia obtusa | AF394527 | Thelohaniidae | C: Cladocera |
| Loma sp. Nil | AF104081 | Glugeidae | F: Encelyopus cimbrius |
| Microgemma caulleryi | AY033054 | Unikaryonidae | F: Hyperoplus lanceolatus |
| Microgemma sp. | AJ252952 | Unikaryonidae | F: Taurulus bubalis |
| Microsporidium sp. JES2002G | AJ438962 | not placed | C: Gammarus chevreuxi |
| Microsporidium prosopium | AF151529 | not placed | F: Prosopium williamsoni |
| Microsporidium sp. exGammarus | AF397404 | not placed | C: Gammarus duebeni duebeni |
| Microsporidium sp. JES2002B | AJ438955 | not placed | C: Gammarus duebeni celticus |
| Microsporidium sp. JES2002C | AJ438957 | not placed | C: Echinogammarus berilloni |
| Microsporidium sp. RSB1 | AJ295323 | not placed | F: Pagrus major |
| Nosema algerae | AF069063 | Nosematidae | I: Anopheles stephensi |
| Nosema bombycis | D85504 | Nosematidae | I: Bombyx mori |
| Nucleospora salmonis | AF186002 | Enterocytozoonidae | F: Salmo salar |
| Perezia nelsoni | AJ252959 | Pleistophoridae | C: Litopenaeus setiferus |
| Pleistophora mirandellae | AJ252954 | Pleistophoridae | F: Rutilus rutilus |
| Pleistophora sp. TB | AJ252957 | Pleistophoridae | F: Taurulus bubalis |
| Spraguea lophii | AF056013 | Spragueidae | F: Lophius piscatorius |
| Tetramicra brevifilum | AF364303 | Tetramicridae | F: Scophthalmus maximus |
| Thelohania contejeani | AF492594 | Thelohaniidae | C: Astacus fluviatilis |
| Thelohania montirivulorum | AY183664 | Thelohaniidae | C: Cherax destructor destructor |
| Thelohania parastaci WA1 | AF294780 | Thelohaniidae | C: Cherax destructor albidus |
| Thelohania solenopsae | AF031538 | Thelohaniidae | I: Solenopsis invicta |
| Thelohania sp. | AF031537 | Thelohaniidae | I: Solenopsis richteri |
| Unidentified microsporidian GHB1 | AJ295324 | not placed | F: Sparus aurata |
| Unidentified microsporidian S1 | AJ295328 | not placed | C: Metapenaeus joineri |
| Visvesvaria acridophagus | AF024658 | not placed | I: mosquito |
| Visvesvaria algerae | AF024656 | not placed | I: mosquito |
| Weiseria palustris | AF132544 | Caudosporidae | I: Cnephia ornithophilia |



Figure 5.1: Spores of Thelohania butleri under light microscopy, showing always eight spores per sporophorous vesicle (SV), and the persistence of the SV. Large arrow = sporophorous vesicles (SV) containing eight spores; Small arrow = SVs persisting once spores are gone. Scale bar $=10 \mu \mathrm{~m}$.


Figure 5.2: Developmental stages of Thelohania butleri undergoing binary fission, showing cells in the early 2-cell stage (center), one with a diplokaryotic nucleus (large black arrow) sharing a sporophorous vesicle (SV) with a sister-sporont (note golgi-like vesicles, small black arrow); also showing 2 pairs of cells at the 4 -cell stage (top left and top center), showing unikaryotic nuclei and the presence of dark, finely granular material (hollow black arrow) between stages surrounded by finely fibrous material. Note the round, dark structure in the parasite cytoplasm of the top, center cell (white arrow). Transmission electron micrograph. Scale bar $=1 \mu \mathrm{~m}$.


Figures 5.3-5.6: Developmental stages of Thelohania butleri. Fig. 5.3 Early stages (2 and 4cell stages) in the musculature of smooth pink shrimp showing small, dark granular vesicles (small white arrows) in the parasite cell cytoplasm. Dark, finely granular material (hollow black arrows) is forming between cells). Fig. 5.4 Early 8 -cell stage ( 5 cells visible here within a sporophorous vesicle). Fig. 5.5 Sporoblast (large solid arrow) with advanced polar filament formation, and granular or perforated material (large hollow arrow) and dark vesicles (small arrows) in episporontal space. Fig. 5.6 Mature spores and earlier developmental spaces within SV membranes, showing perforated dark vesicles (hollow arrow). Scale bars $=1 \mu \mathrm{~m}$.


Figures 5.7-5.8 Mature spores of Thelohania butleri. Fig. 5.7 Mature spores with 13 turns of the polar filament in 2 rows (small arrows) and showing details of episporontal inclusions, showing striations of fibrous material (black hollow arrow) and round, dark-material-filled, granular inclusions (hollow white arrow). Fig. 5.8 Mature spore filled with 13 turns of the isofilar polar filament in 2 rows ( $5+6$ plus 2 ). Scale bars $=1 \mu \mathrm{~m}$.

Table 5.2: Maximum parsimony bootstrap consensus tree results for various different input data sets, showing the effects on tree topology and bootstrap support for the fish/crustacean (F/C) clade of adding or removing from the alignment: regions with uncertain alignment (U), regions with missing and ambiguous characters (M), or positions with gaps (G); as well as the effects of adding a matrix that encodes gap end positions as characters (+G), using Gap Matrix in REALEM 1.01 (this study). $-\mathrm{U}=$ uncertain alignment regions removed; $-\mathrm{M}=$ missing/ambiguous characters removed; $-\mathrm{G}=$ positions with gaps removed; $+\mathrm{G}=$ gap matrix added (no nucleotide characters); $\mathrm{F} / \mathrm{C}=$ fish/crustacean clade; $\mathrm{F}=$ fish/Gammarus subclades within F/C (see Fig. 5.11); sister to F = branching after the Ameson michaelis clade within F/C clade; $\mathrm{G}=$ Gammarus subclade within F/C (see Fig. 5.11); * gap matrix analyzed alone, with no nucleotide substitutional information; ** Nucleospora salmonis included within F/C

| Data <br> Set | \# of <br> alignment <br> positions | Is T. butleri <br> within F/C <br> clade? | Bootstrap <br> support for <br> F/C clade | T. butleri's <br> place in F/C <br> clade |
| :---: | :---: | :---: | :---: | :---: |
| Without gap matrix: |  |  |  |  |
| -U -M -G | 524 | y | 90 | within F |
| -U -G | 783 | y | 75 | sister to F |
| -U | 993 | y | 80 | sister to F |
| full data | 1675 | y | 85 | sister to F |
| With gap matrix added: |  |  |  |  |
| only +G* | 147 | y | $<50 * *$ | within F |
| -U -M -G +G | 671 | y | 70 | sister to G |
| -U +G | 1140 | $y$ | 79 | sister to F |

Figure 5.9: SSU rDNA results from 524 alignment positions, showing relative branch length differences and the position of Thelohania butleri always within the fish/crustacean "F/C" clade (branch shown in bold) far from clades T and V (enclosed in boxes) containing other Thelohania species (heuristic search trees using logDet/paralinear distance, maximum likelihood GTR $+\mathrm{I}+\mathrm{G}$ model, and one of two very similar most parsimonious trees). Branch to Ameson michaelis is quite long in distance tree, and it falls outside F/C clade in this tree (see asterisk). Host groups are indicated on the maximum likelihood tree in gray typeface. Not indicated on figure, for clarity of figure, are Nucleospora salmonis = Fish host, Nosema bombycis = Insect host, Antonospora scoticae $=$ Insect host.



Figure 5.10: Distance 50\% majority-rule consensus tree from bootstrap heuristic search (1000 replicates) in PAUP*, with bootstrap values $>50 \%$ shown on branches for 524 SSU rDNA positions. Bootstrap support for the fish/crustacean ( $\mathrm{F} / \mathrm{C}$ ) clade is low (bold), and support is low for branches (asterisks) leading to Thelohania butleri in the F/C clade. "G"= Gammarusparasitic sub-clade. Boxes show strongly supported Thelohania spp. clades ( $\mathrm{T}=$ clade with crayfish-parasitic Thelohania spp.; V = clade with ant-parasitic Thelohania spp.)


Figure 5.11: Maximum parsimony $50 \%$ majority rule consensus tree from bootstrap heuristic search ( 1000 replicates) in PAUP*, with bootstrap values $>50 \%$ shown on branches from 524 nucleotide positions of SSU rDNA. Notice high support for the fish/crustacean (F/C) clade (bold), and low support for branches (asterisks) leading to Thelohania butleri in the F/C clade. $\mathrm{G}=$ Gammarus-parasitic subclade; F = Fish-parasitic subclade (see Table 5.2); M = Marine decapod-parasitic subclade. Boxes show strongly supported Thelohania spp. clades ( $\mathrm{T}=$ clade with crayfish-parasitic Thelohania spp.; $\mathrm{V}=$ clade with ant-parasitic Thelohania spp.).


Figure 5.12: Maximum likelihood 50\% majority rule bootstrap consensus tree (faststep, 100 replicates) in PAUP* with the GTR $+\mathrm{I}+\mathrm{G}$ model of substitution, with bootstrap values $>50 \%$ shown on branches for 524 alignment positions of SSU rDNA. Support is low (bold and asterisks) for some branches leading to Thelohania butleri. $\mathrm{G}=$ Gammarus-parasitic subclade; $\mathrm{F}=$ Fish-parasitic subclade (see Table 5.2); $\mathrm{M}=$ Marine decapod-parasitic subclade. Boxes show strongly supported Thelohania spp. clades ( $\mathrm{T}=$ clade with crayfish-parasitic Thelohania spp.; V = clade with ant-parasitic Thelohania spp.).

Table 5.3: Comparison of unconstrained and monophyly-constrained trees from 524 alignment positions of SSU rDNA using the Approximately Unbiased or "AU" test, weighted and unweighted Kishino-Hasegawa or "KH" and Shimodaira-Hasegawa or "SH" tests in the CONSEL software package. Tree numbers correspond to trees shown in Figure 5.13.

| Tree <br> $\#$ | Constraint | p-values from CONSEL tests |  |  |  |  |
| :---: | :--- | :--- | :--- | :--- | :--- | :--- |
|  | AU | KH | SH | WKH | WSH |  |
| $\mathbf{1}$ | Ameson michaelis | 0.539 | 0.524 | 0.940 | 0.524 | 0.944 |
| 2 | Species from decapod crustaceans | 0.001 | 0.002 | 0.312 | 0.002 | 0.011 |
| 3 | Thelohania species from crayfish | 0.001 | 0.002 | 0.312 | 0.002 | 0.011 |
| 4 | All species from crustaceans | $1 \mathrm{e}-008$ | 0 | $1 \mathrm{e}-004$ | 0 | 0 |
| $\mathbf{5}$ | Unidentified microsporidium S1 | $4 \mathrm{e}-090$ | 0 | 0 | 0 | 0 |
| $\mathbf{6}$ | All members of genus Thelohania | $3 \mathrm{e}-006$ | 0 | 0 | 0 | 0 |
| 7 | All members of family Thelohaniidae | $5 \mathrm{e}-027$ | 0 | 0 | 0 | 0 |

Figure 5.13: Tree topologies used in testing various hypothetical groups. Monophyletic groups (shown in boxes), corresponding to various hypotheses, were created as constraints prior to heuristic tree searches. From these topologies likelihood scores (site likelihood scores in PAUP* 4.0-10b Swofford, 2000) were calculated under the best-fit substitution model using parameters estimated from the data in Modeltest V. 3.06 (Posada \& Crandall, 1998) for statistical tests of best trees using the CONSEL V.0.1.f software package (Shimodaira \& Hasegawa, 2001) (results shown in Table 7).



# Chapter 6: A new microsporidian, Vossbrinckus richardi n. gen., n. sp. (Microsporidia; Golbergiidae) from a copepod: ultrastructural features and phylogenetic placement. 

## INTRODUCTION

Microsporidians are small, single-celled, obligate intracellular parasites of a wide range of taxa, including many fishes, insects and crustaceans. Freshwater crustaceans, especially Decapoda, Copepoda, and Cladocera, commonly serve as either definitive or intermediate hosts for microsporidians (Sprague, 1977; Becnel \& Andreadis, 1999; Vossbrinck et al., 2004). It has been suggested that every pond or freshwater pool with a community of cladocerans or copepods worldwide has a unique community of microsporidians yet to be discovered (Larsson, 1996). Larsson ( $1996 ; 1999$ ) estimated about $1 / 4$ of all microsporidian genera are from crustacean hosts. These 25 genera include relatively few (about 130) species, suggesting these parasites are diverse at the genus- and higher-levels in these hosts. Furthermore, transmission electron microscopy (TEM) and DNA have shown some species previously known only from light microscopy were incorrectly placed together (Sprague, 1977; Larsson, 1999; Weiss \& Vossbrinck, 1999; Lom, 2002; Lom \& Nilsen, 2003; Moodie et al., 2003), suggesting what is known from taxonomic diversity may not reflect the true phylogenetic diversity. However, recent rDNA evidence shows microsporidia from aquatic Crustacea occur throughout the microsporidian tree (Refardt et al., 2002), and confirms Larsson's (1996) view that crustaceanparasitic microsporidia represent an unusually broad array of forms. This study examines an unusual new microsporidian found in a freshwater copepod, and considers questions about morphological adaptation in related microsporidia.

Freshwater cyclopoid copepods (family Cyclopidae) frequently serve as obligate intermediate hosts in the life cycles of microsporidians of the large ( $>90$ species) Amblyospora group (Andreadis, 1983; Andreadis, 1988a, b; Sweeney et al., 1990; Darwish \& Canning, 1991;

Becnel, 1994; Lucarotti \& Andreadis, 1995; Becnel \& Andreadis, 1998; Micieli et al., 2000; Micieli et al., 2001; Andreadis \& Vossbrinck, 2002; Andreadis, 2002). Careful transmission studies (Andreadis, 1985; Andreadis, 1988a; b; Sweeney et al., 1990; Micieli et al., 2000) and rDNA studies (Vossbrinck et al., 1998; Andreadis \& Vossbrinck, 2002; Vossbrinck et al., 2004) confirmed that some Amblyospora-like species have complex permutations of a two- or threehost life cycle involving vertical and horizontal transmission at different stages (Becnel, 1994; Becnel \& Andreadis, 1999). The most complex life-cycles occur in Amblyospora-like species that alternate between 3 hosts: copepods, larval mosquitoes, and adult mosquitoes. This occurs in at least 12 of the "true" Amblyospora species. Simpler life-cycles occur in close-relatives. For example, Amblyospora-like species that fall within the true Amblyospora-clade may use only mosquito hosts (e.g. Edhazardia aedis and Culicospora magna), and species just outside the Amblyospora-clade may use only aquatic crustacean hosts (e.g. Trichotuzetia guttata) or only mosquito hosts (e.g. Hazardia milleri) (Becnel \& Andreadis, 1999; Vossbrinck et al., 2004).

Some molecular phylogenies have suggested the complex life-cycle, involving a copepod intermediate host, is the most primitive state in the microsporidia (Baker et al., 1997; Nilsen \& Chen, 2001), based on evidence that the Amblyospora-like group appears to branch early in the tree (Vossbrinck et al., 2004). Other evidence suggests a nearby group of species, called the "Aquatic Outgroup" and including mostly one-host (monoxenic) species from freshwater aquatic crustaceans, may be earlier-branching than two- or three-host Amblyosporalike species (Vossbrinck et al., 2004). Nevertheless, many molecular loci and other evidence now place microsporidia within the Fungi (Edlind et al., 1996; Keeling \& Doolittle, 1996; Germot et al., 1997; Hirt et al., 1997; Hirt et al., 1999; Keeling et al, 2000), which leads one to speculate whether the earliest microsporidians came from Fungi that parasitize aquatic organisms. Conversely, other phylogenies have placed monoxenic species from terrestrial hosts, most notably Antonospora scoticae Fries, Paxton, Tengö, Slemenda, da Silva, \& Pieniazek, 1999, from a bee, in a basal place in the tree (Weiss \& Vossbrinck, 1999). Such relationships seem to differ depending on the choice of both ingroup and outgroup taxa. Thus, the question of which host condition is most primitive in the microsporidia remains open. An understanding of evolution and host-use of microsporidia from aquatic Crustacea may, therefore, be of interest on several levels.

To date, few molecular studies have focused on microsporidia from copepods. Yet microsporidia from copepods may be of interest not only phylogenetically, but also for
ecological reasons. For example, copepod-parasitic microsporidians may play an important role in host population regulation (Andreadis, 1988a; Becnel et al., 1995; Andreadis, 1999; Dunn \& Smith, 2001; Micieli et al, 2001; Andreadis, 2002) for either copepod or mosquito host. Where the host is a pest or potentially serious disease vector (e.g. mosquito species), the microsporidia may be significant epidemiological interest. Even for species that use only a copepod, the host may still play an important role in the dynamics of pond or lake communities; hence, these may be worth investigation. Molecular tools can, therefore, be developed for specific identification (diagnosis of species) in the field as well as for phylogenetic study.

This study examined a new microsporidian from a cyclopoid copepod found in Vancouver, British Columbia, Canada. This microsporidian's morphological and DNA features were different enough from those of other genera to require erection of a new genus and species. First, this new genus and species was described and differentiated from others using ultrastructural data from light and transmission electron microscopy (TEM). The second goal of this study was to use the molecular data from three ribosomal DNA (rDNA) regions, including the small subunit (SSU), internal transcribed spacer (ITS) and partial large subunit (LSU) genes, to substantiate evidence for the new taxon, and examine relationships in the aquatic crustaceanparasitic microsporidia using the AU test of Shimodaira \& Hasegawa (2001). The AU test is explained in Appendices 5 to 9, and justified in Appendix 6 (based on arguments in Kishino \& Hasegawa, 1989; Goldman et al., 2000; Shimodaira \& Hasegawa, 2001; Shimodaira, 2002). The third goal of this study was to consider and discuss evidence for a definitive mosquito host for this species. Finally, the last goal was to provide rDNA sequence that can be used to develop a species-specific molecular ecological tool in future work on this copepod parasite.

## MATERIALS AND METHODS

## Specimen collection

Copepods were collected from a roadside fresh-water ditch in Southlands, north of the Fraser River North Arm, in Vancouver, British Columbia, Canada, at four times per year, in late February, May, August and November for two successive years. The ditch was approximately 1 $m$ deep $\times 2 \mathrm{~m}$ wide $\times 10 \mathrm{~m}$ long and supported a year-round community of fauna, including small fishes. Each sample consisted of about 30 L of water and organisms collected using a fine-mesh dip-net. Water was filtered first through coarse, then successively finer mesh, and finally through a $50 \mu \mathrm{~m}$ mesh nitex filter to select copepod-sized organisms. About 50-300 cyclopoid copepods were examined per sampling month. Live copepods were examined under dissection microscope before being gently placed under cover slips and examined at high magnification for spores. Uninfected and infected copepods were then squashed to better examine tissues for spores. When copepods were heavily infected, the surrounding liquid containing spores was carefully removed from under cover slips, and some of this spore-filled liquid was pipetted into microcentrifuge tubes for DNA extraction, while the remaining copepod body was fixed for transmission electron microscopy (TEM) as described below.

## Light and transmission electron microscopy (TEM)

Tissue for transmission electron microscopy was placed in 4\% glutaraldehyde and cacodylate ( 0.1 M ) for 12 hours, then post fixed in $1 \%$ osmium tetraoxide for 1 hour. Tissue was embedded in standard Spurr's resin. Ultra thin sections were lifted onto copper grids and stained with $2 \%$ aqueous uranyl acetate and standard Reynold's lead citrate. Ultrastructural features were observed and photographed on a Zeiss 10C Transmission Electron Microscope and negatives were scanned at high resolution (1200 dpi) and opened in Adobe Photoshop 6.0 for closer examination and measurement.

Fresh spores were measured from drawings at 1000X magnification using a drawing tube with scale bars calculated using a slide micrometer and ocular micrometer to an estimated accuracy of $\pm 0.05 \mu \mathrm{~m}$. Spores fixed in glutaraldehyde and embedded in Spurr's resin were measured from photographs taken at 1000X magnification, scanned at high resolution (2000 dpi), opened in Adobe Photoshop 6.0 and blown up with the zoom feature to about 15 cm for measurement on screen.

## DNA isolation

Fresh spores were heated in $30 \mu \mathrm{~m}$ lysis buffer ( 10 mM Tris, 1 mM EDTA, 10 mM $\mathrm{NaCl}, 1 \% \mathrm{SDS}$ ) at $95-100^{\circ} \mathrm{C}$ for $2-5 \mathrm{~min}$ to break them open, and immediately digested with $0.5 \mathrm{mg} / \mathrm{ml}$ proteinase K for 2 hours at $37^{\circ} \mathrm{C}$ in a rotating incubator. DNA was phenol chloroform extracted twice with phenol: chloroform: isoamyl alcohol 25:24:1, and once with chloroform: isoamyl alcohol $24: 1$, precipitated in cold $95 \%$ ethanol, washed twice with $70 \%$ ethanol, vacuum dried, resuspended in $10 \mu \mathrm{l}$ distilled water and stored for use at $-20^{\circ} \mathrm{C}$.

## Polymerase chain reaction (PCR)

PCR for sequencing and cloning of ribosomal DNA (rDNA) was performed in a Perkin Elmer Cetus DNA Thermal Cycler 480 in $25 \mu \mathrm{l}$ reactions with standard PCR buffer 2.5 mM $\mathrm{MgCl}_{2}, 0.2 \mathrm{mM} \mathrm{dNTP}, 15 \mathrm{pmol}$ of each primer, and 1-3 units of Taq DNA polymerase (Gibco BRL), conditions: $95^{\circ} \mathrm{C}$ for $2 \mathrm{~min}, 3$ cycles of $95^{\circ} \mathrm{C}$ for $50 \mathrm{sec}, 50^{\circ} \mathrm{C}$ for $40 \mathrm{sec}, 72^{\circ} \mathrm{C}$ for 2 min, then 35 cycles of $95^{\circ} \mathrm{C}$ for $50 \mathrm{sec}, 54^{\circ} \mathrm{C}$ for $30 \mathrm{sec}, 72^{\circ} \mathrm{C}$ for 90 sec , and a final extension of $72^{\circ} \mathrm{C}$ for 5 min . Primers amplifying the small subunit (SSU), internal transcribed spacer (ITS) and large subunit (LSU) ribosomal DNA (rDNA) were: forward M5P - CAC CAG GTT GAT TCT GCC pos. 1-18 at $5^{\prime}$ end of the SSU (18eMIC in Docker, Kent et al. 1997); Seq1f CGT TGT AGT TCT AGC AGT pos. 719-736 in the SSU (provided by M. F. Docker); and reverse SeqR - AAC AGG GAC KYA TTC ATC pos. 1218-1235 in the SSU (this study); 580R GGT CCG TGT TTC AAG ACG G pos. 1847-1865 in the LSU (Vossbrinck et al., 1987). PCR products were visualized in ethidium bromide stained $1.5-2 \%$ agarose gels. Products of the correct size were cloned as described below or excised from gels and freeze-thaw extracted from agarose to be sequenced directly.

## Cloning

PCR products were isolated in $0.8 \%$ agarose and cleaned for ligation using Ultraclean 15 MOBIO DNA Purification Kit (BIO/CAN Scientific Inc. Mississauga, ON) and cloned using the TOPO TA Cloning PCR Version 2.1 (Invitrogen Corp., Carlsbad, CA) using $1 / 2$ volume. Clones were screened for presence of the insert in $10 \mu \mathrm{PCR}$ reactions using Taq DNA Polymerase (Invitrogen Corp., Carlsbad, CA) with standard reagents and screening primers M13-20 and M13 Rev (conditions: $94^{\circ} \mathrm{C}$ for $2 \mathrm{~min}, 34$ cycles of $92^{\circ} \mathrm{C}$ for $45 \mathrm{sec}, 55^{\circ} \mathrm{C}$ for 45 $\sec 72^{\circ} \mathrm{C}$ for 1 min 30 sec , followed by $72^{\circ} \mathrm{C}$ for 5 min ). Positive clones from master plates were grown in 3 ml of standard LB culture with 50 mM ampicillin by shaking at 220 rpm at 37
${ }^{\circ} \mathrm{C}$ overnight. Plasmids were isolated for sequencing using the Rapid Plasmid Miniprep System (Gibco BRL, Gaithersburg, MD) following directions of the manufacturer.

## DNA sequencing

Sequencing was performed on the ABI PRISM 377 DNA automated sequencer using BigDye Terminator Version 3.1 fluorescent dye-labelled terminators with forward and reverse primers and PCR conditions as recommended for the Taq terminators. Multiple PCR products and multiple clones were sequenced in both directions to check for Taq or sequencer errors.

## Flip algorithm

Sequencing directly from PCR products sometimes produced sequences that were double, having two different nucleotide signals at almost every site after a stretch of normal single nucleotide signals. The correct signals from the doubled portion of sequences were extracted using Flip Analyzer of REALEM Version 1.01 (developed this study). For more details, see Appendix 1.

## Alignment

Ribosomal DNA sequences were aligned with a large set of rDNA sequences from microsporidia, fungi, and other outgroups from Genbank by first grouping them into clusters of taxa known to be related from previous phylogenetic studies, then by aligning more conserved regions by eye using ESEE 3.2s (Eyeball SEquence Editor, Eric Cabot, 1998). Less conserved regions of the rDNA were aligned with emphasis on conservation of secondary structure using models from other microsporidians in the rDNA structural database (Van de Peer et al., 2000) and also using Clustal W (version 1.74) (Thompson et al., 1994) under a variety of parameters emphasizing transitions over transversions for more closely related taxa or allowing frequent gaps near rDNA loop regions. A preliminary phylogenetic analysis was performed to help choose taxa that could be removed to speed analyses. All taxa with Blast (NCBI) nucleotide similarity to the new species were included, as were many members of the Amblyospora and nearby clades, as well as several representative taxa from clades distant from Amblyospora species (Table 6.1).

## Phylogenetic analysis

Phylogenetic analysis was performed in PAUP*4.0b10 (Swofford, 2001), using all three types of optimality criteria, maximum parsimony (MP), distance (ME) and maximum likelihood (ML) in with heuristic search, random stepwise sequence addition and TBR branch swapping with 10 repetitions, or only one repetition for ML. Unweighted maximum parsimony was
followed by bootstrap resampling with 1000 replicates reported on $50 \%$ majority rule trees. The $\operatorname{logDet}$ /paralinear model was used to calculate pairwise distances and trees were estimated under the minimum evolution (ME) criterion. Bootstrap was replicated 1000 times and reported on a 50\% majority rule tree. Maximum likelihood analysis was performed by first estimating the best model predicted for the data using Modeltest Version 3.06 (Posada \& Crandall, 1998). Maximum likelihood heuristic searches were run using the best model and estimated parameters, and were followed by 100 replicates of bootstrap resampling (faststep search) shown on $50 \%$ majority-rule consensus trees. (For further details see Appendices 2 to 4.)

Different input data sets were created using the character exclusion feature of PAUP* to analyze the effects of missing and ambiguous character removal, gap removal (gap stripping), and removal of different regions of the gene considered to be less or more conserved. Several additional input data sets were created using Gap Matrix of REALEM Version 1.01 (developed this study) to analyze the effects of adding a matrix containing gap information to either the full alignment or to the gap stripped input, as well as to examine a phylogeny estimated from gap information only. See further details in Appendix 1.

## Monophyly constraints and AU tests

Alternate tree topologies corresponding to clades comprised of species of the same host group, genus or family designation were created using text-view versions of PAUP* output trees (TreeView Version 1.6.6; Page, 1996). These altered parenthetic trees were pasted into PAUP* input data along with the monophyly constraint command. Log likelihood branch length scores (lscores) were calculated and pasted into CONSEL Version 0.1 f (Shimodaira \& Hasegawa, 2001), to statistically compare unconstrained and constrained trees using the Approximately Unbiased test (AU test), and other similar tests. For further details and justification for use of this test, see Appendices 5 to 9 . Output test probabilities less than 0.01 are used as evidence for rejection of the null hypothesis that trees are equally likely. Trees included in these comparisons were inspected to be sure they had equal numbers of nodes (and therefore branches), and were topologically alike in clades outside the modified clades.

## RESULTS

## DESCRIPTION ${ }^{1}$

## Genus Vossbrinckus n. gen.

Diagnosis: Spores of three forms in groups of four, eight or 16 without sporophorous vesicle, held together by mucocalyx-like substance in no apparently organized arrangement in haemocoel of cyclopoid copepod. Most common spore pyriform, about 3.7-4.2 $\mu \mathrm{m} \times 2.1-2.8$ $\mu \mathrm{m}$, uninucleate. Second most common spore ovoid and elongate, about $1 / 20$ as frequent as pyriform spores, elongate ovoid and $5.94-6.73 \times 2.11-2.51 \mu \mathrm{~m}$, uninucleate. Rare spore form ovoid, about $3 / 4$ as long as elongate spores and about $2.5 \mu \mathrm{~m}$ wide, binucleate. Pre-sporulation sequence not observed.
Type species: V. richardi n. gen., n. sp.

Etymology: Generic name after Dr. Charles R. Vossbrinck for his contributions to molecular phylogenetics of microsporidia, and for providing the initial idea to look for this microsporidian. Genus Vossbrinckus n. gen. should be regarded as masculine for the purpose of gender agreement with species-group names.

## Vossbrinckus richardi n. gen., n. sp.

(Figs. 6.1-6.27, Table 6.2)

Spores of three forms throughout haemocoel of cyclopoid copepod (COPEPODA; Cyclopidae) in Vancouver, British Columbia, Canada. Most common spores pyriform (Figs. $6.2,6.4,6.5,6.14,6.16,6.18$ ), uninucleate, fresh $3.8 \pm 0.1 \mu \mathrm{~m}$ (range 3.7-4.2 $\mu \mathrm{m}$ ) long x $2.4 \pm$ $0.2 \mu \mathrm{~m}$ (range $2.1-2.8 \mu \mathrm{~m}$ ) wide ( $\mathrm{n}=10$ ); or fixed in glutaraldehyde and embedded in Spurr's resin $3.12 \pm 0.15 \mu \mathrm{~m}$ (range $2.78-3.78 \mu \mathrm{~m}$ ) long x $1.86 \pm 0.16 \mu \mathrm{~m}$ (range $1.54-2.75 \mu \mathrm{~m}$ ) wide $(\mathrm{n}=14)$ with round, wide posterior end, and tapered anterior end, moderately thin endospore ( 90

[^3]-120 nm ), a roughly undulating exospore (Figs. 6.14, 6.16,6.18) in at least two thin layers (about 13 nm and 17 nm ), cytoplasm packed with polyribosomes, small vesicular and lamellar polaroplasts, a posterior vacuole filling up to $1 / 3$ of spore, polar filaments in six isofilar, single coils, uniformly arranged, 120-145 nm wide, with six - eight layers (Fig. 6.17). Elongate spores (Figs. 6.2-6.4, 6.15,6.19) about $1 / 20$ as common as pyriform spores, uninucleate, in resin $6.34 \pm 0.17 \mu \mathrm{~m}$ (range 5.94-6.73 $\mu \mathrm{m}$ ) long $\times 2.35 \pm 0.09 \mu \mathrm{~m}$ (range $2.11-2.51 \mu \mathrm{~m}$ ) wide ( $\mathrm{n}=9$ ), elongated and ovoid, incurved, with bluntly rounded ends with thick endospore 200 460 nm wide, thinner at anterior end, a roughly undulating exospore, cytoplasm packed with large amount of lamellar and vesicular polaroplast, polar filaments isofilar, 100-150 nm wide, in at least four single coils arranged at about 45 degrees from longitudinal axis of spore, a large, centrally located polar capsule (Fig. 6.19), having a round, electron-dense structure (Fig. 6.15) of unknown nature, about $0.5 \mu \mathrm{~m}$ wide at periphery near middle of spore. Shorter ovoid, binucleate spores (Figs. 6.3, 6.20, 6.21) observed more rarely, about same width and $3 / 4$ length compared to elongate spores, with thick endospore (250-350 nm), fairly smooth exospore, five to six isofilar polar filament coils, and packed with vesicular polaroplast. Pre-sporulation stages not observed. Sporulation stages observed (Figs. 6.6-6.13, 6.18) include sporonts (Figs. 6.66.8 ) with thin, electron-dense surface coat and a low level of vacuolation, early, thin-walled early sporoblasts (Figs. 6.9, 6.10), late sporoblasts (Figs. 6.11-6.13) with relatively undifferentiated thick walls (200-600 nm ), and spores, some stages with a single, round, electron-dense structure (Fig. 6.6, 6.8, 6.12, 6.15) of unknown nature, about $0.5 \mu \mathrm{~m}$ wide peripheral to nucleus, resembling an early polar capsule primordium. Spores and sporoblasts in clusters of four, eight or 16 cells (Figs. 6.1, 6.2) held together in mucous-like material, or singly (perhaps an artifact from mechanical damage during squashing), in direct contact with host cytoplasm. Sporophorous vesicles not observed. Mucous-like material amorphous and relatively electron-lucent, in streams holding spores together under light and electron microscopy (Figs. 6.2, 6.5, 6.18). Sporulation sequence unknown. Small subunit ribosomal DNA sequence similar ( $5.9 \%$ difference) to Marssoniella elegans sensu Vossbrinck et al. (2004) (syn. Gurleya elegans of Vossbrinck et al., 2002 Genbank submission AY090041), and phylogenetic placement is together with Hazardia Weiser, 1977 species, within a larger clade containing other species of Gurleya Doflein, 1898.

## Taxonomic summary

## Type host: COPEPODA; Cyclopidae.

Type locality: Vancouver, British Columbia, Canada.
Site of infection: Haemocoel, throughout body.
Prevalence: 0-3 \%
Material deposited: (will be submitted to Canadian Museum of Nature, Invertebrate Collection (Parasites), Ottawa, Canada; ribosomal DNA sequence information will be submitted to Genbank prior to publication).
Etymology: Generic name after Dr. Charles R. Vossbrinck as above. Species name after late mathematician and evolution hobbyist, Dr. Richard L. W. Brown.

## Ecology and pathology

From 50-300 copepods examined at each of four intervals (in February, May, August, and November) over two years (three times in the month of August), infection was observed only in August, at a prevalence of about 0-3\%. Infected copepods could not be distinguished from uninfected copepods by the naked eye or dissecting microscope at 25 X magnification against a black background, as is the case with some microsporidian species. Infections were recognized only by the presence of spores at 250 X or higher magnification after copepods were pressed between a slide and cover slip. Several hundred male copepods were examined, but infections were seen only in female adult (or larger sized) copepods. A large number of females with egg sacs were examined ( $>200$ ), and spores were not observed in egg sacs or haemocoel of these copepods. All infected copepods appeared to be without egg sacs. Heavily infected copepods had spores distributed throughout the haemocoel (Fig. 6.1). Site of infection in lightly infected cases was difficult to observe due to squashing necessary to expose spores under light microscopy.

## Ribosomal DNA sequence characteristics

Of the 1785 nucleotide positions sequenced, 1293 were from the small subunit (SSU), 26 from the internal transcribed spacer (ITS), and 466 from the large subunit (LSU) rDNA. A low level of intra-clone (intraspecific) difference ( $0.118 \%$ ) was observed in eight similar sequences obtained from clones and directly sequenced PCR products. A 1 bp indel difference between sequences was detected in both cloned and directly sequenced PCR products, at position 1631 in the LSU region (see Flip algorithm in Materials and Methods). Since intraspecific variation was low, a single representative sequence will be used in phylogenetic analyses, and referred to hereafter as Vossbrinckus richardi n. gen., n. sp. Percent GC content was $56.9 \%$ in the SSU region of this sequence. This sequence was most similar to that from Marssoniella elegans sensu Vossbrinck et al. (2004) (syn. Gurleya elegans of Vossbrinck et al., 2002 Genbank
submission AY090041) from the copepod Cyclops vicinus from the Czech Republic, having 5.9 \% difference in the SSU region (78 substitutions and one indel over 1293 bp ).

## Phylogenetic analysis of rDNA

Vossbrinckus richardin. gen., n. sp. fell consistently within a group consisting of species of Hazardia, Gurleya, Marssoniella, Larssonia, and Berewaldia, hereafter referred to as clade "A" (after "Aquatic Outgroup") (Figs. 6.24-6.26). Clade A plus Amblyospora sp. CHE and Vairimorpha sp. were equivalent to the "Aquatic Outgroup" of Vossbrinck et al. (2004), a sistergroup to the large Amblyospora clade. Clade A was obtained in all analyses of SSU rDNA, including parsimony (MP), distance (ME), and maximum likelihood (ML) (best-fit model from Modeltest was GTR $+\mathrm{I}+\mathrm{G}$ ), and using a variety of input data sets. For example, clade A was obtained from an input consisting of 82 alignment positions of gap matrix information (only gap end information, no nucleotide characters) although with < 50\% bootstrap support.

Phylogenetic analyses were performed with input data sets having fewer or more characters (e.g. 825 to 2426 alignment positions) through elimination or addition of regions with uncertain alignment removing missing or ambiguous data, gaps, or addition of a gap matrix. These produced topologies that were similar and had roughly the same level of bootstrap support for clade A (e.g. 65-73 \% bootstrap support in MP analysis), hence a conservative input set consisting of 1022 positions was used for all further trees shown in this study.

Within clade A, V. richardin. gen., n. sp. was always paired with Marssoniella elegans sensu Vossbrinck et al. (2004) with high ( $>97 \%$ ) bootstrap support (Figs. 6.24 to 6.26 ), and these species were always grouped with two Hazardia species, except in some analyses using a reduced input ( 825 positions) in which the Hazardia species fell basally to the remaining taxa in clade A, but such placement always had low ( $<50 \%$ bootstrap) support. Vossbrinckus richardi n. gen., n. sp. grouped together with Hazardia species to form a clade hereafter referred to a clade "H" (after "Hazardia"), that was sister to clade "G" (after "Gurleya"), containing Gurleya species, Larssonia obtusa, and Berwaldia schaefernai. This arrangement was more strongly supported in distance (ME) analyses (Fig. 6.25), less strongly supported in parsimony (MP) (Fig. 6.24), and weakly supported in maximum likelihood (ML) searches ( $<50 \%$ bootstrap support) (Fig. 6.26).

Tree topologies and bootstrap support were similar in analyses repeated without the fungal outgroups.

## Monophyly constraints and AU-tests of rDNA region

Alternative hypothetical topologies were constructed around the clades of interest to test the robustness of relationships (Fig. 6.27). Trees were compared using AU, SH, KH, and weighted SH and KH tests of Shimodaira \& Hasegawa (2001). See Appendices 5 to 9 for further explanation. These tests were performed on the ML log likelihood branch scores (using the GTR $+\mathrm{I}+\mathrm{G}$ substitution model) for the original ML tree versus multiple alternative hypothetical-relationship trees listed in Table 6.2. All alternate placements of $V$. richardi n . gen., n. sp. produced low AU test p-value test results ( $<0.01$ ) suggesting these topologies should be rejected. The unweighted SH test could not reject placement of V. richardi n. gen., n. sp. together with Gurleya species and M. elegans sensu Vossbrinck et al. (2004) Genbank submission; however, for reasons discussed in Appendix 6, the SH test tends to be overconservative in that it fails to reject many invalid trees (Shimodaira \& Hasegawa, 2001).

## Remarks

Vossbrinckus richardi n. gen., n. sp. bears some resemblance to Marssoniella elegans Lemmermann, 1900 (syn. Gurleya sp. of Vávra, 1963; Gurleya marssoniella Vávra, 1968 in Komárek \& Vávra, 1968; Gurleya elegans (Lemmermann, 1900) Vávra, 1976) (see Sprague et al., 1992 for explanation of synonymy, nomenclature and recognition of Marssoniella over Gurleya for this species), with several major distinctions, which warrant its placement in a separate genus.

Marssoniella elegans Lemmermann, 1900 was examined microscopically by a number of authors (Lemmermann, 1900; Lom \& Vávra, 1963; Komárek \& Vávra, 1967; Komárek \& Vávra, 1968; Vávra \& Barker, 1977; Vávra \& Barker, 1980), whereas, the Marssoniella elegans from Genbank accession AY090041 (Table 6.1), originally submitted to Genbank as Gurleya elegans by Vossbrinck et al., 2002, was presented as only rDNA sequence not accompanied by morphological or other data describing how it was identified. The only publication mentioning this specimen replaced the name Gurleya with Marssoniella (Vossbrinck et al., 2004) and showed its molecular phylogenetic position, host, and locality, but gave no data on morphology or features used to identify it. The absence of information on how this specimen was identified is of concern because its copepod host, Cyclops vicinus, may be infected with more than one species of microsporidian at that locality (Vossbrinck et al., 2004). Furthermore, this specimen was from a different host and locality, C. vicinus in Czech Republic, than the type species M. elegans Lemmermann, 1900 (from Cyclops strenuus in Germany). Therefore, without further data, it is not clear whether M. elegans from Genbank AY099941 was correctly identified. To
prevent possible confusion from equating 2 different species, one of which may have been misidentified, I will reserve the name M. elegans Lemmermann, 1900 (or, simply M. elegans) for specimens for which morphological identification is certain, whereas I will use the name $M$. elegans sensu Vossbrinck et al. (2004) for the specimen from which only rDNA sequence is known.

## Morphological justification for $\boldsymbol{V}$. richardi n. gen., n. sp.

Vossbrinckus richardi n . gen., n. sp., is similar to M. elegans Lemmermann, 1900 (Lom \& Vávra, 1963; Komárek \& Vávra, 1967; Komárek \& Vávra, 1968; Vávra \& Barker, 1977; Vávra \& Barker, 1980) in its production of spores in clusters embedded in a mucocalyx, ranging in number from four to eight, 12 , or 16 , lacking a vesicle holding spores together, and both occurring in a cyclopoid copepod host. These species differ in some significant ways, particularly in that $V$. richardi n. gen., n. sp. produces three different types of spores. While it is possible that rarer spore forms analogous to those in V. richardin. gen. n. sp. were missed in examinations of M. elegans by Komárek \& Vávra (1968) and others, they would be more rare than a frequency of about 1 per 20 (as for ovoid, elongate to pyriform spores, respectively, in $V$. richardi n. gen., n. sp.), otherwise those authors would probably have seen them. For example, we can see the absence of non-pyriform spores in all photographic plates and accompanying text from the references above (Vávra \& Barker, 1980, plates I \& II; Lom \& Vávra, 1963, plate I; Komárek \& Vávra, plate I), and absence of any reference to distinct spore shapes or sizes in all papers cited above. Although differences in spore size of the magnitude observed here for $V$. richardi n . gen., n . sp. may have been small enough to be overlooked using light microscopy, the general shapes of all three spore forms were different from all records for M. elegans, in that the latter have long and sharply pointed (lanceolate) spores. This was clearly illustrated by photographs in Komárek \& Vávra (1968, Fig. 1, 2a, 2b, 5a, 5b). None of the three spore forms in V. richardi n. gen., n. sp. were long and pointed. The predominant pyriform spore form produced in $V$. richardi n . gen., n. sp., was also distinctly smaller, about $3.8 \pm 0.09 \mu \mathrm{~m}$ long (pyriform, fresh), compared to 5-9 $\mu \mathrm{m}$ in various populations of M. elegans (see Komárek \& Vávra, 1968).

The ability to produce several different kinds of spores per sporogony in $V$. richardi n . gen., n. sp. suggests differences in the process and timing of sporogony such that there are several sporulation "sequences" (Komárek \& Vávra, 1968). Although sporulation sequences have never been documented in M. elegans Lemmermann, 1900, the presence of one spore form,
in contrast to three forms, suggests there is only one sporulation sequence in this species. However, there are two known mechanisms for producing multiple forms of spore: (1) multiple sporulation sequences (here there would have to be three such sequences in $V$. richardi n. gen., n . sp.); and (2) production of macrospores and microspores by using a single sporulation sequence which terminates early to produce larger-sized macrospores, and later, after further nuclear divisions, to produce identically shaped but smaller-sized microspores (Cali \& Takvorian, 1999; Vávra \& Larsson, 1999). Whereas multiple sporulation sequences result in spores that are not identical in shape or other features, microspores and macrospores are invariably identical in shape (Cali \& Takvorian, 1999; Larsson, 1999; Vávra \& Larsson, 1999). Hence, mechanism (1) above probably constitutes a more significant difference between species than mechanism (2), which presumably can evolve through a small change in the signal to terminate sporulation. Therefore, the presence of spores of three distinct forms of differing shape, nuclear condition, and spore-wall features in V. richardin. gen., n. sp., is a major character difference distinguishing this species from M. elegans, and one of the strongest arguments for the requirement of a separate genus Vossbrinckus n. gen. Differences in inferred sporulation sequence, combined with numerous differences in spore morphology (shape, size and arrangement), have generally been considered more than sufficient to place species in separate genera (Sprague, 1977; Canning \& Lom, 1986; Larsson, 1988; Sprague et al., 1992).

Another significant difference between these two species is that in $V$. richardi n . gen., n . sp. spores were not released in radial or "stellate" clusters having anterior ends pointing outward and posterior ends bound together by tufts of fibers within the mucocalyx, as in M. elegans (Lom \& Vávra, 1963; Komárek \& Vávra, 1967; Komárek \& Vávra, 1968; Vávra \& Barker, 1977; Vávra \& Barker, 1980) or longitudinally arranged end-to-end (Lemmermann, 1900; Komárek \& Vávra, 1968), but instead appeared clustered in disordered masses with no special orientation of the anterior ends. Furthermore, M. elegans appears to prefer the fat body and oocytes, unlike V. richardi n. gen., n. sp., which was always found in the haemocoel. However, Komárek \& Vávra (1968) observed that in rare, heavy cases, M. elegans infections, which usually began in the oocytes, could later fill the haemocoel. Since fat body cells and oocytes were not systematically examined in this study, and most infections were heavy, it is possible that spores of $V$. richardi n. gen., n. sp. could occur in these locations as well.

The geographic localities of these two species differ significantly. Whereas V. richardi n. gen., n. sp. was found in northwestern North America, the type locality of M. elegans was Lake Summt-See, near Berlin, Germany (Lemmermann, 1900). Others report M. elegans from
various parts of central Europe, Russia, and northeastern USA (Komárek \& Vávra, 1968). This geographic range seems unusually broad for a single species parasitic in freshwater copepods that exist in small, isolated ponds, and might have resulted from mis-identification of species. Indeed, taxonomists probably underestimated the number of species using only light microscopy (Larsson, 1996; 1999).

In summary, morphological data suggest $V$. richardi n. gen., n. sp. resembles M. elegans morphologically, in having spores arranged in clusters attached within a distinctive mucocalyx, which can be seen as "streaks" (Vávra \& Barker, 1977) between spores, and in its use of a cyclopoid copepod host; however, V. richardi n. gen., n. sp. does not belong in the genus Marssoniella Lemmermann, 1900, for several reasons, most importantly because of differences in sporulation sequence and the presence of three spore forms, none of which form stellate clusters in $V$. richardi n. gen., n. sp., as discussed earlier.

Besides Marssoniella Lemmermann, 1900, morphological similarity warrants consideration of two other possible genera for this new species: Hazardia Weiser, 1977, and Gurleya Doflein, 1898. Hazardia species undergo three sporulation sequences resulting in the most common spores being pyriform and thin-walled and contained in a sporophorous vesicle; however, Hazardia species are found exclusively in mosquito larvae. Gurleya species parasitize aquatic Crustacea, but are primarily found in the hypodermis of Cladocera (Daphnia spp.), and have only one known sporulation sequence (Sprague et al., 1992; Becnel \& Andreadis, 1999). Vossbrinckus richardi n. gen., n. sp. differs from Hazardia and Gurleya in that it has three sporulation sequences all in the haemocoel of a cyclopoid copepod, each producing variable numbers of spores. Vossbrinckus richardi n. gen., n. sp. undergoes all three sequences in the copepod, and so presumably does not require mosquito larvae or adults in its life-cycle, whereas the type species H. milleri (Hazard \& Fukuda, 1974) Weiser, 1977, has been shown, experimentally, to undergo all sequences exclusively in mosquitoes (Becnel \& Andreadis, 1999). These species also differ in that V. richardi n. gen., n. sp., unlike Hazardia spp., has no sporophorous vesicle as far as could be discerned, and the most common spores have thicker walls, and instead, produce a mucocalyx to hold spores together. Similarly, by definition Gurleya species (type species G. tetraspora) always have one sporulation sequence that produces four spores (Sprague, 1977). However, Becnel \& Andreadis (1999) and Sprague (1977) suggested this sole defining character of genus Gurleya - one sporulation sequence producing four spores - is too broadly inclusive. It is unclear whether there is a sporophorous
vesicle in the type species, G. tetraspora (Sprague, 1977); thus, Gurleya is not a suitable group for several reasons.

## Molecular and phylogenetic justification for V. richardi n. gen., n. sp.

One question to be addressed first was whether the less-common spore forms in the new species could be from a co-infection with more than one species. This is particularly important given the observation that a single copepod species can be the host for several species of microsporidian (Vossbrinck et al., 2004). To test for possible co-infection in these copepods DNA was obtained from mixtures containing all three spore forms, then many PCR products and clones were sequenced from this DNA. The result from three separate genetic loci, including: SSU, ITS and LSU rDNA; $\alpha$-tubulin; and $\beta$ - tubulin (the latter 2 loci were amplified and sequenced by Dr. P. J. Keeling, personal communication) was a consistently low ( $\sim 0.1 \%$ ) intra-clone difference, strongly suggesting these spores represent a single species.

Molecular results were consistent with morphological data, showing that $V$. richardi n . gen., n. sp. differs from species of Marssoniella, Hazardia, Gurleya and other genera at a high level (at least 5.9 to 12.2 \% SSU rDNA difference), but is most similar in rDNA sequence to $M$. elegans sensu Vossbrinck et al. (2004) Genbank submission AY090041. Vossbrinckus richardi n. gen., n. sp. and M. elegans sensu Vossbrinck et al. (2004) clustered together, with high support, in all analyses (Figs. 6.24-6.26), suggesting these are close sister-genera. These two species were the only ones from clade A from copepods (see Fig. 6.26). This shared difference in host, compared to others in the clade presumably helped the authors of the Genbank submission to re-consider the genus placement of Gurleya elegans (Vossbrinck et al., 2002, Genbank submission), changing it to Marssoniella (Vossbrinck et al., 2004). Phylogenetic results here confirmed that Gurleya is inappropriate for M. elegans, as was suggested previously (Sprague et al., 1992).

Phylogenetic results and AU tests strongly suggested $V$. richardi n. gen., n. sp. together with M. elegans sensu Vossbrinck et al. (2004) form a sister-group to two Hazardia species, one of these being the type species H. milleri (Hazard \& Fukuda, 1974) Weiser, 1977 (Figs. 6.24 6.26, 6.27 and Table 6.2). This phylogenetic result is consistent with the morphological similarities of these genera. The \% SSU rDNA difference between $V$. richardi n. gen., n. sp. and H. milleri was a little higher than that between $V$. richardi n . gen., n . sp . and M. elegans sensu Vossbrinck et al. (2004), these being 8.6 \% vs. 5.9 \% difference, respectively; however, to test the relative molecular similarity of these pairs of species, H. milleri and V. richardi n. gen., n. sp. were forced together as close sister-taxa (Fig. 6.27, tree \#3), and the result was rejection of
this tree with high confidence in AU and other tests. This result suggests $V$. richardin. gen., n. sp. and M. elegans sensu Vossbrinck et al. (2004) are indeed closest-sisters among these taxa, despite dissimilarities in sporulation sequence, spores and other morphological features. As discussed earlier, sporulation sequence appears to be more similar between $H$. milleri and $V$. richardi n . gen., n. sp. than between M. elegans and V. richardin. gen., n. sp. Phylogenetic results, therefore suggest sporulation sequence may be a less important character than host at the genus- and higher levels. This finding is consistent with other studies that suggest general host group (e.g. host family) is an important indicator of relatedness in microsporidia (Baker et al. 1997; Vossbrinck et al. 2004), although microsporidia do not strictly or exclusively co-speciate with their hosts (see host groups Fig. 6.26; and host groups in Chapter 5 Fig. 5.9; Chapter 4 Fig. 4.20). These data suggest that perhaps the tendency of microsporidia from cyclopoid copepods to group together may result less from co-speciation, and more from co-adaptation (or convergence) to host-habitat or ecological role.

Vossbrinckus richardi n. gen., n. sp. and M. elegans sensu Vossbrinck (2004) did not cluster closely with two other Gurleya species, G. daphniae and G. vavrai, that resemble the type species of genus Gurleya in their host (Cladocera species) and morphology. When forced together (Figs. 6.27, trees \#1 and 5), clades containing V. richardi n. gen., n. sp. and Gurleya species were not supported, confirming morphological results suggesting $V$. richardin. gen., n. sp. and M. elegans sensu Vossbrinck et al. (2004) are closer to Hazardia than Gurleya species.

## Justification for placement in family Golbergiidae 1ssi, 1986

Phylogenetic and morphological similarities between $V$. richardi n. gen., n. sp. and $H$. milleri (Hazard \& Fukuda, 1974) Weiser, 1977 suggest placement of V. richardi n. gen., n. sp. into the family Golbergiidae Issi, 1986, into which Issi (1986) placed genus Hazardia. For example, H. milleri and V. richardi n. gen., n. sp. share features such as three spore forms, the most common of these being uninucleate and pyriform, and another more rare form of these being thicker-walled, oval and binucleate, although they differ in important ways that distinguish the genera (H. milleri is found exclusively in mosquitoes, producing a vesicle around spores, and having differences in spore morphology). Family Golbergiidae was defined mainly for features of $H$. milleri; however, Sprague et al. (1992) observed that some of these features really only apply to Hazardia. The present study suggested inclusion of $V$. richardi n. gen., n. sp. in the family Golbergiidae because it displays most of the family features, including: concurrent sporulation sequences with the dominant one resulting in uninucleate spores (for $V$.
richardi n . gen., n. sp. these would be pyriform spores) in clusters of various sizes, one resulting in binucleate spores (here the shorter, ovoid spores).

Phylogenetic evidence presented here also suggested the family Golbergiidae Issi, 1986 should be expanded to include M. elegans sensu Vossbrinck et al. (2004), or M. elegans Lemmermann, 1900 (provided the species was correctly identified). In future, researchers should re-examine features of all members of the family Golbergiidae (Hazardia, Golbergia, and Vossbrinckus n. gen.), and also M. elegans sensu Vossbrinck et al. (2004) or M. elegans Lemmermann, 1900, to determine to what extent diagnostic features of this family require revision. Family Gurleyidae Sprague, 1977, which was considered to be too broadly defined and therefore listed as "taxa incertae sedis" by Sprague et al. (1992) was used by authors of Genbank submissions for Hazardia and Marssoniella (Table 6.1); however, family Gurleyidae, defined as "four spores usually formed within the pansporoblastic membrane" (Sprague, 1977), is here considered unsuitable for these genera.

## Life-cycle evolution in the "Aquatic Outgroup"

Although the life-cycle of $V$. richardin. gen., n. sp. was not directly studied, several features of the life-cycle can be inferred from its morphology and phylogenetic place. Vossbrinckus richardi n . gen., n. sp. was surrounded phylogenetically by species known or suspected to have two- or three-host (heteroxenous) life-cycles. For example, most species in the Amblyospora group, a sister group to the "Aquatic Outgroup" (and clade A, Figs. 6.246.26 ) in which $V$. richardi n. gen., n. sp. fell, require multiple hosts. Parathelohania anophelis, a species that branches basally to both the Aquatic Outgroup and Amblyospora group in most analyses (MP and ML in this study; and Refardt et al., 2002), requires three hosts. Furthermore, three species within the Aquatic Outgroup are suspected to require further hosts, since they cannot be transmitted to their original hosts in laboratory: Gurleya vavrai and Larssonia obtusa from Daphnia spp. (Ebert, unpublished cited in Refardt et al., 2002), and Vairimorpha sp. from the black fire ant, Solenopsis richteri (Moser, personal communication cited in Refardt et al., 2002). No data are available to determine the likelihood of a second host in Amblyospora sp. CHE from Simulium sp., or M. elegans sensu Vossbrinck et al. (2004). The only relative of $V$. richardi n . gen., $\mathrm{n} . \mathrm{sp}$. that has been demonstrated to require only one host species (monoxenous) is Hazardia milleri. Hence, these phylogenetic data suggest multiple-host, heteroxenous life-cycles are more primitive than monoxenous life-cycles in the AmblyosporalParathelohania/Aquatic Outgroup clades (Baker et al., 1997; Vossbrinck et al.,
2004), although hosts can be easily lost. For example, a copepod was almost certainly lost in some species from mosquitoes Culicospora magna, Edhazardia aedis, and Hazardia milleri, (Becnel, 1994; Becnel \& Andreadis, 1999; Vossbrinck et al., 2004). Was the mosquito host lost in species from copepods or other aquatic Crustacea in the Aquatic Outgroup clade, or is it more likely that a definitive host has yet to be found?

The morphology of $V$. richardin. gen., n. sp. is unusual for species from the Aquatic Outgroup, in that features seem to point to present or past use of a second host. For example, the presence of a mucocalyx suggests these spores may float and endure for a long time until a more seasonally ephemeral, water surface-dwelling host (e.g. adult or larval mosquito) is available. By contrast, it is not as clear how to interpret the presence of three spore forms. In fact, $V$. richardi n. gen., n. sp. was the first microsporidian from a copepod found to produce three forms of spore in a single host. Knowing the phylogenetic proximity of V. richardin. gen., n . sp. to species that undergo more than one sporulation sequence to produce different spore forms (Amblyospora-like species and Hazardia species) destined for different functions or hosts, one can speculate about the possible functions of these spores. For example, usually two spore forms at most are produced per host, one specialized for autoinfection and the other specialized for the external environment. Where a third or even fourth spore form is found (e.g. four exist in Edhazardia), species either use two hosts (e.g. usually a mosquito and a copepod) or multiple generations of one host (e.g. parent and offspring mosquito). Although V. richardi n. gen., n. sp. is unusual in producing three spore forms in one host, Hazardia species, from larval mosquitoes, seem to also produce three spore forms within the same host. The complexity in Hazardia spp. is a remnant of an ancestrally complex life-cycle rather than evidence of a missing (not yet discovered) host. Future studies should critically examine hostuse in V. richardi n. gen., n. sp. and other species in the Hazardia/Vossbrinckus clade (clade H), as these species may be close relatives that have lost opposite hosts.

## Usefulness and design of a molecular-ecological PCR-probe

Studies of geographic- and host-range of a species normally precede studies that develop a species-specific molecular probe (e.g. a PCR-probe); however, for $V$. richardi n. gen., n. sp., geographic- and host-range studies might benefit from first having a species-specific probe available, to help detect an alternate host or the destiny of spores in the environment. A preliminary molecular probe is particularly essential because $V$. richardi n. gen., n. sp. produces several forms of spore that, without a probe, could be mistaken for a different species.

Furthermore, a probe would help if some copepods are infected with multiple, perhaps cryptic, species of microsporidia, as has been suggested (Vossbrinck et al., 2004).

By visually inspecting alignments of rDNA from large numbers of species, including $V$. richardi n . gen., n. sp. and its closest relatives, variable regions were identified that can be used in designing a molecular (PCR) probe that should preferentially amplify this species. Normally the ITS is a good target region for designing a species-specific molecular probe; however in $V$. richardin. gen., n . sp. the ITS appeared to be too short. Although the exact boundaries of the ITS were difficult to discern, the best guess of the ITS boundaries here, based on alignments using secondary structure, suggest the ITS of $V$. richardin. gen., n. sp. is only 26 bp long. This short length suggests the ITS is at its minimum functional length, and therefore may be highly constrained and invariant within, and perhaps even between, closely related species (as was the case for Loma salmonae; see Chapter 2). In contrast, the SSU and LSU appeared to show normal interspersed regions of high variability. Therefore, two regions were identified that might be ideal for distinguishing $V$. richardi n. gen., n. sp. in future studies: positions 112-118 in the SSU and 1773-1781 in the LSU. An intraspecific indel found in a highly variable region at position 1641 in the LSU suggested this region is not suitable.

## CONCLUSION

Data presented here suggested that the new genus Vossbrinckus n. gen., and type species Vossbrinckus richardi n . gen., n . sp. was distinctly different from other available genera and species of microsporidia, both based on morphological characters of development and spores, and based on the host, geographic locality, and rDNA sequence. More data are needed on the early development and ecology of this species. These results suggested $V$. richardi n. gen., n. sp. had sufficient morphological and rDNA similarity to be placed in the family Golbergiidae Issi, 1986. Results also suggested M. elegans sensu Vossbrinck et al. (2004) Genbank submission AY090041 should be placed in the same family, based on rDNA evidence. Morphological data were unavailable for M. elegans sensu Vossbrinck et al. (2004), and will be needed to determine whether it was correctly identified as M. elegans Lemmermann, 1900, or whether it was identified incorrectly, or perhaps resembles the new genus Vossbrinckus. The life-cycle of $V$. richardi n . gen., n. sp. is unknown, but its close phylogenetic relationship with Hazardia species, together with production of three spore forms and the presence of a
mucocalyx suggested this species may have either ancestrally used or currently requires a secondary host.

## FURTHER INVESTIGATION

Many questions remain to be addressed regarding the life-cycle of this parasite. For example, what is its host breadth and does it use a second dipteran (e.g. mosquito) host? Ribosomal DNA sequences obtained here should be promising in future development of a specific PCR-probe to answer these questions. Such a probe could also help studies of early developmental morphology to find cryptic stages (e.g. meronts) in the copepod or other hosts. Studies suggest microsporidians from aquatic crustaceans commonly undergo vertical transmission and cause sterilization or castration of the host (Refardt et al., 2002; Dunn \& Smith, 2001; Dunn et al., 2001). Others also suggest crustacean-parasitic microsporidia can dramatically affect host population dynamics (Becnel et al., 1995; Andreadis, 2002; Micieli et al, 2001; Andreadis, 1999). These phenomena should be examined in $V$. richardi n. gen., n. sp. This species' close relative, M. elegans Lemmermann, 1900 may help to indicate directions for further study. For example, Komárek \& Vávra (1968) suggested rare (6-8\%) instances of spores found in groups of more than four were the result of earlier division during sporogony, releasing immature spores to float via the mucocalyx up into the plankton, where they remain viable for over one year (Lom \& Vávra, 1963; Komárek \& Vávra, 1968; Vávra \& Barker, 1980). Immature spores of $V$. richardi n. gen., n. sp., covered by thick (possibly mucous) coats, apparently ready to be released (Figs. 6.12, 6.13), were observed, and may be similar in function to those observed in M. elegans, or may be produced for infection of alternate hosts. Malformed or abortive spores (Figs. 6.22, 6.23) may be indicative of an abortive meiosis as is seen in some microsporidia (Becnel \& Andreadis, 1999), though further data are needed to confirm this.

## ACKNOWLEDGEMENTS

The Natural Sciences and Engineering Research Council of Canada strategic grant 582073 to M. L. Adamson supported this work. I am grateful to Dr. P. J. Keeling and his laboratory for their assistance with cloning, and to Elaine Humphrey and technicians at the BioImaging Facility, UBC for their help with TEM and Michael Coury for help with software
design. I gratefully acknowledge Dr. Charles S. Vossbrinck for the idea to sample roadside ditches and I thank Charles Vossbrinck and Bettina Debrunner-Vossbrinck for their generosity and hospitality on the visit that motivated this work, and for inspiration to persevere when the work was mired in lost-DNA and inhibitors.

## LITERATURE CITED

Andreadis, T. G. 1985. Experimental transmission of a microsporidian pathogen from mosquitoes to an alternate copepod host. Proceedings of the National Academy of Sciences US.A 82:5574-5577.

Andreadis, T. G. 1983. Life cycle and epizoology of Amblyospora sp. (Microspora: Amblyosporidae) in the mosquito, Aedes cantator. Journal of Protozoology 30(3):509-518.

Andreadis, T. G. 1999. Epizootiology of Amblyospora stimuli (Microsporidia: Amblyosporidae) infections in field populations of a univoltine mosquito, Aedes stimulans (Diptera: Culicidae), inhabiting a temporary vernal pool. Journal of Invertebrate Pathology 74(2):198-205.

Andreadis, T. G. 2002. Epizootiology of Hyalinocysta chapmani (Microsporidia: Thelohaniidae) infections in field populations of Culiseta melanura (Diptera: Culicidae) and Orthocyclops modestus (Copepoda: Cyclopidae): a three-year investigation. Journal of Invertebrate Pathology 81(2):114-21.

Andreadis, T. G. and Vossbrinck C. R. 2002. Life cycle, ultrastructure and molecular phylogeny of Hyalinocysta chapmani (Microsporidia: Thelohaniidae), a parasite of Culiseta melanura (Diptera: Culicidae) and Orthocyclops modestus (Copepoda: Cyclopidae). Journal of Eukaryotic Microbiology 49(4):350-64.

Andreadis, T. G. 1988a. Amblyospora connecticus sp. nov. (Microsporida: Amblyosporidae): horizontal transmission studies in the mosquito Aedes cantator and formal description. Journal of Invertebrate Pathology 53:90-101.

Andreadis, T. G. 1988b. Comparative susceptibility of the copepod Acanthocyclops vernalis to a microsporidian parasite, Amblyospora connecticus, from the mosquito Aedes cantator. Journal of Invertebrate Pathology 52:73-77.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Maddox, J. V. 1997. Phylogenetic position of Amblyospora Hazard \& Oldacre (Microspora: Amblyosporidae) based on small subunit rRNA data and its implication for the evolution of the microsporidia. Journal of Eukaryotic Microbiology 44(3):220-225.

Becnel, J. J. and Andreadis, T. G. 1998. Amblyospora salinaria n. sp. (Microsporidia: Amblyosporidae), parasite of Culex salinarius (Diptera: Culicidae): its life cycle stages in an intermediate host. Journal of Invertebrate Pathology 71:258-262.

Becnel, J. J. and Andreadis, T. G. 1999. Microsporidia in insects. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 447501.

Becnel, J. J., Garcia, J. J. and Johnson, M. A. 1995. Edhazardia aedis (Microspora: Culicosporidae) effects on the reproductive capacity of Aedes aegypti (Diptera: Culicidae). Journal of Medical Entomology 32(4):549-553.

Becnel, J. J. 1994. Life cycles and host-parasite relationships of Microsporidia in culicine mosquitoes. Folia Parasitologica 41:91-96.

Cali, A. and Takvorian, P. M. 1999. Developmental morphology and life cycles of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 85-128.

Canning, E. U. and Lom, J. 1986. The Microsporidia of Vertebrates. Academic Press, London. 289 pp.
Darwish, A. and Canning, E. U. 1991. Amblyospora sp. (Microspora, Amblyosporidae) infecting nerve ganglia of Culex pipiens (Diptera, Culicidae) from Egypt. Journal of Invertebrate Pathology 58:244-251.

Dunn, A. M. and Smith, J. E. 2001. Microsporidian life cycles and diversity: the relationship between virulence and transmission. Microbes and Infection 3:381-388.

Dunn, A. M., Terry, R. S. and Smith, J. E. 2001. Transovarial transmission in the microsporidia. Advances in Parasitology 48:57-100.

Edlind, T. D., Li, J., Visvesvara, G. S., Vodkin, M. H., McLaughlin, G. L. and Katiyar, S. K. 1996. Phylogenetic analysis of beta-tubulin sequences from amitochondrial protozoa. Molecular Phylogenetics \& Evolution 5:539-367.

Fries, I., Paxton, R. J., Tengö, J., Slemenda, S. B., da Silva, A. J. and Pieniazek, N. J. 1999. Morphological and molecular characterization of Antonospora scoticae n. gen., n. sp. (Protozoa, Microsporidia) a parasite of the communal bee, Andrena scotica Perkins, 1916 (Hymenoptera, Andrenidae). European Journal of Protozoology 35:183-193.

Germot, A., Philippe, H. and Le Guyader, H. 1997. Evidence for loss of mitochondria in microsporidia from a mitochondrial-type HSP70 in Nosema locustae. Molecular and Biochemical Parasitology 87:159-68.

Goldman, N., Anderson, J. P. and Rodrigo, A. G. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49(4):652-70.

Hazard, E. I. and Fukuda, T. 1974. Stempellia milleri sp. n. (Microsporida: Nosematidae) in the mosquito Culex pipiens quinquifasciatus Say. Journal of Protozoology 21:497-504.

Hirt, R. P., Healy, B., Vossbrinck, C. R., Canning, E. U. and Embley, T. M. 1997. Identification of a mitochondrial Hsp70 orthologue in Vairimorpha necatrix: molecular evidence that microsporidia once contained mitochondria. Current Biology 7:995-998.

Hirt, R. P., Logsdon, J. M., Healy, B., Dorey, M. W., Doolittle, W. F. and Embley, T. M. 1999. Microsporidia are related to fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proceedings of the National Academy of Sciences USA 96:580-585.

Issi, I. V. 1986. Microsporidia as a phylum of parasitic protozoa, p. 6-136 In Beyer, T. V. and Issi, I. V. (eds) Protozoology, vol. 10. Nauka, Leningrad, USSR. [in Russian with English summary]

Keeling, P. J. and Doolittle, W. F. 1996. Alpha-tubulin from early-diverging eukaryotic lineages and the evolution of the tubulin family. Molecular Biology \& Evolution 13:1297-305.

Keeling, P. J., Luker, M. A. and Palmer, J. D. 2000. Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. Molecular Biology \& Evolution 17:23-31.

Kishino, H. and Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29:170-179.

Komárek, J. and Vávra, J. 1967. Identity of a blue-green alga Marssoniella elegans with a microsporidian Gurleya sp. Journal of Protozoology 14(Suppl):36.

Komárek, J. and Vávra, J. 1968. In memoriam of Marssoniella Lemm, 1900. Archiv fuer Protistenkunde 111(1):1217.

Larsson, J. I. R. 1988. Identification of microsporidian genera (Protozoa, Microspora) a guide with comments on the taxonomy. Archiv fuer Protistenkunde 136:1-37.

Larsson, J. I. R. 1996. Microsporidia in microcrustaceans. Scandinavian Section Society of Protozoologists 18th Annual Meeting. 202.

Larsson, J. I. R. 1999. Identification of microsporidia. Acta Protozoologica 38:161-197.
Lemmermann, E. 1900. Beiträge zur Kenntnis der Planktonalgen. Berichte der Deutschen Botanischen Gesellschaft 18:272-275.

Lom, J. 2002. A catalogue of described genera and species of microsporidians parasitic in fish. Systematic Parasitology 53(2):81-99.

Lom, J. and Nilsen F. 2003. Fish microsporidia: fine structural diversity and phylogeny. International Journal for Parasitology 33(2):107-27.

Lom, J. and Vávra, J. 1963. Mucous envelopes of spores of the subphylum Cnidospora (Doflein, 1901). Vestnik Ceskoslovenske Spolecnosti Zoologicke 27:4-6.

Lucarotti, C. J. and Andreadis, T. G. 1995. Reproductive strategies and adaptations for survival among obligatory microsporidian and fungal parasites of mosquitoes: a comparative analysis of Coelomomyces and Amblyospora. Journal of the American Mosquito Control Association 11:111-121.

Micieli, M. V., García, J. J. and Becnel, J. J. 2000. Horizontal transmission of Amblyospora albiafasciati García and Becnel, 1994 (Microsporidia: Amblyosporidae), to a copepod intermediate host and the neotropical mosquito, Aedes albifasciatus (Macquart, 1837). Journal of Invertebrate Pathology 75:76-83.

Moodie, E. G., Le Jambre, L. F. and Katz, M. E. 2003. Thelohania montirivulorum sp. nov. (Microspora: Thelohaniidae), a parasite of the Australian freshwater crayfish, Cherax destructor (Decapoda: Parastacidae): fine ultrastructure, molecular characteristics and phylogenetic relationships. Parasitology Research 91:215-228.

Nilsen, F. and Chen, W. J. 2001. rDNA phylogeny of Intrapredatorus barri (Microsporida: Amblyosporidae) parasitic to Culex fuscanus Wiedemann (Diptera: Culicidae). Parasitology 122:617-623.

Page, R. D. 1996. TreeView: an application to display phylogenetic trees on personal computers. Computer Applications in Biosciences 12:357-358.

Posada, D. and Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14(9):817-818.

Van de Peer, Y., De Rijik, R., Wuyts, J., Winkelmans, T. and De Wachter, R. 2000. The European small subunit ribosomal RNA database. Nucleic Acids Research 28:175-176.

Refardt, D., Canning, E. U., Mathis, A., Cheney, S. A., Lafranchi-Tristem, N. J. and Ebert, D. 2002. Small subunit ribosomal DNA phylogeny of microsporidia that infect Daphnia (Crustacea: Cladocera). Parasitology 124:381-389.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology 51(3):492508.

Shimodaira H. and Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17(12):1246-7.

Sprague, V. 1977. Annoted list of species of Microsporidia. In Comparative Pathology. 2. Systematics of the Microsporidia. Plenum Press, New York, New York, 333 p.

Sprague, V., Becnel, J. J. and Hazard, E. I. 1992. Taxonomy of phylum Microspora. Critical Reviews in Microbiology 18(5/6):285-395.

Sweeney, A. W., Doggett, S. L. and Piper, R. G. 1990. Host specificity studies of Amblyospora indicola and Amblyospora dyxenoides (Microspora: Amblyosporidae) in mosquitoes and copepods. Journal of Invertebrate Pathology 56:415-418.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.

Vávra, J. and Barker, R. J. 1977. The demonstration of the microsporidian mucocalyx in the scanning electron microscope (SEM). Journal of Protozoology 24(4):54A.

Vávra, J. and Barker, R. J. 1980. The microsporidian mucocalyx as seen in the scanning electron microscope. Folia Parasitologica 27:19-21.

Vávra, J. and Larsson, J. I. R. 1999. Structure of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 7-84.

Vossbrinck, C. R., Maddox, J. V., Friedman, S., Debrunner-Vossbrinck, B. A. and Woese, C. R. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326(6111):411-414.

Vossbrinck, C. R., Andreadis, T. G. and Debrunner-Vossbrinck, B. A. 1998. Verification of intermediate hosts in the life cycles of microsporidia by small subunit rDNA sequencing. Journal of Eukaryotic Microbiology 45:290-292.

Vossbrinck, C. R., Andreadis, T. G., Vávra, J. and Becnel, J. J. 2004. Molecular phylogeny and evolution of mosquito parasitic microsporidia (Microsporidia: Amblyosporidae). Journal of Eukaryotic Microbiology 51(1):88-95.

Weiss, L. M. and Vossbrinck, C. R. 1999. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In The microsporidia and microsporidiosis. Wittner M., and Weiss L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 129-171.

Table 6.1: Microsporidian rDNA sequences included in phylogenetic analysis with Genbank accession numbers, current family designation, and host name. $\mathrm{F}=$ fish, $\mathrm{I}=$ insect, $\mathrm{C}=$ Crustacea, $\mathrm{B}=$ Bryozoa. * Marssoniella elegans was originally submitted to Genbank under the name Gurleya elegans by Vossbrinck et al. (2002); however, Vossbrinck et al. (2004) later referred to this species as M. elegans, presumably based on the arguments in Sprague et al. (1992).

| Species | Accession | Family | Host |
| :--- | :--- | :--- | :--- |
| Amblyospora californica | U 68473 | Amblyosporidae | I: Culex tarsalis |
| Amblyospora connecticus | AF025685 | Amblyosporidae | I: Aedes cantator |
| Amblyospora sp. BAK | U 68474 | Amblyosporidae | I: Culex salinarius |
| Amblyospora sp. CHE | AJ252949 | Amblyosporidae | I: Simulium sp. |
| Amblyospora stimuli | AF027685 | Amblyosporidae | I: Aedes stimulans |
| Antonospora scoticae | AF024655 | not placed | I: Antonospora scoticae |
| Berwaldia schaefernai | AY090042 | Amblyosporidae | C: Daphnia galeta |
| Bryonosema plumatellae | AF484692 | Pseudonosematidae | B: Plumatella nitens |
| Coelomomyces stegomyiae (Fungi) | AF322406 | Coelomomycetaceae | N/A |
| Conidiobolus coronatus (Fungi) | AF296753 | Ancylistaceae | N/A |
| Culicosporella lunata | AF027683 | Caudosporidae | I: Culex pilosus |
| Edhazardia aedis | AF027684 | not placed | I: Aedes aegypti |
| Flabelliforma magnivora | AJ302318 | Dubosqiidae | C: Daphnia magna |
| Flabelliforma montana | AJ252962 | Dubosqiidae | I: Phlebotomus ariasi |
| Gurleya daphniae | AF439320 | Gurleyidae | C: Cladocera |
| Gurleya vavrai | AF394526 | Gurleyidae | C: Cladocera |
| Hazardia milleri | AY090067 | Gurleyidae | I: Culex quinquefasciatus |
| Hazardia sp. | AY090066 | Gurleyidae | I: Anopheles crucians |
| Hyalinocysta chapmani | AF483837 | Thelohaniidae | I: Culiseta melanura |
| Intrapredatorus barri | AY013359 | Amblyosporidae | I: Culex fuscanus |
| Larssonia obtusa | AF394527 | Thelohaniidae | C: Cladocera |
| Marssoniella* elegans | AY090041 | Gurleyidae | C: Cyclops vicinus |
| Ordospora colligata | AF394529 | Ordosporidae | C: Cladocera |
| Parathelohania anophelis | AF027682 | Amblyosporidae | I: Anopheles quadrimaculatus |
| Polydispyrenia simulii | AJ252960 | not placed | I: Simulium sp. |
| Pseudonosema cristatellae | AF484694 | Pseudonosematidae | B: Cristatella mucedo |
| Schroedera plumatellae | AY135024 | Mrazekiidae | B: Plumatella fungosa |
| Trichonosema pectinatellae | AF484695 | Pseudonosematidae | B: Pectinatella magnifica |
| Vairimorpha sp. | AF031539 | Burenellidae | I: Solenopsis richteri |
| Weiseria palustris | AF132544 | Caudosporidae | I: Cnephia ornithophilia |
|  |  |  |  |



Figures 6.1-6.5: Light micrographs of spores from toluidine-stained, thick resin sections. Figs. 6.1 and 6.2 show location of spores in small, loose clusters (solid white arrow) or large, tight clusters (hollow white arrow) in haemocoel of copepod. Figs. 6.2-6.5 show spores of 3 forms, elongate (hollow black arrow), pyriform (solid black arrow), and ovoid (solid black arrowhead), spores caught in mucous (large white arrow). Scale bars $=10 \mu \mathrm{~m}$.


Figures 6.6-6.9: Transmission electron micrographs showing sporulation stages (with nuclei N). Fig. 6.6 shows a round stage containing a round, dark structure that may be an early polar capsule primordium (white arrow). Fig. 6.7 shows a sporont or early sporoblast with electrondense surface coat and cytoplasm with a low level of vacuolation. Fig. 6.8 shows a sporont (shape deformed due to fixation) with greater vacuolation drawn out in a stream of mucous (m) with a dark structure resembling that seen in other stages, perhaps an early polar capsule primordium (white arrow). Fig. 6.9 shows an early sporoblast (deformed due to fixation), with early signs of polar filament formation (hollow white arrows) and the beginnings of a thin electron-lucent endospore (black arrow). Scale bars $=1 \mu \mathrm{~m}$.


Figure 6.10-6.13: Transmission electron micrographs showing sporulation stages (nuclei N ). Fig. 6.10 shows a mature spore in transverse section next to an early sporoblast (deformed due to fixation) with early polar filament formation and a thin electron-lucent endospore, both stages being embedded in a stream of mucous (m). Fig. 6.11 shows a sporoblast in transverse section with a thick, multi-layered coat, embedded in a stream of mucous (m). Fig. 6.12 shows another sporoblast of the elongate form, having a thick, multi-layered coat (black arrow), single nucleus, various cytoplasmic membranes, with a dark round capsule of unknown function (white arrow). Fig. 6.13 shows a sporoblast of the pyriform type with early signs of polar filament formation (hollow white arrow) and a crescentric nucleus (CN). Scale bar $=1 \mu \mathrm{~m}$.


Figure 6.14-6.17: Transmission electron micrographs of mature pyriform (Figs. 6.14, 6.16 \& 6.17 ) and elongate (Fig. 6.15) spores. Figs $6.14,6.16$ ( \& 6.17 close up of 6.16 ) show pyriform spores with 6 turns of the isofilar polar filament having about 6-8 layers of different electron density, a distinctly rough, undulating exospore (ex) of several thin layers, moderately thick endospore (en), a compressed nucleus ( N ), posterior vacuole ( pv ) that fills about $1 / 3$ of the spore, spore cytoplasm is packed with chains of polyribosomes (ri). Fig. 6.15 shows an elongate spore, with a round dark structure of unknown function in the middle (black arrow), distinct lamellar and vesicular polaroplasts (white arrows), and polar filament isofilar (seen in graze here) with at least 4 turns, and thick exospore that is thinner at anterior end (hollow white arrow). Scale bar $=1 \mu \mathrm{~m}$.


Figure 6.18-6.21: Fig. 6.18 Four spores caught in a stream of mucous (arrow), beside one thick-walled solitary sporoblast (sb). Scale bar $=10 \mu \mathrm{~m}$. Figs. 6.19-6.21 Details of mature spores (nuclei N). Fig. 6.19 shows anterior end of an elongate spore with polar capsule (pc), laminar polaroplasts (white arrows), and rugose exospore in graze. Fig. 6.20 shows an ovoid spore in slightly off-transverse section showing a large, obvious nucleus, abundant vesicular polaroplast (v) and smooth exospore (black arrow). Fig. 6.21 shows another ovoid spore with 2 separated nuclei. Scale bars $=1 \mu \mathrm{~m}$.


Figure 6.22-6.23: Transmission electron micrographs of malformed or abortive spores. Fig. 22 shows polar filament coil (7 turns) in non-linear arrangement (white arrows), and an unusually small or not completely formed posterior vacuole (pv). Fig. 6.23 shows polar filament coils (white arrows) at an unusually steep angle from presumed perpendicular axis (shown with black arrow pointing towards presumed axis). Scale bar $=1 \mu \mathrm{~m}$.


10
Figure 6.24: Maximum parsimony tree (one of 2 similar most parsimonious trees, which do not vary in labeled clades) from heuristic search in PAUP* showing position of Vossbrinckus richardi n. gen., n. sp. using 1022 alignment positions of SSU rDNA. Bootstrap values $>50 \%$ ( 1000 replicates) are indicated on nodes in common with $50 \%$ majority rule consensus tree. Note bootstrap values in bold leading to V. richardi n. gen., n. sp. and relative support for clades A, H and G. * Marssoniella elegans sensu Vossbrinck et al. (2004).


Figure 6.25: Distance tree (logDet/paralinear) from heuristic search in PAUP* showing position of Vossbrinckus richardi n. gen., n. sp. using 1022 alignment positions of SSU rDNA. Bootstrap values $>50 \%$ ( 1000 replicates) are indicated on nodes in common with $50 \%$ majority rule consensus tree. Note bootstrap values in bold leading to $V$. richardi n. gen., n. sp. and relative support for clades A, H and G. * Marssoniella elegans sensu Vossbrinck et al. (2004).


Figure 6.26: Maximum likelihood tree from heuristic search in PAUP* (GTR+I+G model parameters estimated using Modeltest) using 1022 alignment positions of SSU rDNA. Bootstrap values $>50 \%$ ( 100 replicates) are indicated on nodes in common with $50 \%$ majority rule consensus tree. * Marssoniella elegans sensu Vossbrinck et al. (2004). Hosts and hostbased groups are shown in gray typeface ( $\mathrm{C}=$ Copepod-parasitic; $\mathrm{M}=$ Mosquito-parasitic; $\mathrm{D}=$ Cladocera-parasitic; A = Ant-parasitic; $\mathrm{S}=$ Simulium spp. parasitic).

Table 6.2: Comparison of unconstrained and monophyly-constrained trees from 1022 alignment positions of SSU rDNA, using the Approximately Unbiased or "AU", weighted and unweighted Kishino-Hasegawa or "KH" and Shimodaira-Hasegawa or "SH" tests in the CONSEL software package (Shimodaira \& Hasegawa, 2001). Tree numbers correspond to trees shown in Figure 6.27. ${ }^{*}$ M. elegans $=$ Marssoniella elegans sensu Vossbrinck et al. (2004).

|  | Constraint |  |  |  |  |  |
| :---: | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Hypo- <br> thetical <br> Tree \# | Vossbrinckus richardi n. gen., <br> n. sp. grouped with: | AU | KH | SH | WKH | WSH |
|  | (unconstrained) | 1 | 0.998 | 1 | 0.998 | 1 |
| 1 | Gurleya spp. \& M. elegans* | $4 \mathrm{e}-004$ | 0.002 | 0.211 | 0.002 | 0.005 |
| 2 | not Gurleya spp. \& M. elegans | $4 \mathrm{e}-090$ | 0 | $3 \mathrm{e}-004$ | 0 | $9 \mathrm{e}-005$ |
| 3 | Hazardia milleri | $6 \mathrm{e}-099$ | 0 | $1 \mathrm{e}-004$ | 0 | 0 |
| 4 | Larssonia obtusa | $2 \mathrm{e}-049$ | 0 | 0 | 0 | 0 |
| 5 | Gurleya vavrai | $3 \mathrm{e}-068$ | 0 | 0 | 0 | 0 |




Figure 6.27: Tree topologies used in AU and related tests, showing just the taxa near clade A of Figs. 6.24-6.26. Boxes show monophyletic groups, corresponding to various hypotheses created as constraints prior to heuristic tree searches. Likelihood scores from these topologies were calculated using a best-fit substitution model from Modeltest and parameters estimated from the data to statistically test trees (results shown in Table 6.2). * Marssoniella elegans sensu Vossbrinck et al. (2004).

## Chapter 7: General Conclusions

## Overview

The purpose of this chapter is to present some recurrent themes in the results of the thesis, to discuss these in the context of broader evolutionary theory, and to relate these themes to practical questions about preventing microsporidial disease. After a brief introduction, this chapter is divided into five sections. The first section examines taxonomic conclusions from the thesis. It begins by discussing how species were evaluated and discusses the importance of intermediates between candidate species. It acknowledges the necessary vagueness in the underlying species definition, given our lack of knowledge of sex in the microsporidia. The remainder of this first section summarizes arguments for new species presented in the thesis and then discusses characters used in these arguments, including those that would be appropriate for field-diagnosis of species and those that may be developmentally plastic or evolutionarily convergent at some taxonomic levels. The second section examines biological conclusions, particularly those relating to the importance of host-parasite interactions at macro-evolutionary (co-evolution, biogeography) and micro-evolutionary (epidemiology, evolution of pathogenesis) scales. The third section discusses peculiarities in the DNA sequences examined throughout the thesis, highlighting the importance of indels (gaps), ribosomal DNA (rDNA) paralogs, and the implications of rDNA recombination and variable rates of molecular evolution in the microsporidia. The fourth section presents some new questions that emerged from this work. The final section gives a short summary of key points of this chapter.

## Introduction

Microsporidia examined in this thesis are of practical interest as pathogens causing mortality and commercial losses to fisheries worldwide (Olson \& Lannan, 1984; Kent et al., 1989; Fomena et al., 1992; Shaw \& Kent, 1999), and of more fundamental interest as model species showing remarkable potential to adapt in life cycle, development and morphology (Ebert, 1994; Agnew \& Koella, 1997; Koella \& Agnew, 1997; Koella et al., 1998). Microsporidia, like most intracellular parasites, typically show dramatic reduction in form associated with parasitism, or "parasitic loss" of complexity in cellular and biochemical features, and genome size (Biderre et al., 1995; Peyretaillade et al., 1998; Méténier \& Vivarès, 2001). Microsporidia appear to have undergone rapid and major morphological changes over their
evolutionary history as a result of parasitic loss. Microsporidians in this thesis showed variation among close and distant relatives at a level that suggests they may be under significant selective pressure or free to drift. For example, morphological variation was surprisingly high and was a poor predictor of genetic relatedness at the genus level (see text of preceding chapters, Appendix 12, and especially Tables 3.3, 3.4, 4.3; Figs. 3.2-3.75, 4.4, 4.8, 4.14, 4.17, 5.6, 6.26; and Fig. 7.1 of this chapter). This unexpectedly high level of morphological variation was both confounding and informative taxonomically here and in previous studies, depending on the taxonomic level in question (Canning \& Lom, 1986; Weiss \& Vossbrinck, 1999; Sánchez et al., 2001a).

This thesis addressed objectives listed in Chapter 1, by comparing molecular phylogenetic and morphological data. In every chapter, these data appeared to disagree. For example, Chapters 2 and 4, the variant "SV" from brook trout was expected to be conspecific with Loma salmonae based on gross morphology (see Sánchez et al., 2001a), but instead, it was demonstrated to be a separate species (Chapter 2 and Tables 4.8-4.10). In contrast, in Chapters 3 and 4, Loma species were more similar genetically (Fig. 4.4; Table 4.12) than had been expected by morphological analysis (Figs. 3.2-3.75; Tables 3.3, 3.4). In Chapter 5, Thelohania butleri was expected to group with other Thelohania species, based on morphological classification, but instead, it grouped distantly from other Thelohania species, nestled among a largely fish-parasitic clade (Fig. 5.9). Similarly, in Chapter 6, Vossbrinckus richardi ${ }^{1}$ was expected to resemble another species in a copepod host based on phylogenetic data and general morphology of the Amblyospora-like species, but instead, it stood alone, having unique morphology for this group (Figs. 6.1-6.5).

Thus, the first major section of this chapter examines the limits of morphology-based taxonomy of the microsporidia, considering cases in which methodology and measurement differences may be responsible for apparent morphological variation among microsporidia. This section then examines evidence for characters that display non-heritable variation (developmental plasticity) or heritable but reticulating variation (convergence). For example, the number of spores per parasitophorous vacuole may be plastic, depending on timing of host reproduction (see Fig. 7.1 column titled "sp", and Cali \& Takvorian, 1999). This same character may show evolutionary homoplasy (convergence) (compare tree to character "sp" in Fig. 7.1)

[^4]and so be largely taxonomically uninformative or misleading. If taxonomic studies examine too few individuals or species from too narrow a geographic range, such plasticity or homoplasy may be missed.

Molecular phylogenies can help highlight such problems. Homoplasy and plasticity are not only a challenge to taxonomists, but they may also give clues about the adaptive environment of the organism. The second major section of this chapter speculates on questions of adaptive environment and life-history strategy by examining biological conclusions from this thesis. This section briefly explores epidemiological models and pathological features that might help explain the morphological variability in microsporidia. For example, the ability of microsporidia to adapt should not be surprising, because it is generally believed that parasites must evolve quickly to keep up with the defenses of their hosts. This process is analogous to running as fast as they can only to stand still, which has been called the "Red Queen Hypothesis" (Hamilton, 1980; Bell, 1982; Lively, 2001). The mechanistic explanation for the phenomenon of parasitic loss suggested by the completed genome sequence of the microsporidian E. cuniculi (Katinka et al., 2001) is that microsporidia passively or actively lost most of their epistatic regulation cascades on route from their fungal ancestors. There remains the paradox of how they maintain such evolutionary plasticity with such reduced regulatory DNA and biochemical complexity. One explanation of the paradox is that part of this adaptation may be achieved as a secondary consequence of the production of more propagules (i.e. in microsporidia, the spores) per unit time combined with short generation times of parasites relative to that of the host. These general life-history traits common in microparasites, together create more variation for natural selection to act upon. This would be a natural selection-based explanation for the paradox of how microsporidia exhibit so much variation with so little machinery.

It can be difficult to resolve groups phylogenetically where the morphological change of interest happens because of genome size reduction. All chapters of this thesis illustrated the limits of phylogeny reconstruction as a method for studying morphological change in the microsporidia. For example, peculiarities of the microsporidial genome, such as frequent interspecific indels (Table 4.6), interspecific variation in the tendency to find rDNA paralogs (Figs. 4.6, 4.7), and reduced SSU and ITS variation and length affected DNA analyses. Therefore, it seems that in intracellular parasites the history of parasitic loss (molecular phylogenetic history) is most difficult to reconstruct in the lineages in which most of the loss has
happened; thus, there is a "Catch 22 " in studying the history of loss: where it happens, the critical evidence has usually disappeared.

## Taxonomic Conclusions

The central challenge in taxonomy is to interpret correctly how natural variation in observable features (phenotype, DNA, etc.) distributes itself relative to true relationships between species. The greatest obstacle in dealing with characters is recognizing those that converge (homoplaseous characters) or have non-heritable, environment-driven variation (developmental plasticity). When such characters are common, or when character evolution is poorly known, as is the case in the microsporidia, the usual solution is to examine many independent characters in many individuals. By analyzing such characters together, opposite directional biases from homoplasy or plasticity cancel each other out, so the net result may be a reasonable estimate of true relationships. This approach was employed separately for morphological and DNA data in Chapters 3 and 4 to assess species boundaries. By examining evolution of morphological characters plotted on DNA-based phylogenies that used relatively neutral housekeeping genes (e.g. rDNA, EF-1 $\alpha$, and RPB1), we can uncover clues about character evolution. For example, genetic and morphological data are compared in Loma species in simplified form in Figure 7.1. These results are explored further in this section.

## Defining and evaluating species

With few exceptions, explicit statement of a species definition is of great value to taxonomists, because it allows species hypotheses to be tested and disproved (Sites \& Crandall, 1997). Debate over species concepts, according to Mayr (2000), arises from confusion between species category definitions (theoretical species concepts that address how we determine whether a given set of populations constitute a full species or only a lower-level group) and species taxon definitions (working species concepts that address how we define criteria for organisms to belong together under a species name). The working definition can function separately from the underlying species concept upon which it was based and may be necessary when sexuality is not known. In such cases, a working definition may still be helpful even when a species category definition is not possible. For species examined in Chapters 2 to 6 , sex, chromosomal behaviour and ploidy were unknown, making species concepts that apply only to sexual, but not asexual species problematic. Therefore, only a working species concept was given. Depending on whether sex is present, the underlying species category definition (theoretical concept) may be consistent with the Biological Species Concept (BSC) of Mayr
(1940; and see Mayr, 2000) or with the lineage-based concepts, such as the Phylogenetic Species Concept (see Wheeler \& Platnick, 2000), or a combination such as the Evolutionary Species Concept (Wiley \& Mayden, 2000).

The working species concept used in this thesis was a modification of definitions discussed elsewhere (arguments in Wheeler \& Meier, 2000) and has two parts. The first criterion for distinguishing one species from another when in sympatry is that each must possess separate discrete characters or overlapping characters with a statistically separate mean. The second criterion is that the suite of characters that provides evidence for species must agree across the sampled populations. Chapter 3 focused mainly on morphological evidence for species, much of which was from overlapping characters that differ in the mean, whereas Chapter 4 focused on genetic evidence, which was evaluated as explained in Appendix 5. However, the unusual biology of the microsporidia affected how well species could be evaluated by this definition. For example, because microsporidians are obligate intracellular parasites, host individuals and species can be important as "islands" that create barriers that often correlate with genetic and morphological discontinuities. Thus, many groups of microsporidia tend to be highly host-specific and the host species is often a valid, easily recognized taxonomic character, representing 'host preference'. So here, as in many other studies, the assumption of host preference or specificity was used at the start to form testable hypotheses; however, this does not suggest that host was assumed to be a valid character. This is because the existence of the parasite in a host does not demonstrate specificity for this host (i.e. a parasite may use multiple host species). In Chapter 3, this problem was addressed by surveying potential alternate hosts for presence of the parasite. Furthermore, using host as a starting point could be a problem if morphological features arise from developmental plasticity shaped by the host (i.e. hostdependent phenotypic plasticity) because these may not be valid taxonomic characters if they are not be shared across the population. This problem was addressed by sampling a wide range of characters and developmental stages across many individuals (cells) and host isolates. However, ideally, before considering host or any host-correlated morphological differences as valid characters, we should test host preference (host specificity) by experimental transmission to alternate hosts, survey all potential alternate hosts at a wide range of times and locations for presence of the parasite perhaps using a diagnostic genetic marker, and also identify genetic differences between species that correlate with host or morphological differences. Generally, experimental transmission is often too technically difficult or impossible, and in the few cases where it was possible in this study (see next section), the results may not closely reflect
transmission in nature. Alternate hosts were surveyed over three years and genetic differences were examined for two independent loci (Chapter 4); however, these approaches were limited by time, resources, and the availability of suitably variable genetic markers, and so they do not represent an ideal test for the influence of host on morphology. Nevertheless, evidence for species was tested at least as rigorously here as in other studies of the microsporidia.

Using the working definition, the more complex hypothesis of two or more species, rather than one, was only accepted if it better fit the data. Failure to reject monophyletic clades for candidate species (Chapter 4) was interpreted as insufficient evidence for species; however, there was sufficient evidence to support these candidate species (Chapter 3) based on statistical analyses of morphological characters. Chapters 3 and 4 examined candidate species in three groups, each consisting of species having potential contact (i.e. sympatry) as specified in the working definition. These were: a Pacific marine group, containing seven Loma species from 12 fishes (Pacific cod, walleye pollock, Pacific tomcod, lingcod, sablefish, shiner perch, and six species of Pacific salmon and trout of Oncorhynchus spp.), an Atlantic marine group, containing two Loma species from Atlantic cod and haddock, and a freshwater eastern Canadian group, containing two Loma species from brook trout and rainbow trout (the latter being introduced to eastern North America). Species in the first two groups were considered to be in potential contact, because they produced xenomas in hosts collected not only in the same general geographic region, but also often in the same trawls. Spores released from these hosts can survive and be viable for up to a year in water (Shaw et al., 2000c), and so would be readily available to alternate hosts. Species in the eastern Canadian group were different in that hosts and possibly also their parasites were from the same geographic region, but infections were obtained by experimental contact (transmission between host) rather than natural contact; however, results in Chapter 2 suggest experimental transmission was not the source of infection.

Indirect indicators of species-integrity or interbreeding, such as presence of intermediates between sympatric populations, are not easy to interpret (Mayr, 2000). For example, populations may have distinct morphological or genetic differences but not yet be genetically isolated (a single polytypic species); or, may be reproductively isolated without yet showing morphological or genetic differences (cryptic or sibling species). Therefore, more direct experimental tests that can rule out these possibilities are preferable. Usually, for sexual species such tests involve attempting to interbreed two candidate species, whereas for microsporidia, whose sexuality is unknown, a possible test is to attempt transmission of two
candidate species to reciprocal hosts. This experimental test was performed on several species examined in this study as will be discussed in the next section (and see pgs 112-113, Chapter 3).

For the genus or higher categories (e.g. for species examined in Chapters 5 and 6), or for distantly related species or species in allopatry, boundaries between groups are more ambiguous. These groups are defined in relative terms using characters appropriate to the taxonomic level, as determined by specialists most familiar with the group.

## New species

Morphology-based classification of the microsporidia generally disagrees with geneticbased classification (see Weiss \& Vossbrinck, 1999). As long as morphological characters are examined as apomorphies at appropriate levels of classification as in Chapters 3 and 6 and not as synapomorphies used for phylogeny reconstruction in higher classification, the inability of morphological characters to predict molecular phylogenies should not be problematic. Thus, for close sister species, morphological characters were used as synapomorphies clustering individuals together into species groups (Chapter 3), but not for estimating higher-level (e.g. intrageneric) relationships. Morphological data provided more information by indirectly representing many more nucleotides, genes, and loci, whereas the DNA loci examined here appeared to be constrained, showing low variability. For example, Figure 4.20 suggested rDNA SSU, ITS and LSU evolves at a slow rate in Loma species and Tables 2.3 and 4.5 showed exceptionally low genetic variation in L. salmonae compared to other Loma species. This is surprising considering its wide geographic distribution. However, the low variation in $L$. salmonae may be the result of recent geographic spread of this parasite. Thus, these results indicated that the morphological evidence for species was more compelling than the DNA evidence, if morphological differences are not merely plastic responses to different hosts.

Five new species were described and two pairs of sibling-species were distinguished in Loma in Chapters 2 to 4, according to the working definition of species given above. In Chapter 6, a new species and genus was described based on differences at the genus-level. For new species described in this thesis, genetic and morphological features were compared among related species for which information was available, including the type species. The nine Loma species were compared to sympatric sister-species, whereas there were no described sympatric or sister species for the species described in Chapter 6. In Chapter 3 sister-species Loma
pacificodae ${ }^{2}$ from Pacific cod and Loma wallae from walleye pollock were distinguished by differences in developmental stages and vesicles of the host-parasite interface, and sister-species Loma lingcodae from lingcod and Loma richardi from sablefish were distinguished by xenoma features, vesicle and spore sizes. Loma kenti from Pacific tomcod was quite different from these four species, particularly in the nature of its vesicles, developmental timing, and spore features. Genetic analyses distinguished these five morphospecies into three clear groups but did not reliably resolve members of the two sister-species pairs into their respective speciesmonophylies (Figure 4.12 \#2, \#3, \#4). Failure to reject these monophyletic groups (Table 4.12) suggested species might be too recently diverged to have accumulated sufficient mutations at these loci (SSU, ITS and LSU rDNA, and EF-1 $\alpha$ ).

Previously described species from Atlantic cod and haddock in eastern Canada were found to be separate species based on DNA (Figs. 4.3, 4.4), as was originally proposed based on morphology (Morrison \& Sprague, 1981a). Given this result, the type species of genus Loma should be Loma morhua as proposed by Morrison \& Sprague (1981a) and not Loma branchialis as was proposed by Canning \& Lom (1986) who did not recognize these species as separate. Similarly, two specimens considered to be conspecific, L. salmonae and "L. salmonae variant SV', by Sánchez et al. (2001a; b) and Speare \& Daley (2003), were shown to be separate species. Chapters 2 and 4 showed that isolate "SV" from brook trout was separated from $L$. salmonae. This was demonstrated in phylogenetic tests (Tables 4.8-4.10), and by the level of genetic difference ( 19 bp or $1.4 \%$ in the SSU and 10.5 bp or $2.3 \%$ in the LSU) compared to that observed between L. salmonae and two sympatric sister-species, Loma embiotocia and $L$. lingcodae, that were not transmissible to reciprocal hosts (Shaw et al., 1997; Shaw \& Kent, 1999; Shaw et al., 2000a; R. W. Shaw, personal communication). More importantly, no isolates genetically resembling "SV" were found at locations (farms, wild, laboratory) from which "SV" supposedly originated. Therefore, "SV" was not a variant of $L$. salmonae from British Columbia salmon, as had been thought, but instead was a separate species, possibly a cryptic infection with Loma fontinalis from local brook trout.

Experimental transmission to alternate hosts should be performed to further resolve the question of reproductive boundaries. Some specimens collected in this study have already been tested in transmission experiments (Shaw \& Kent, 1999; Shaw et al., 2000a). For example, $L$.

[^5]lingcodae was shown experimentally to not infect the hosts of two genetically and morphologically similar sympatric species, L. salmonae and L. embiotocia (Shaw \& Kent, 1999; R. W. Shaw, personal communication). Similarly, L. pacificodae did not experimentally infect these hosts; however, the latter result was inconclusive because the original host, Pacific cod, was unavailable as an experimental control (R. W. Shaw, personal communication). These results, combined with previous transmission studies, and a survey of wild Pacific fishes for Loma-like infections (Table 3.1) suggest Loma species tend to have narrow host-breadth. The genus category, in contrast to the species category, is created for organizing species and has no biological meaning. Therefore, the criterion for recognizing a genus was that it constituted a monophyletic group of species possessing characters recognized by experts on the group. Accordingly, genus boundaries for Loma and Thelohania were examined in Chapters 4 and 5; with the result that Loma acerinae was shown to not be a valid member of the genus, and genus Thelohania was shown to be polyphyletic, forming three separate groups. These results are not surprising given how common polyphyletic genera are in the microsporidia (Baker et al., 1994; Baker et al., 1995; Baker et al., 1997; Docker et al., 1997; Pomport-Castillon et al., 1997; Nilsen et al., 1998; Nilsen, 2000; Pomport-Castillon et al., 2000; Bell, et al., 2001; Lom \& Nilsen, 2003).

The new genus and species $V$. richardi in Chapter 6 were described on the basis of their unique morphology (polyspory with three spore types, the commonest of these being pyriform, distinctly smaller and differently arranged than those of the nearest relative, Marssoniella elegans) and for their genetic distance ( $5.9 \%$ difference in the SSU rDNA or 78 substitutions) from M. elegans. As discussed in Chapter 6, these genetic distances are generally consistent with genus-level differences in this group.

## Field-diagnostic characters

Misidentification of microsporidian species is a widespread problem that can be solved by taxonomists providing field-diagnostic characters with species descriptions (Larsson, 1999). Field-diagnostic characters should be recognizable with little preparatory work, for example, using low magnification light microscopy. Throughout this thesis, many traditional recognition characters were found to overlap under low- and high-power light microscopy and therefore be unsuitable field-diagnostic characters (e.g. some spore features and the ITS rDNA sequence for some species). However, several good field-diagnostic characters were identified. For Loma species, host, and xenoma size and location appeared to be reasonable field-diagnostic characters; however, xenomas should be mature or nearly mature and measured to about 1 to 2
$\mu \mathrm{m}$. Statistically comparable xenoma size data are not yet available for many Loma species, including those needed to distinguish Loma sp. from brook trout from L. salmonae, and $L$. branchialis in haddock from L. morhua in Atlantic cod. For T. butleri and V. richardi (Chapters 5 and 6), reliable field-diagnostic characters could not be provided beyond characters given in the species descriptions because close relatives or sympatric sibling-species are unknown for these species.

## Plastic or convergent characters

When genetic and morphological characters suggest different relationships among specimens, the possible explanations fall on a continuum between two extremes: either (a) two specimens could appear morphologically identical but be significantly different in DNA sequence, or (b) two specimens could appear different morphologically but be identical in DNA sequence. Morphologically identical specimens in (a) may suggest convergent evolution, whereas genetically identical specimens in (b) may be evidence of developmental plasticity. Thus, when a morphological character is distributed in a pattern that does not reflect relatedness, the character may be plastic or convergent. Such characters are common in microsporidia (Weiss \& Vossbrinck, 1999) and occurred in every chapter of this thesis. The situation is shown most clearly in Loma species, as illustrated in Figure 7.1. In Figure 7.1, host group, xenoma size and number of tubules per vesicle (shaded columns) were distributed in a pattern consistent with intrageneric (interspecific) relationships, whereas all other morphological characters (un-shaded columns, e.g. spore ultrastructure, early and late development, host-parasite interface, and pathology and ecology) were not distributed in any obvious pattern with respect to deeper relationships among Loma species.

Morphological characters were distributed in a pattern consistent with relationships elsewhere in this thesis. For example, in Chapter 2, spore size and site in the host mistakenly suggested strain "SV" was conspecific with L. salmonae; however, Speare et al. (1998a) and Sánchez et al. (2001a; b) identified the specimens based on low-magnification light microscopy, but more detailed study may be necessary. Similarly, in Chapters 3 and 4, spores and gross morphology of $L$. branchialis and L. morhua mistakenly suggested synonymy of these species (Fig. 3.7; Table 4.3), which are clearly distinguished from one another, though not from all other Loma species, by DNA sequence (Figs. 4.3, 4.4). It is not clear whether these cases represent plasticity or convergence or merely imprecise measurement. In contrast, features of the hostparasite interface have been examined closely and appear to often misrepresent relatedness except among close sibling-species. In Chapter 4 the interface of $L$. acerinae falsely placed it
with the type species of Loma (Lom \& Pekkarinen, 1999), whereas two independent genetic loci now clearly show it to belong well outside the genus (Figs. 4.14-4.16, and Lom \& Nilsen, 2003). Figure 7.1 suggests other features of the host-parasite interface are poor genus-level characters in this group. Chapter 5 showed for genus Thelohania sporulation features and the number of spores per sporophorous vacuole were not indicative of relatedness; however, the genus appears to be too broadly defined and polyphyletic. Finally, in Chapter 6, differences in development and morphology (particularly sporulation sequence) turned out to be large between close sister-genera. Further data on sporulation sequences for relatives in the groups represented in Chapters 5 and 6 are needed to examine plasticity or convergence of these characters.

Differences in methodology, specimen fixation and sample size could explain some of the results discussed above. However, in Chapter 3, methodology was standardized, spore shrinkage due to fixation was estimated, potential artifacts of TEM sectioning were considered, and sample size was reported so that such measurement artifacts were minimized. These data (Fig. 7.1) are more likely to reflect real convergence rather than merely measurement differences. It is not known whether such convergence could be caused by functional or energetic limits (i.e. fitness tradeoffs) or by drift, rather than by unidirectional selection. Tests of phenotype-environment trait correlation would be needed to test this (via the comparative method of Harvey \& Pagel, 1991, see Poulin, 1995, and Pigliucci, 2001). Characters exhibiting high intraspecific variance (see Fig. 7.1) could be phenotypically plastic; however, demonstration of plasticity would require controlled experiments on genetic clones raised in environments that select for either extreme of the plastic character (Pigliucci, 2001). The observed variation and possible convergence in host-parasitic interface features could reflect some energetically expensive processes involved during the formation of interfacial structures. Several authors have suggested materials involved in chitinous spore wall formation are exchanged between parasite and host via the tubules during sporulation (Morrison \& Sprague, 1981a, b; Vávra \& Larsson, 1999). This could explain the pattern in interfacial features (e.g. tubules and vesicles) in Figure 7.1.

Nevertheless, the pattern of inter- and intra-specific variation in Figure 7.1 suggests Loma species retain great potential to adapt morphologically over short evolutionary times. They presumably do so with limited epistatic or pleiotropic regulation machinery, given their small genomes (Katinka et al., 2001). Further study of the most variable characters that have confounded taxonomists in the past may show these to be the most interesting and informative biological features, providing clues about selection and adaptation in the microsporidia;
however, testing whether these features are adaptive or merely result from drift would be difficult.

## Biological Conclusions

Once taxonomic questions have been addressed, we can build our knowledge of the species we have named and grouped. Biological questions were addressed by plotting host, geographic and other data on the organismal history inferred by molecular phylogenies. Mallet (1995) warns that data used to define groups and make inferences about the biology of such groups should be separate, to avoid circular reasoning. Thus, this section examines macro- and micro-evolutionary strategies using morphological and ecological data not central to taxonomic arguments for new species. This section also considers practical concerns of parasitologists studying aquaculture pathogens relating to controlling disease. Diseases evolve through the close interaction of parasite with host. This intimate association over the long-term results in host-parasite co-evolutionary (and biogeographic) patterns, whereas over the short term, it results in patterns of epidemiological (virulence, pathogenicity and transmissibility) change. The following sections review long- and short-time scale patterns in microsporidian biology observed in this thesis.

## Co-evolution and biogeographic patterns

Phylogenetic results from all chapters suggested host group remains one of the best taxonomic characters at both higher- and lower-levels, presumably because of host-parasite parallel evolution consistent with broad sense co-evolution; however, results also showed that host-switching may be important. Both parallel host-parasite evolution and host-switching appear to be important in the microsporidia (Baker et al., 1997; Baker et al., 1998; Weiss \& Vossbrinck, 1999; Lom \& Nilsen, 2003). Even close-to-perfect host-parasite phylogenetic congruence could have arisen by chance. Thus, results that were consistent with parallel evolution or co-evolution (Figs. 4.20, 5.9, 6.26) should be tested statistically when sufficient data become available (e.g. more species of Loma from gadids or further molecular phylogenetic data for non-gadid hosts), using programs such as TreeMap, TreeFitter or ParaFit (see Desdevises et al., 2002).

Results consistent with host-parasite parallel evolution or co-evolution occurred at all levels of relatedness in this thesis. For example, at the lowest level, host population subdivision (see Nielsen et al., 1997) into west coast and interior populations mirrored the genetic pattern in L. salmonae populations (Table 2.3). Similarly, host relatedness among Pacific salmon
predicted $L$. salmonae prevalence (Table 2.1). Host also predicted relatedness among isolates of L. salmonae and "SV" (see text of Chapter 2, Table 2.3 and 4.8-4.10). Again, host-relatedness predicted the closeness of $L$. lingcodae and $L$. richardi (Figs. 4.3, 4.4, and Appendix 14), and host relatedness in Gadiformes not only predicted relatedness in five Loma species, but also suggested they may have co-speciated with their hosts during colonization of the Pacific basin about 12 and 3.5 million years ago (Fig. 4.20). Higher-level relationships showed similar patterns. Host relatedness and ecology predicted relatedness among Thelohania-like species with the exception of a species from an isopod, which grouped with species from marine decapods (Figs. 5.9, 5.12). Similarly, host predicted the relationship between M. elegans and $V$. richardi (Fig. 6.26). However, the genetic and geographic distances between Thelohania-like and related crustacean-parasitic species are large, suggesting undiscovered intermediate species probably exist.

Parallel evolution or co-evolution is not a new observation, but has often been observed in the microsporidia (Baker et al., 1997, Baker et al., 1998; Weiss \& Vossbrinck, 1999), and for parasites in general (Klassen, 1992). Fahrenholz's rule, the null model that hosts and parasites co-speciate perfectly, producing perfectly congruent phylogenies, was proposed to explain this. Under tight host-parasite co-evolution, the host habit, biogeography, population genetics and physiology, in theory, might exert stronger selection on the parasite than any other extrinsic or intrinsic quality. If all of the apparent cases of co-evolution in this thesis are shown to be valid, these lead to predictions on micro-evolutionary scales. For example, host-specificity and prevalence may be high and vertical transmission is more likely to be found. Vertical transmission should select for more benign parasites under some circumstances (Bull et al., 1991; Ebert, 1994; Lipsitch et al., 1996; Dunn \& Smith, 2001). The next section discusses this in more detail.

Parasitologists, epidemiologists, and evolutionary biologists may be more interested in exceptions to Fahrenholz's rule, or cases of host-switching, because these better illuminate evolutionary innovations and the potential for new diseases. These are so common that some suggest the rule itself is generally wrong (see Desdevises et al., 2002; and Zietara \& Lumme, 2002). Given this, it is not surprising that host-switches (host-capture speciation) occur in the microsporidia examined here. For example, L. embiotocia from a shiner perch (order Perciformes) and Loma sp. from Australian surf bream (also order Perciformes) do not cluster together (Figs. 4.3, 4.4) and others from Perciformes (L. diplodae and L. boopsi) are expected not to group with L. embiotocia (Lom \& Pekkarinen, 1999). While molecular phylogenies are
still unavailable for many hosts, other obvious cases of host-switching exist. For example, fishparasitic microsporidia form several major clades that have probably colonized fishes independently more than once (Lom \& Nilsen, 2003). Also, crustacean-parasitic species, Daphnia-parasitic and isopod-parasitic (Fig. 5.12) seem to be host-switchers as a rule, rather than the exception (Refardt et al., 2002).

The pre-conditions for host-switching in the few cases where it appears to have happened may be the most interesting questions in co-evolution (Poulin, 1998). Pre-conditions could consist of changes in host geographic locality, food chain, fresh water versus marine environment, or use of multiple hosts by the parasite. The phylogenetic dispersion of crustacean-parasitic microsporidians (Fig. 5.9, and Refardt et al., 2002; Moodie et al., 2003) and multiple invasions of fishes (Lom \& Nilsen, 2003) together with the discovery of a new Nucleospora species in sea lice (Crustacea; Caligidae) (see Freeman et al., 2003) suggest crustaceans could play a role in the host-switching in the mostly fish-parasitic clade (clade F/C Fig. 5.9). Crustacean-parasitism could simply be ancestral to fish-parasitism in each case of fish invasion, or crustacean intermediate hosts of two-host life cycles could facilitate invasions of fishes by secondary host-switching (e.g. switching from insects to fish secondary hosts). The latter situation could be followed by loss of the intermediate host phase of the life-cycle, resulting in direct transmission in fishes. This process has been proposed by others (Baker et al., 1997) to be responsible for loss of the crustacean-portion of the life-cycle in microsporidia.

Host-switching can be associated with biogeographic phenomena, like the host colonizing a new environment. For example, Figure 4.4 (see clade L) suggests $L$. salmonae may have recently colonized the anadromous Oncorhynchus species from fresh water salmonid hosts, and Scorpaeniformes-parasitic Loma species may have arisen from fresh-water salmonidparasitic or L. salmonae-like ancestors, or L. embiotocia-like ancestors in the Pacific region about 8.2 to 4.6 million years ago (see Chapter 4). Genetic data (Fig. 4.20) also suggested gadid-parasitic Loma species may have originated in the Atlantic in tomcod-like hosts about 12 million years ago, given the early-branching position of $L$. branchialis.

## Epidemiology and pathology

Various epidemiological and pathological features were examined in this thesis. Many hosts of species studied here are commercially important, including salmon, trout, cod, haddock, pollock, lingcod, sablefish, and pink shrimp, so this discussion will focus largely on how to reduce the microsporidian infection. Transmission is one of the most important parameters in parasite life-histories, as it strongly affects the evolution of virulence according to
epidemiological models developed by Anderson \& May (1981). Little is known about differences in transmission or virulence among species studied here, except that Loma species are thought to be directly transmitted (Bekhti \& Bouix, 1985; Kent et al., 1989; Shaw et al., 1998; Shaw \& Kent, 1999; Shaw et al., 2000c; Lom, 2002) and known to cause high mortality in farms (Magor, 1987; Speare et al., 1989; Markey et al., 1994; Bruno et al., 1995; Bader et al., 1998; Kent et al., 1998; Kent, 2000). Thelohania butleri is also thought to be directly transmitted and cause significant mortality and possibly sterilization of hosts (Vernick et al., 1977; Johnston et al., 1978; Butler, 1980; Olson \& Lannan, 1984; J. Boutillier, personal communication). Transmission in $V$. richardi is unknown, but host mortality and sterilization would be assumed to be high in heavy infections. This section will review features relating to adaptations for transmission or host immune system evasion in the microsporidia studied here.

Features associated with transmission were among the more variable studied here. For example, polar filament length should be adapted for target hosts or cell types (Cali \& Takvorian, 1999), thus length (number of turns) showed inter- and intraspecific variation in this study (Fig. 3.8). Timing of sporulation and xenoma formation should be optimized to match the window of transmission opportunity in the host life-cycle. Not surprisingly, timing of sporulation and possibly also timing of xenoma formation appeared to be variable within and between species (see Chapters 2 and 3). Host breadth and use of reservoir hosts varied between species studied here (compare host breadth of L. salmonae to that of other Loma species). Host breadth should be inversely related to the importance of vertical transmission.

The simplest model for directly transmitted microparasites is $R_{0}=\beta \mathrm{N} /(\alpha+b+v)$, based on Anderson \& May (1981). Here, $R_{0}$ is the parasite's net reproductive rate, which is partly represented by the number of propagules (spores) per starting meront in microsporidia. $R_{0}$ is also a rough approximation for fitness (Poulin, 1998). Then, $\beta$ is the transmission coefficient (probability parasite will be transmitted during contact between hosts times the frequency of contact). The remaining terms in the equation are: $\mathrm{N}=$ the number of hosts in the population (or host density), $\alpha=$ the host death rate due to the parasite (or virulence), $b$ is the host death rate in the absence of the parasite, and $v$ is the recovery rate (or immunity). The two terms $R_{0}$ (reproductive rate) and $\beta$ (transmission) are two of the most important terms to epidemiologists, followed by $\alpha$ (virulence).

This model may be extremely simplistic and therefore limited in predictive value, but it can still be useful in allowing us to sort life-history traits categorically, and compare its
predictions to our observations. For example, it predicts that for competing strains of microparasites, those with the highest $R_{0}$ (reproductive rate) will win. $R_{0}$ would be highest if $\alpha$ (virulence) were zero, but $\alpha$ and $\beta$ (transmission) are often genetically correlated because of the biological processes inherent in parasitism (production and release of propagules tends to necessitate damage to the host). So $R_{0}$ is often maximized at $\alpha>0$. In crowded conditions, e.g. fish farms, one component of $\beta$, the rate of contact between hosts, will rise, and so all other factors being unchanged, virulence will increase. This is especially problematic for microsporidia because they can cause high mortality (Becnel \& Andreadis, 1999; Shaw \& Kent, 1999). Similarly, the model predicts that increased reservoirs for the parasite, either in the spore bank or reservoir hosts, will increase $\beta$ and allow $\alpha$ to increase. In cases where any increase in virulence $(\alpha)$ decreases transmission $(\beta)$ and $R_{0}$, parasites effectively can have large increases in $\beta$ with small increases in $\alpha$ and therefore increase reproductive rate ( $R_{0}$ ), thus, such a parasite strain can win against others without increasing virulence (e.g. for vertically transmitted parasites that depend on high survival of the parental generation). This may be unusual, except where vertical transmission is highly efficient in allowing transmission to offspring while producing few propagules (i.e. spores in microsporidia), thereby causing little damage to the hosts. Signs of vertical transmission are common in the microsporidia (Mangin et al., 1995; Terry et al., 1998; Dunn et al., 2001; Dunn \& Smith, 2001). Under most conditions, the model predicts vertically transmitted parasites should become more benign because larger numbers of surviving hosts provide more opportunities for transmission of the parasites (Bull et al., 1991; Lipsitch et al., 1995, 1996).

While there exist many elaborations of this general model, in its simple form the model can help us to interpret and make use of other observations in this thesis. Theory predicts that to fight a disease, we should inhibit transmission by (1) reducing crowding of hosts (reduce N ) (however, this may not be financially feasible and may reduce the effect of anti-parasitic drugs), (2) determining the geographic source of the infection (a component of $\beta$ ) using a molecular probe like that of Chapter 2 so that contaminating materials can be quarantined, (3) determining reservoir or alternate host species or host breadth (which affects $\beta$ ) as was done in Chapters 2 to 4 , (4) determining the spatial and temporal range of infection and persistence of spores in the environment (also a component of $\beta$ ); for example, see Figure 3.1, (5) determining the natural genetic variation corresponding to strains and boundaries between populations (or species) as was done throughout the thesis so we can prevent selecting for more virulent strains. The model
also predicts that when natural host mortality, $b$, rises, with intense fishing over a long enough time (e.g. coastal salmon or cod fishing), this could eventually increase virulence, $\alpha$, of the parasite (Poulin, 1995; 1998) or instead, $R_{0}$ would possibly not be maintained at the same level.

Prevalence differences among Atlantic and Pacific hosts (Chapter 4) or among different salmon species (Chapter 2) and differences in pathological signs (Chapter 3) could suggest pathogenicity (virulence) or life-cycle timing differences between and within Loma species. These differences could be caused by cryptic or dormant stages (e.g. in heart) or by differences in other mechanisms of host-immune system evasion such as host-parasite interfacial features (see previous sections), or genetic features associated with detection (antigenic surface features). Alternatively, the prevalence or pathogenicity variation could be from differences in the parasite's journey outside the host. Such variation would be seen in features of spores, such as dimensions and endo- and exo-spore wall features (examined in all chapters), which are adapted for dissemination and persistence in the environment prior to transmission.

If species can modify their life-history to increase fitness by varying the relative roles of vertical and horizontal transmission, then they may converge on similar strategies, and therefore have similar morphology (see previous sections). There was indirect evidence in this thesis for vertical transmission in Loma species (spores in gonads of L. pacificodae and L. salmonae), and in $V$. richardi. This model could be used to examine the observation that $L$. pacificodae seemed to elicit greater pathological reaction than $L$. kenti even though these species were closely related and both formed large xenomas, therefore perhaps having similar $R_{0}$. Loma pacificodae formed xenomas in the gonads, so may employ vertical transmission. However, in this case the model would predict the opposite result (that L. pacificodae should elicit less pathological reaction than $L$. kenti). This apparent contradiction might be explained by considering $R_{0}$ from the first point of infection of a host, in which a host gut cell becomes infected to the final formation of gill xenomas, instead of considering only the gill phase (gill meront to gill xenoma). Considering reproductive rate from the first point of infection, L. pacificodae would probably have much higher $R_{0}$ than L. kenti, because the hosts of L. pacificodae (Pacific cod) are much larger than $L$. kenti's hosts (Pacific tomcod). Of course, other factors may differ between these species.

The survey of Pacific fishes (Table 3.1) suggested all seven Pacific Loma species might be quite host-specific (although L. salmonae is a generalist of Oncorhynchus spp.). Table 3.2 also shows that these species are at high prevalence in nature. This could be a sign that all these species may be similar in ecology and may all use vertical transmission to some extent.

Whether vertical transmission actually occurs in species studied here and how it might affect developmental sequence or other characters, such as those in the un-shaded columns in Figure 7.1, is unknown. Vertical transmission in the microsporidia is thought to be associated with condensed developmental sequences or loss of a host and the associated spore form from the life-cycle (Baker et al., 1997; Sweeney et al., 1989; Iwano \& Kurtti, 1995; Koella et al., 1998). Nevertheless, this thesis supports the finding by others that convergence in life-history strategy has probably been underestimated in the microsporidia (Weiss \& Vossbrinck, 1999), as has vertical transmission (Dunn et al., 2001).

## DNA Peculiarities

Knowledge of the microsporidian genome has advanced dramatically since this work began. Some of this knowledge explains peculiarities in DNA sequences observed throughout this thesis. For example, the next sections will show that paralogs and indels were unexpectedly frequent in rDNA, particularly in the $5^{\prime}$ LSU and ITS regions, and rDNA appeared to evolve at a slow rate. These can be explained by two observations. First, rDNA copies are dispersed in the genome, rather than in tandem, in many microsporidia (Peyretaillade et al., 1998; Vivarès \& Méténier, 2000; Méténier \& Vivarès, 2001). This dispersion decreases the rate of homogenization of rDNA mutations between copies and increases the divergence among rDNA paralogs. Second, the rDNA, and especially the ITS, is extremely short in many microsporidia. The recently completed full genome sequence of $E$. cuniculi shows all genes are severely compacted, having very little regulatory DNA, little non-coding DNA, and fewer and shorter proteins than other eukaryotes (Katinka et al., 2001). This compaction leaves little non-coding sequence and may explain the lack of resolution of sister-species in Chapter 4. These observations are explained in further detail below.

## Indels and paralogs

Intra-individual heterogeneity in rDNA (paralogs) was observed in all chapters of this thesis. The most direct evidence of this was within-individual rDNA polymorphisms in DNA extracted from a single xenoma (Chapter 4). Single xenomas from isolates named "Gi" from $L$. pacificodae and "Ai" from L. morhua, each produced two unique rDNA (SSU, ITS and partial LSU) sequences differing by seven substitutions and one indel, six base pairs long (Chapter 4). Since a single xenoma is likely to be the result of asexual divisions from a single starting spore (sporoplasm), these sequences are almost certainly intra-genomic rDNA paralogs. There was also less direct evidence of rDNA paralogs. For example, multiple (up to 10) unique rDNA
sequences were found per isolate (a small piece of infected host tissue from a single host). These unique sequences usually differed by short indels, one to six base pairs long, as well as nucleotide substitutions. These unique intra-isolate rDNA sequences formed copy-based groups rather than species-based groups in phylogenetic analyses (Figs. 4.6, 4.7), and the speciesrelationships were retrieved within each copy-based group, as shown in Figure 4.7. This phylogenetic result is strong evidence that these are paralogs that diverged before divergence of the species. These paralogous relationships appear to have interfered with phylogenetic resolution of relationships among gadid-parasitic Loma species.

Species seemed to differ in the degree of divergence between paralogs. For example, paralogs in gadid-parasitic Loma species were more diverged than those in L. salmonae and other Loma species (Table 4.5). Paralog divergence is expected to vary with degree of rDNA copy dispersion. When rDNA is dispersed rather than in tandem in the genome, copies tend to be less well homogenized by processes of concerted evolution (unequal crossing over and gene conversion) (Dover, 1982; Hancock \& Dover, 1990; Buckler et al., 1997; Hughes \& Peterson, 2001). Ribosomal DNA is distributed differently in different lineages and even among close relatives in the microsporidia, probably due to chromosomal rearrangements (Méténier \& Vivarès, 2001); thus, the results in Figure 4.7 are not surprising.

Efficient homogenization of rDNA copies or prevention of paralog divergence may still occur when copies are dispersed. The process implicated in dispersed-copy homogenization, which appears to prevent rDNA divergence for species with dispersed rDNA, is gene conversion (Liao, 2000). Intra-individual rDNA divergence is extremely common in intracellular parasites with reduced genomes, probably due to loss of recombinational gene conversion apparatus that normally homogenize the rDNA copies (Dale et al., 2003). Defective repair genes, such as recA and recF, have been associated with rDNA paralogs in parasitic intracellular bacterial (Liao, 2000; Shigenobu et al., 2000, Tamas et al., 2002). The same process has been suggested for eukaryotes (Hughes \& Peterson, 2001; Dale et al., 2003). Loss of repair gene efficacy due to relaxed selection for repair in the intracellular environment, where parasites are protected from many sources of DNA damage, has been proposed as the cause of paralog divergence in bacteria (Dale et al., 2003). However, the higher paralog divergence in gadid-parasitic Loma species could instead be caused by interspecific hybridization between sympatric sister-species. Several studies, where interspecific hybridization is the known cause of rDNA paralog divergence, show similar results to these (Sang et al., 1995; Hugall et al., 1999; Hughes \& Peterson, 2001).

On a practical level, different rDNA copy homogenization rates among different organismal or rDNA-copy lineages (e.g. gadid-parasitic versus non-gadid parasitic group, or alternate copy groups) could affect phylogenetic results by violating the assumption of equal rDNA evolutionary rates among lineages. Differences in rates across lineages could arise from small changes in repair genes or from larger changes, such as chromosomal rearrangements that affect rDNA dispersion. Another possibility is that intra-genomic differences in rDNA evolution could arise if different paralogs are adapted for different purposes and are thereby under different selective pressure (e.g. multiple ribosome forms which fold differently or bind to different riboproteins).

Indels were often phylogenetically informative in this thesis both among and within species; however, they were a practical problem in two ways. First, indels obscured sequences produced directly from PCR products, necessitating the design and use of the Flip program (Appendix 1) or cloning of all PCR products to retrieve the correct sequence (Figs. A1.1-A1.9). Second, indels (or alignment gaps) required design and use of a program GapMatrix (Appendix 1) to count their phylogenetic information (Figs. A1.10-A1.19). Others have demonstrated that when indels occur at a high rate, they often hold phylogenetic, recombinational, and interspecific hybridizational information that is not available in substitutional data (Giribet \& Wheeler, 1999; Cheynier et al., 2001). The results here (Tables 4.6, 4.7, 5.2, Fig. 4.6) confirm the importance of including indel information whenever possible.

## Ribosomal DNA evolutionary rate

This is the first instance where microsporidian evolutionary rDNA rate has been calibrated based on geological data and molecular and phylogenetic data from the host (Fig. 4.20). The congruent branching pattern of gadid-parasitic Loma species and their hosts was used to estimate divergence times under the simple assumption that hosts and parasites cospeciated. Although co-speciation was not statistically tested and congruence only appears in part of this tree, geological and biogeographic data of the hosts (Carr et al., 1999) suggested speciation in Loma and their hosts occurred between 3 and 12 million years ago, and give a rate for SSU rDNA divergence of just less than two base pairs out of 1344 (or $0.146 \%$ ) per million years. However, these rates may not apply to other microsporidia, as rDNA rates appear to differ across the microsporidian tree. For example, branch lengths seem to be highly variable and fish-parasitic species, in particular, they seem to have especially short branch-lengths compared to other groups (Figs. 4.14, 5.9, and see Cheney et al., 2000; Lom \& Nilsen, 2003). In addition, rDNA regions appear to differ in rate between species. For example, isolates of $L$.
$k e n t i$ from Pacific tomcod were more divergent in the upstream SSU rDNA compared to isolates of other species, and the ITS region was less variant in salmonid-parasitic Loma species than in other Loma species.

## New Questions

Three topics emerged from this chapter that produce questions for further investigation.

1. This chapter highlighted the fact that there is much we still do not know about transmission, reproductive boundary formation, and potential developmental plasticity in different hosts or under different conditions. The thesis showed that xenomas are key structures that evolve with the formation of species, yet they may show great developmental and evolutionary plasticity. For example, xenomas may not form in some individuals (see Chapter 2) and can be of different forms in a single host (see Chapter 3). Over longer timescales xenomas show plasticity in that they appear to have arisen independently several times in the microsporidia (Lom, 2002; Lom \& Nilsen, 2003). This independent evolution of the xenoma could be through non-homologous character evolution (convergence). Developmental plasticity can be tested with transmission experiments to alternate tissues or sites in host. Experimental transmission to different host cell types may be used to see if this is a developmentally plastic trait. This could address the question: does the site of infection determine or limit xenoma size? An alternate possibility is that xenoma size could determine the site of infection. Understanding determinants of xenoma size could be of interest to epidemiologists because xenoma size probably correlates with the number of spores from a starting meront, thus positively correlating with transmission $\beta$ or $R_{0}$, when all other things are equal. Data from this study suggested that longer sporulation with more division cycles could produce larger xenomas, whereas a shorter sporulation sequence with fewer divisions would produce smaller xenomas. Comparative studies could address whether sporulation sequence length and xenoma size are truly positively correlated. Internal features of the xenoma may also be studied in this way, as they appear to vary both among and between species (see Fig. 7.1). While we might consider how xenoma cytoplasmic features (e.g. vesicles and tubules) could be adaptive, this may be difficult to test. Results here also suggested we should consider correlations between xenoma size and the intracellular location of merogonic divisions; for example, some species may develop in the host endoplasmic reticulum whereas others may develop free in the host cytoplasm. Another feature that may be related to the xenoma is spore size, which could either depend on packing of the spore contents (polar filament, nucleus, and materials to help spores persist in the environment),
or, alternately on packing of spores into the xenoma. Transmission tests may help address all these questions. However, these may be hard to perform for species that are not easily maintained in the laboratory (e.g. sablefish), or that are difficult to maintain free of parasites (see Shaw et al., 2000b). In general, manipulation of microsporidian infections tends to be labourintensive.
2. Future studies of the systematics of microsporidia should be more strategic in selecting taxa to address other questions. In this study, species were selected to address general biological and taxonomic questions of fisheries and aquacultural pathogens. The questions were taxonomic, so the focus was on collecting species resembling these known pathogens, as well as specimens from the type locality and type species. But for phylogenetic questions of broader interest to microsporidian systematists, taxa should be selected accordingly. For example, only a tiny fraction of the species in groups that are intermediate or transitional between the major microsporidian groups (e.g. host-parasite groups of Baker et al., 1997) have been examined. The primitive Metchnikovellidae and taxa basal to mixed-host groups (e.g. Pseudoloma, etc.) as well as species in the copepod and micro-crustacean-parasitic groups and fresh water host groups have been underrepresented. Data from these species will be needed to understand major transitions in microsporidian evolution. If future studies are to focus on species boundaries, the origins of new species or strains, or new host invasions (host-switching), then species should be collected accordingly. Because we typically study very little of a species' total natural variation (across individuals and traits), we must attempt to be strategic in collecting samples and choosing traits to measure. Chapter 3 showed the importance of collecting more isolates, examining more early stages (meronts), and modifying protocols to maximize the data gathered from the available material (see Appendix 20). If future studies focus on the question of prerequisites for evolutionary innovations or adaptation, then a comparative approach could be used to distinguish true correlation from merely random associations in phylogenies. A comparative approach could be used to study parasitic loss (loss of size, life-cycle complexity, and genome size). For example, we might ask whether microsporidians with the most reduced mitochondrial remnants have undergone stronger selection to be reduced in size or are freer to evolve to become smaller than those with less reduced mitochondria-like structures (see Williams et al., 2002).
3. While the full genome sequence of E. cuniculi has provided much information about how the microsporidian genome is structured, we still know very little about how recombination and repair occur or are regulated. This discussion illuminated the possibility that genetic
recombination and repair may be variable in the microsporidia and may affect how molecular markers evolve, thus affecting how they should be interpreted. Furthermore, we do not yet know whether sexual crossing-over occurs in some groups, and affects marker segregation and the spread of mutations. We also do not know how the reduced genome size is tolerated, although this would be difficult to address. Future studies could compare species for the presence of non-coding or highly repeated DNA sequences to examine relative reduction in genome (thereby infer the strength of selection or drift contributing to genome compaction). Future studies could also compare full rDNA genes from single xenomas, looking for signs of rDNA recombination using SplitsTree and LARD. We could then address questions such as whether more recombination produces more indels (Cheynier et al., 2001). While indels may be created frequently in microsporidia because of defects in repair machinery, we do not know whether they are better tolerated in microsporidia than in other eukaryotes.

## Summary

Microsporidia seem to retain high levels of morphological variation despite having the smallest genomes of any eukaryotes. This morphological variation occurs over both microevolutionary and macroevolutionary time scales. This thesis used morphology and DNA sequence to describe and distinguish five new species of Loma in Chapter 3 and one new genus and species in Chapter 6, and to distinguish several cryptic species-pairs in Chapters 2 and 4, and to identify polyphyletic genera in Chapters 4 and 5. Morphology was unexpectedly more informative than DNA for close relationships among Loma species and served to distinguish sister species in different hosts. Thus, careful re-analysis of inter- and intra-specific morphological variation, incorporating statistical tests and considering whether molecular phylogenetics support morphological character-based groups is critically important for these microsporidia. While several characters (e.g. host, xenoma size, xenoma location, and number of tubules per vesicle) were distributed in a pattern consistent with deeper genetic relationships in genus Loma, most were not (see Fig. 7.1). However, other characters that were more homoplaseous at higher levels in the genus (e.g. number of spores per parasitophorous vacuole, 'spore ultrastructure, early and late developmental stages, and features of the host-parasite interface) were useful for distinguishing closely related Loma species (see Chapter 3). These results were in contrast with previous studies that suggested morphology generally overlaps at a level that makes it impossible to distinguish good species. Although morphological variation might be expected to be somewhat limited in microsporidia because of the reduced cellular
complexity and lack of regulatory genes, it probably provided more taxonomic information than genetic data because morphological characters were the product of many independent genes. Furthermore, it may provide more information in this study because more characters, stages, individuals and geographic isolates were examined here compared with previous studies.

For phylogenetic analysis, morphological data tends to be unsuitable, particularly in the microsporidia, because it is so often homoplaseous at deeper levels. So, even though genetic data provided less information than morphology, molecular information gathered here was from cellular housekeeping genes that presumably evolve relatively neutrally with respect to wholeorganism or species differences, and hence are more appropriate for phylogeny reconstruction. Species were considered valid based on a combination of data, including the correlation of host and morphology, a reduced presence of genetic or morphological intermediates in sympatry, and where possible, based on the failure to experimentally transmit to reciprocal hosts. Molecular data were evaluated statistically against a working species concept to explore species boundaries in Chapters 2 to 4 . Genus boundaries examined in Chapters 5 and 6 were determined phylogenetically and based on differences in traditionally recognized genus-level characters. Ribosomal DNA had been expected to vary significantly at a level that could distinguish species and strains, based on previous studies, but results in Chapter 4 suggest for Loma species, this may not be the case. While rDNA, EF-1 $\alpha$ and RPB1 were in agreement about relationships and were informative at low- and high-levels of taxonomy, they were unable to resolve some relationships. Resolution with rDNA was particularly affected by paralogs and indels and overall low variation for closely related species. This suggests a greater continuous length of DNA must be sequenced to provide sufficient data to separate species and distinguish paralogs. Estimation of evolutionary rates in rDNA from gadid-parasitic Loma species provides further information for future study of the rapid radiations of fish-parasitic microsporidia compared to the slower-evolving but higher rDNA rate of the Amblyospora-like clade.

By examining morphological features independently of genetic phylogeny reconstruction, this study was able to examine possible cases of developmental plasticity or convergence. Even if morphology-based classification schemes of the past have been toppled by newer genetic data, they were not in vain; on the contrary, morphological data obtained in the past and in this study (in an attempt to classify the microsporidia) highlights important morphological adaptations. Similarities in host and parasite phylogenies found in this thesis may be consistent with co-evolution. If host-parasite co-evolution is important in shaping the evolution of these microsporidia, this could explain the morphological diversity observed here at
both higher and lower taxonomic levels. Epidemiological models predict that the selective environment, which may be intensified with increased dependence on the host, may accelerate adaptive evolution. Thus, processes of morphological evolution in the microsporidia could be examined by studying host-parasite co-evolution versus host-switching, the role of host group, and the role of vertical vs. horizontal transmission. Future studies will be needed to test these ideas, employing experimental transmission, completing parts of the tree, and using the comparative method to determine correlations of trait and environment (e.g. host), while eliminating the confounding factor of non-independence due to shared ancestry.

## LITERATURE CITED

Agnew, P. and Koella, J. ¿. 1997. Virulence, parasite mode of transmission, and host fluctuating asymmetry. Proceedings of the Royal Society of London Series B 264:9-15.

Anderson, R. M. and May, R. M. 1981. The population dynamics of microparasites and their invertebrate hosts. Proceedings of the Philosophical Transactions of the Royal Society 291:451-524.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Maddox, J. V. 1997. Phylogenetic position of Amblyospora Hazard \& Oldacre (Microspora: Amblyosporidae) based on small subunit rRNA data and its implication for the evolution of the microsporidia. Journal of Eukaryotic Microbiology 44(3):220-225.

Baker, M. D., Vossbrinck, C. R., Maddox, J. V. and Undeen, A. H. 1994. Phylogenetic relationships among Vairimorpha and Nosema species (Microspora) based on ribosomal RNA sequence data. Journal of Invertebrate Pathology 64:100-106.

Baker, M. D., Vossbrinck, C. R., Didier, E. S., Maddox, J. V. and Shadduck, J. A. 1995. Small subunit ribosomal DNA phylogeny of various microsporidia with emphasis on AIDS related forms. Journal of Eukaryotic Microbiology 42(5):564-570.

Baker, M. D., Vossbrinck, C. R., Becnel, J. J. and Andreadis, T. G. 1998. Phylogeny of Amblyospora (Microsporida: Amblyosporidae) and related genera based on small subunit ribosomal DNA data: a possible example of host parasite cospeciation. Journal of Invertebrate Pathology 71:199-206.

Becnel, J. J. and Andreadis, T. G. 1999. Microsporidia in insects. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 447501.

Bekhti, M. and Bouix, G. 1985. Loma salmonae (Putz, Hoffman et Dunbar, 1965) et Loma diplodae n. sp., microsporidies parasites de branchies de poissons téléosteens: implantation et données ultrastructurales. Protistologica 21(1):47-59.

Bell, A. S., Aoki, T. and Yokoyama, H. 2001. Phylogenetic relationships among microsporidia based on rDNA sequence data, with particular reference to fish-infecting Microsporidium Balbiani 1884 species. Journal of Eukaryotic Microbiology 48(3):258-265.

Biderre, C., Pages, M., Méténier, G., Canning, E. U. and Vivarès, C. P. 1995. Evidence for the smallest nuclear genome ( 2.9 Mb ) in the microsporidium Encephalitozoon cuniculi. Molecular and Biochemical Parasitology 74:229-231.

Bruno, D. W., Collins, R. O. and Morrison, C. M. 1995. The occurrence of Loma salmonae (Protozoa: Microspora) in farmed rainbow trout, Oncorhynchus mykiss Walbaum, in Scotland. Aquaculture 133:341-344.

Buckler, E. S., Ippolito, A. and Holtsford, T. P. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145:821-832.

Bull, J. J., Molineux, I. J. and Rice, W. R. 1991. Selection of benevolence in a host-parasite system. Evolution. 45(4):875-882.

Butler, T. H. 1980. Shrimps of the Pacific coast of Canada. in Canadian Bulletin of Fisheries and Aquatic Sciences 202.280 pp .

Cali, A. and Takvorian, P. M. 1999. Developmental morphology and life cycles of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 85-128.

Canning, E. U. and Lom, J. 1986. The microsporidia of vertebrates. Academic Press, London. 289 pp.

Carr, S. M., Kivlichan, D. S., Pepin, P. and Crutcher, D. C. 1999. Molecular systematics of gadid fishes: implications for the biogeographic origins of Pacific species. Canadian Journal of Zoology 77(1):19-26.

Cheney, S. A., Lafranchi-Tristem, N. J. and Canning, E. U. 2000. Phylogenetic relationships of Pleistophora-like microsporidia based on small subunit ribosomal DNA sequences and implications for the source of Trachipleistophora hominis infections. Journal of Eukaryotic Microbiology 47:280-287.

Cheynier, R., Kils-Hütten, L., Meyerhans, A. and Wain-Hobson, S. 2001. Insertion/deletion frequencies match those of point mutations in the hypervariable regions of the simian immunodeficiency virus surface envelope gene. Journal of General Virology 82:1613-1619.

Dale, C., Wang, B., Moran, N. and Ochman, H. 2003. Loss of DNA recombinational repair enzymes in the initial stages of genome degradation. Molecular Biology \& Evolution 20(8):1188-1 194.

Desdevises, Y., Morand, S., Jousson, O. and Legendre, P. 2002. Coevolution between Lamellodiscus (Monogenea: Diplectanidae) and Sparidae (Teleostei): the study of a complex host-parasite system. Evolution 56(12):2459-2471.

Docker, M. F., Kent, M. L., Hervio, D. M. L., Khattra, J. S., Weiss, L. M., Cali, A. and Devlin, R. H. 1997. Ribosomal DNA sequence of Nucleospora salmonis Hedrick, Groff and Baxa, 1991 (Microsporea: Enterocytozoonidae): implications for phylogeny and nomenclature. Journal of Eukaryotic Microbiology 44(1):55-60.

Dover, G. A. 1982. Molecular drive: a cohesive mode of species evolution. Nature 299:111-117.
Dunn, A. M. and Smith, J. E. 2001. Microsporidian life cycles and diversity: the relationship between virulence and transmission. Microbes and Infection 3:381-388.

Dunn, A. M., Terry, R. S. and Smith, J. E. 2001. Transovarial transmission in the microsporidia. Advances in Parasitology 48:57-100.

Ebert, D. 1994. Virulence and local adaptation of a horizontally transmitted parasite. Science 265:1084-1086.
Fomena, A., Coste, F. and Bouix, G. 1992. Loma camerounensis new species (Protozoa: Microsporida) a parasite of Oreochromis niloticus Linnaeus 1757 Teleost Cichlidae in fish-rearing ponds in Melen Yaounde Cameroon. Parasitology Research 78(3):201-208.

Freeman, M. A., Bell, A. S. and Sommerville, C. 2003. A hyperparasitic microsporidian infecting the salmon louse, Lepeophtheirus salmonis: an rDNA-based molecular phylogenetic study. Journal of Fish Diseases 26:667676.

Giribet, G. and Wheeler, W. C. 1999. On gaps. Molecular Phylogenetics and Evolution 13(1):132-143.
Hancock, J. M. and Dover, G. A. 1990. 'Compensatory slippage' in the evolution of ribosomal RNA genes. Nucleic Acids Research 18(20):5949-5954.

Hugall, A., Stanton, J. and Moritz, C. 1999. Reticulate evolution and the origins of ribosomal internal transcribed spacer diversity in apomictic Meloidogyne. Molecular Biology and Evolution 16:157-164.

Hughes, K. W. and Peterson, R. H. 2001. Apparent recombination or gene conversion in the ribosomal ITS region of a Flammulina (Fungi, Agaricales) hybrid. Molecular Biology and Evolution 18:94-96.

Johnston, L. B., Vernick, S. H. and Sprague, V. 1978. Light and electron microscope study of a new species of Thelohania (Microsporida) in the shrimp Pandalus jordani. Journal of Invertebrate Pathology 32:278-290.

Katinka, M. D., Duprat, S., Cronillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J. and

Vivarès, C. P. 2001. Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. Nature 414:450-453.

Kent, M. L. 2000. Marine netpen farming leads to infections with some unusual parasites. International Journal for Parasitology 30:321-326.

Kent, M. L., Elliot, D. G., Groff, J. M. and Hedrick, R. P. 1989. Loma salmonae (Protozoa: Microspora) infections in seawater reared coho salmon Oncorhynchus tshawytscha. Diseases of Aquatic Organisms 20:231-233.

Kent, M. L., Traxler, G. S., Kieser, D., Richard, J., Dawe, S. C., Shaw, R. W., Prosperi-Porta, G., Ketcheson, J. and Evelyn, T. P. T. 1998. Survey of salmonid pathogens in ocean-caught fishes in British Columbia, Canada. Journal of Aquatic Animal Health 10:211-219.

Klassen, G. J. 1992. Coevolution: A history of the macroevolutionary approach to studying host parasite associations. Journal of Parasitology 78:573-587.

Koella, J. C. and Agnew, P. 1997. Blood-feeding success of the mosquito Aedes aegypti depends on the transmission route of its parasite Edhazardia aedis. Oikos 78:311-316.

Koella, J. C., Agnew, P. and Michalakis, Y. 1998. Coevolutionary interactions between host life histories and parasite life cycles. Parasitology 116, S47-S55.

Larsson, J. I. R. 1999. Identification of microsporidia. Acta Protozoologica 38:161-197.
Liao, D. 2000. Gene conversion drives within genic sequences: concerted evolution of ribosomal RNA genes in bacteria and archaea. Journal of Molecular Evolution 51:305-317.

Lipsitch, M., Nowak, M. A., Ebert, D. and May, R. M. 1995. The population dynamics of vertically and horizontally transmitted parasites. Proceedings of the Royal Society of London. Series B. Biological Sciences 260:321-327.

Lipsitch, M., Siller, S. and Nowak, M. A. 1996. The evolution of virulence in pathogens with vertical and horizontal transmission. Evolution 50(5):1729-1741.

Lively, C. M. 2001. Parasite-host interactions. In Evolutionary Ecology Concepts and Case Studies. Fox, C. W., Roff, D. A. and Fairbainn, D. J. (eds.). Oxford University Press, Oxford, UK.

Lom, J. 2002. A catalogue of described genera and species of microsporidians parasitic in fish. Systematic Parasitology 53:81-99.

Lom, J. and Nilsen, F. 2003. Fish microsporidia: fine structural diversity and phylogeny. International Journal for Parasitology 33:107-127.

Lom, J. and Pekkarinen, M. 1999. Ultrastructural observations on Loma acerinae (Jirovec, 1930) comb. nov. (Phylum Microsporidia). Acta Protozoologica 38:61-74.

Magor, B. G. 1987. First report of Loma sp. (Microsporida) in juvenile coho salmon (Oncorhynchus kisutch) from Vancouver Island, British Columbia. Canadian Journal of Zoology 65:751-752.

Mallet, J. 1995. A species definition for the modern synthesis. Trends in Ecology and Evolution 10:294-299.
Mangin, K. L., Lipsitch, M. and Ebert, D. 1995. Virulence and transmission modes of two microsporidia in Daphnia magna. Parasitology 111(2):133-142.

Markey, P. T., Blazer, V. S., Ewing, M. S. and Kocan, K. M. 1994. Loma sp. in salmonids from the eastern United States: associated lesions in rainbow trout. Journal of Aquatic Animal Health 6:318-328.

Mayr, E. 1940. Speciation phenomena in birds. The American Naturalist 74:249-278.

Mayr, E. 2000. The biological species concept. In Species concepts and phylogenetic theory. Wheeler, Q. D. and Meier, R. (eds.). Columbia University Press. New York, NY. 230 pp.

Meier, R. and Willmann, R. 2000. The Hennigian species concept. In Species concepts and phylogenetic theory. Wheeler, Q. D. and Meier, R. (eds.). Columbia University Press. New York, NY. 230 pp.

Méténier, G. and Vivarès, C. P. 2001. Molecular characteristics and physiology of microsporidia. Microbes and Infection 3:407-415.

Moodie, E. G., Le Jambre, L. F. and Katz, M. E. 2003. Thelohania montirivulorum sp. nov. (Microspora: Thelohaniidae), a parasite of the Australian freshwater crayfish, Cherax destructor (Decapoda: Parastacidae): fine ultrastructure, molecular characteristics and phylogenetic relationships. Parasitology Research 91:215-228.

Morrison, C. M. and Sprague, V. 1981a. Electron microscopical study of a new genus and new species of microsporida in the gills of Atlantic cod Gadus morhua L. Journal of Fish Diseases 4:15-32.

Morrison, C. M. and Sprague, V. 1981b. Light and electron microscope study of microsporida in the gill of haddock, Melanogrammus aeglefinus (L.). Journal of Fish Diseases 4:179-184.

Nielsen, J. L., Fountain, M. C. and Wright, J. M. 1997. Biogeographic analysis of pacific trout (Oncorhynchus mykiss) in California and Mexico based on mitochondrial DNA and nuclear microsatellites. In Molecular systematics of fishes. Kocker, T. D. and Stepian, C. A. (eds.). Academic Press, San Diego. 314 pp.

Nilsen, F. 2000. Small subunit ribosomal DNA phylogeny of microsporidia with particular reference to genera that infect fish. Journal of Parasitology 86(1):128-133.

Nilsen, F., Endresen, C. and Hordvick, I. 1998. Molecular phylogeny of microsporidians with particular reference to species that infect the muscles of fish. Journal of Eukaryotic Microbiology 45:535-543.

Olson, R. E. and Lannan, C. N. 1984. Prevalence of microsporidian infection in commercially caught pink shrimp, Pandalus jordani. Journal of Invertebrate Pathology 43:407-413.

Peyretaillade, E., Biderre, C., Peyret, P., Duffieux, F., Méténier, G., Gouy, M., Michot, B. and Vivarès, C. P. 1998. Microsporidian Encephalitozoon cuniculi, a unicellular eukaryote with an unusual chromosomal dispersion of ribosomal genes and a LSU rRNA reduced to the universal core. Nucleic Acids Research 26(15):35133520.

Pigliucci, M. 2001. Phenotypic plasticity. In Evolutionary ecology concepts and case studies. Fox, C. W., Roff, D. A., Fairbainn, D. J. (eds.). Oxford University Press, Oxford, U.K. 424 pp.

Poulin, R. 1995. Phylogeny, ecology and the the richness of parasite communities in vertebrates. Ecological Monographs 65(3):283-302.

Poulin R. 1998. Evolutionary ecology of parasites. Chapman \& Hall, London, UK. 212 pp.
Pomport-Castillon, C., Romestand, B. and de Jonckheere, J. F. 1997. Identification and phylogenetic relationships of microsporidia by riboprinting. Journal of Eukaryotic Microbiology 44(6):540-544.

Pomport-Castillon, C., De Jonkheere, J. F. and Romestand, B. 2000. Ribosomal DNA sequences of Glugea anomala, G. stephani, G. americanus and Spraguea lophii (Microsporidia): phylogenetic reconstruction. Diseases of Aquatic Organisms 40:125-129.

Refardt, D., Canning, E. U., Mathis, A., Cheney, S. A., Lafranchi-Tristem, N. J. and Ebert, D. 2002. Small subunit ribosomal DNA phylogeny of microsporidia that infect Daphnia (Crustacea: Cladocera). Parasitology 124:381-389.

Sánchez, J. G., Speare, D. J., Markham, R. J. F. and Jones, S. R. M. 2001a. Experimental vaccination of rainbow trout against Loma salmonae using a live low-virulence variant of L. salmonae. Journal of Fish Biology 59:442-448.

Sánchez, J. G., Speare, D. J., Markham, R. J. F. and Jones, S. R. M. 2001b. Isolation of a Loma salmonae variant: biological characteristics and host range. Journal of Fish Biology 59:427-441.

Sang, T., Crawford, D. J. and Stuessy, T. F. 1995. Documentation of reticulate evolution in peonies (Paeonia) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. Proceedings of the National Academy of Sciences USA 92:6813-6817.

Shaw, R. W. and Kent, M. L. 1999. Fish microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 418-446.

Shaw, R. W., Kent, M. L., Docker, M. F., Brown, A. M. V., Devlin, R. H. and Adamson M. L. 1997. A new species of Loma (Microsporea) in shiner perch (Cymatogaster aggregata). Journal of Parasitology 83(2):296-301.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 1998. Modes of transmission of Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 33(2):151-156.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000a. Innate susceptibility differences in chinook salmon Oncorhynchus tshawytscha to Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 43:49-53.

Shaw, R. W., Kent, M. L., Brown, A. M. V., Whipps, C. M. and Adamson, M. L. 2000b. Experimental and natural host specificity of Loma salmonae (Microsporidia). Diseases of Aquatic Organisms 40:131-136.

Shaw, R. W., Kent, M. L. and Adamson, M. L. 2000c. Viability of Loma salmonae (Microsporidia) under laboratory conditions. Parasitology Research 86:978-981.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. and Ishikawa, H. 2000. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. Nature 407:81-86.

Sites, J. W. and Crandall, K. A. 1997. Testing species boundaries in biodiversity studies. Conservation Biology 11(6):1289-1297.

Speare, D. J., Beaman, H. J., Jones, S. R. M., Markham, R. J. F. and Arsenault, G. J. 1998a. Induced resistance in rainbow trout, Oncorhynchus mykiss (Walbaum), to gill disease associated with the microsporidian gill parasite Loma salmonae. Journal of Fish Diseases 21(2):93-100.

Speare, D. J. and Daley, J. 2003. Failure of vaccination in brook trout Salvelinus fontinalis against Loma salmonae (Microspora). Fish Pathology 38(1):27-28.

Tamas, I., Kasson, L., Canback, B., Naslund, A. K., Eriksson, A. S., Wernegreen, J. J., Sanstrom, J. P., Moran, N. A. and Andersson, S. G. E. 2002.50 million years of genomic stasis in endosymbiotic bacteria. Science 296:2376-2379.

Terry, R. S., Smith, J. E. and Dunn, A., M. 1998. Impact of a novel feminizing microsporidian parasite on its crustacean host. Journal of Eukaryotic Microbiology 45:497-501.

Vávra, J. and Larsson, J. I. R. 1999. Structure of the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 7-84.

Vernick, S. H., Sprague, V. and Krause, D. 1977. Some ultrastructural and functional aspects of the golgi apparatus of Thelohania sp. (Microsporida) in the shrimp Pandalus jordani Rahbun. Journal of Protozoology 24:(1):94-99.

Vivarès, C. P. and Méténier, G. 2000. Towards the minimal eukaryotic parasitic genome. Current Opinion in Microbiology 3:463-467.

Weiss, L. M. and Vossbrinck, C. R. 1999. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In The microsporidia and microsporidiosis. Wittner, M. and Weiss, L. M. (eds.). American Society for Microbiology Press, Washington, D.C. p. 129-171.

Wheeler, Q. D. and Meier, R. 2000. Species concepts and phylogenetic theory. Columbia University Press. New York, NY. 230 pp.

Wheeler, Q. D. and Platnick, N. I. 2000. The phylogenetic species concept (sensu Wheeler and Platnick). In Species concepts and phylogenetic theory. Wheeler, Q. D. and Meier, R. (eds.). Columbia University Press. New York, NY. 230 pp.

Wiley, E. O. and Mayden, R. L. 2000. The evolutionary species concept. In Species concepts and phylogenetic theory. Wheeler, Q. D. and Meier, R. (eds.). Columbia University Press. New York, NY. 230 pp.

Williams, B. A. P., Hirt, R. P., Lucocq, J. M. and Embley, M. T. 2002. A mitochondrial remnant in the microsporidian Trachipleistophora hominis. Nature 418:865-869.

Zietara, M. S. and Lumme, J. 2002. Speciation by host switch and adaptive radiation in a fish parasite genus Gyrodactylus (Monogenea, Gyrodactylidae). Evolution 56(12):2445-2458.

Figure 7.1: Phylogeny of 10 Loma species examined in this study showing corresponding host, morphological, developmental, pathological, and ecological character states represented with $+/-$ signs. Shaded columns (host group, xenoma size, and number of tubules per vesicle) indicate characters that follow a pattern consistent with phylogenetic relationships at the intrageneric level. Un-shaded columns do not show an obvious pattern with respect to the phylogenetic tree. Blanks represent unknown data. The + and - categories were designated on either side of an arbitrarily drawn line, to simplify representation of the data. Characters with highest intraspecific variance and overlap between species (see Chapter 3) are indicated with asterisks at the bottom of the table. Species name equivalences across chapters are given in Appendix 12, and host authorities are given in Table 1.1 of Chapter 1. Morphological data were from Chapters 2 to 4 as well as Morrison \& Sprague (1981a; c) and Shaw et al. (1997) and Loma sp. BRO was compared to morphology of L. fontinalis of Morrison \& Sprague (1983). Tree was reconstructed from a single, reference sequence for each species consisting of 1113 to 1846 bp of SSU, ITS and LSU rDNA using a heuristic ML search in PAUP*. Character categories were as follows: ho $=$ host group,$+=$ gadids, $-=$ non-gadids; $\mathbf{x s}=$ mean xenoma size,$+=>75 \mu \mathrm{~m}$, $-=<75 \mu \mathrm{~m} ; \boldsymbol{s s}=$ mean spore length fresh or converted using shrinkage factors, $+=>5 \mu \mathrm{~m},-$ $=<5 \mu \mathrm{~m} ; \mathrm{pf}=$ mean number of polar filament turns, $+=16$ or more,$-=<16 ; \mathrm{dt}=$ relative developmental timing of parasitophorous vacuole,$+=$ earlier, $-=$ later; tv $=$ mean number of tubules per vesicle, $+=>5,-=5$ or fewer; ts $=$ mean number of tubules per sporoblast or spore $,+=>40,-=<40 ;$ vs $=$ mean vesicle size $,+=>0.4 \mu \mathrm{~m},-=<0.4 \mu \mathrm{~m} ; \mathrm{vp}=$ mean number of vesicles per spore, $+=>5,-=<5 ; \mathrm{vb}=$ mean number of vesicles per sporoblast, $+=$ $>3,-=<3 ; \mathbf{s p}=$ number of spores per parasitophorous vacuole $,+=4,-=2$ or $1 ;$ wa $=$ xenoma wall plasmalemma, $+=$ undulating, $-=$ smooth; pa $=$ degree of host pathological response to heavy infection, $+=$ moderate to high, $-=$ few or no signs of host response observed; $\mathbf{p v}=$ prevalence of infection, $+=>10 \%,-=<10 \%$.


## APPENDICES

## APPENDIX 1

## Explanation of three software programs designed in this study.

## Introduction

Ribosomal RNA genes possess features that cause their evolution to be characterized by regions with large numbers of insertions or deletions (indels) interspersed with regions that evolve under strong selection (conserved regions) (see Nunn et al., 1996; Otsuka et al., 1997). Processes of ribosomal RNA (rRNA) folding into secondary and higher-level structures, protein interactions with loop regions, and evolution by slippage and recombination among multiple copies, are all considerations that can be incorporated in alignment and analysis of rDNA in phylogenetic studies. These processes and considerations with respect to phylogenetic methods have been presented elsewhere (Hancock \& Dover, 1988; Ragan, 1988; Hancock \& Dover, 1990; Hillis \& Dixon, 1991; Linares et al., 1991; Rousset et al., 1991; Drouin \& de Sá, 1995; Hancock, 1995; Katiyar, 1995; Kjer, 1995; Nunn et al., 1996; Waters \& Schaal, 1996; Winnepenninckx \& Backeijau, 1996; Buckler et al., 1997; Otsuka et al., 1997; Otsuka et al., 1999; Syvanen, 2002). In the present study, three software programs were designed to be used as tools during analysis of rDNA sequences, and their use on rDNA from microsporidia (small, parasitic eukaryotes with fast-evolving rDNA) is demonstrated.

The first program was developed as a tool for extracting correct DNA sequences from a stretch of double peak signal in DNA sequence results (an automated sequencer "trace" or "chromatograph") when the double signal results from amplification of two copies of a gene that differ by an indel. Such double peaks can occur with direct sequencing of polymerase chain reaction (PCR) products. For multi-copy genes, like rDNA, both amplified products may be correct (as opposed to contamination) and so both copies may also be of interest.

The second program was developed to examine phylogenetic information contained in indels or alignment gaps among species. Insertions or deletions (indels) have been long known to be phylogenetically informative mutations (see review in Giribet \& Wheeler, 1999), but are not widely used phylogenetic analyses, and are instead usually removed from the data by "gap stripping". Gap stripping involves removal of all positions (columns of nucleotides in an alignment) containing a gap in one or more species. This practice may be perfectly acceptable in many instances, for example, for most proteins or highly diverged species, where few indels exist or where indels-containing regions are not important for the analysis. However, in cases
where gaps occur frequently in the data (i.e. at a large proportion of the positions), or where gaps occur at most variable positions (i.e. at most informative positions), or where gaps themselves are the dominant or only phylogenetically informative signal, gap stripping can remove all the information of interest. Gaps are omitted from data because it is impractical to incorporate models of gap evolution into the common models used for nucleotide substitution evolution for aligned DNA data, although models involving gaps (e.g. gap costs and gap expansion costs) are widely used in software that optimizes alignments. The program developed in this study uses a simple model of gaps that assumes gap-end locations are informative (as opposed to gap length), and creates from an alignment block a gap weight matrix that can be used with arbitrary weights or encoded as nucleotide data.

The third program was developed as a tool for exploring the influence of rRNA folding constraints (stems) on phylogeny reconstruction. Ribosomal RNA folding structure (secondary structure) has been shown to impose constraints on paired nucleotides in stems (Rousset et al., 1991; De Rijk et al., 1994; Nunn et al., 1996; Otsuka et al., 1997; Otsuka et al., 1999), such that pairs in stems occur in characteristic locations and undergo predominantly substitutions that maintain the pairing structure, or stability or shape of the stem. Where a mutation has caused a disrupted pairing state (e.g. $C$ with $G$ becomes $G$ with $G$ ) there is thought to be strong selection for a substitution that returns the pair to a more stable bond (e.g. a G becomes C ), known as a compensatory substitution. The models of substitution used most widely for phylogenetic analyses do not consider these aspects of selection (constraint) on nucleotide substitutions in stems, and therefore may be less realistic models for rDNA analysis. The program developed here uses a model of stem pairing that allows the user to explore different relative constraints on rDNA stem pairing with respect to phylogenetic analysis. The relative costs of transitions, transversions or disruptions are modified by the user to produce alternative weighting for changes in stems according to the model, and the result is formatted as a PAUP* (Swofford, 2000) input file.

## Software language

Software was written (according to program designs, below) in Java ${ }^{\text {TM }} 2$ (JDK 1.2 or J2SE) by Michael Coury (D-Wave Systems Inc.). Java was chosen because of its relative simplicity for programmer and user, and because it can be used on any computer (MAC, Windows, UNIX etc.) using an appropriate, free, downloadable Java byte code interpreter.

## Design of Flip Analyzer program

If primers for PCR amplify two copies of a gene that differ by an indel, and if the primers enclose the indel, direct sequencing from such PCR products (without cloning) may produce sequences that have two different nucleotide signals (double peaks) at almost every position. Figure A1.1 shows a sequence trace (or chromatograph) produced from an ABI Automated Sequencer with this sort of signal. Unfortunately, in cases of double signal, such as that shown in the bottom portion of the sequence in Fig. A1.1, the two copies of the gene cannot be extracted simply by recording the higher signal peaks as representing "copy 1 " and the lower signal as "copy 2 ". The reason for this is that higher peaks do not correspond to a single gene copy (and lower peaks do not correspond to the second copy) because, typically, both copies of the gene are amplified at about the same overall signal strength (or peak amplitude) while peak amplitude typically varies significantly from site to site (among positions) in each gene copy due to stochastic effects during the sequencing reaction (i.e. Taq polymerase enzyme activity and stop-terminator incorporation tend to be uneven across the length of sequence and may result in uneven distributions of products of each length, perhaps partly due to secondary structure or other chemical interactions affecting Taq polymerase binding). Hence, peak amplitude often varies as much or more within a copy of the gene as it does between the two gene copies. Thus, the peaks from the copy 1 will be, at some sites, higher than those from copy 2 , while at other sites, lower than those of copy 2. The result is that without knowing which copy of the gene a peak corresponds to, it is difficult or impossible to extract the correct sequences. However, when the sequence of one or both genes is known, the correct sequences can be extracted manually by using the known sequence as a guide and then choosing the higher or lower peak that matches the nucleotide in the guide sequence for each position. The remaining set of peaks (some higher and some lower) that did not match the guide sequence and were not chosen will represent the other known or unknown sequence. If neither sequence of the gene is known, but the sequences are identical except for a difference of an indel, the process of choosing the higher or lower peak still works if the identity between the two signals is determined on a site-by-site basis by swapping higher and lower peaks as needed, using the immediately adjacent signal in one or the other sequences as a guide (see details below).

Thus, the idea for the Flip Analyzer program arose from the observation that swapping ("flipping") the higher and lower peaks at appropriate positions according to a prediction that they should be identical except for an indel. The program can take two rows of nucleotides that appear highly dissimilar and will produce two rows of nucleotides that are identical except for
an indel, in which the top row represents one copy of the gene and the bottom row represents the other copy of the gene. The basic procedure used by the Flip Analyzer is as follows:

Given a stretch of double signal in which site 1 (the first position) has peaks representing C and G (where the first letter is the taller peak), site 2 has G and A, site 3 has G and A, site 4 has $G$ and $A$, site 5 has $A$ and $T$, one can manually extract two identical sequences that differ by a one-site shift, as follows:
site 12345
top peaks CGGGA
bottom peaks GAAAT
Write a diagonal dash with a shift of one position between matched letters in the top and bottom rows and a horizontal dash between adjacent letters that match, as follows:

```
        site 1 2 3 4 5
    top peaks C G-G-G A
        / /
bottom peaks G A-A-A T
```

Swap the letters on top and bottom that are not connected by a diagonal dash, i.e. swap letters at site 3 (shown in gray):


Similarly, for a longer sequence with any size of indel, from a double signal like this:

$$
\begin{array}{lll} 
& \text { sites } & 12345678 \text { etc. } \\
\text { e.g. top peaks } & 1 . & \text { ATGGGGGTTTTTACACGCGGCTGTGTCTCCTGTG } \\
\text { bottom peaks } 2 . & \text { ATGGGTGAGAACACTCTGGGGGCTCTCTTTTCCG }
\end{array}
$$

Identify positions to swap:

$$
\begin{array}{lll}
\text { top peaks } & \text { 1. ATGGGGTTTTTACACGCGGCTGTGTCTECTGTG } \\
\text { bottom peaks } 2 . & \text { ATGGGTGAGAACACGCTGGGGGCTETCTRTTCEG }
\end{array}
$$

Swap top and bottom letters (notice identical sequence, underlined):

1. ATGGGGMTTACACMCGGGGCTGTETCTMCTGCG
2. ATGGGGGTGATACACTGGGGGCTGTCTCTTCMG

Recover two identical sequences (with a 2 bp indel) by sliding and inserting dashes:

## extracted sequence 1 ATGGG--TGATTACACTCGGGGCTGCTCTTCTGCG extracted sequence 2 <br> ATGGGGGRGATTACACTCGGGGCTGTCTCTTCTG

Flip Analyzer uses a similar procedure for a range of indel sizes (both positive and negative shifts) chosen by the user.

## Demonstration of Flip Analyzer

Figures A1.2 through A1.9 show the use of Flip Analyzer in REALEM on real data. To use the program, top and bottom signal peaks from the double-signal region of a sequence trace must be entered by hand as a text file. Figure A1.2 shows the first menu, in which the user chooses the program. Figure A1.3 shows the second menu, in which the user may open the input file from its location on the computer. Figure A1.4 shows a typical input file in which top and bottom rows of letters look dissimilar from one another. Figure A1.5 shows the simple, single menu choice for analysis of the input file. Figure A1.6 shows the menu that pops up prompting the user to give a "shift" range, which refers to the hypothetical indel size, for example, one or two base pairs. The program will analyze the data for all integer sizes of indels below the number selected in both directions, for example, if the user selects " 10 " the program gives analyses for $-10,-9,-8 \ldots$ etc. to 10 . It then uses the Flip procedure explained (above) in which it swaps top and bottom letters at any positions necessary to produce a match the letters in the top and bottom rows taking the shift size and direction into account. The result for shift $=1$ is shown in Fig. A1.7. Asterisks show positions that were flipped. As the program is running (flipping top and bottom letters to make matches), if it determines that a match cannot be made, it proceeds to the next position, leaving the mismatch, and continues analyzing the remaining sequence. The output file has the indel inserted at the start of the sequence. Mismatches between the final sequences are represented by exclamation marks show positions that do not match (are not identical) in the resulting sequences. The "\# errors" refers to the number of exclamation marks, and can be used as an indicator of the likelihood that the result is correct. This result is full of "!"s, which indicates these sequences do not match, given this indel size (i.e. the correct sequences were not recovered). Normally one would analyze a wider range of indel sizes, for example Figs. A1.8 and A1.9 show the results for shift $=10$ for the same input data. Figure A1.9 shows the summary for this example, in which the number of mismatches (\# errors) is very high in all but two shifts. Here, the lowest error score (eight errors) resulted from a shift of -6 . This result would strongly suggest these sequences were, in fact two copies of a
gene with a shift of six base positions caused by an indel. In this case, the eight error positions were re-examined manually on the original sequence trace, and peaks were either not entered correctly (typos) or were low and so were missed. Incorrect results can be recognized by the large number of mismatches ( $40-100 \%$ ) between the two sequences, whereas correct results from real data tested in this study tended to produce about 2 to $15 \%$ mismatches. Typos or overlooked low peaks were generally greater for weaker overall signals. If mismatches remain after sequences are re-examined for typos or low peaks, these may be valid polymorphisms (cases of two or more alleles in the population) in the data. In such cases it is not possible to determine to which sequence (top or bottom) these differences belong to, so for more than one mismatch between two sequences, polymorphisms have to be indicated with degenerate code.

Flip analyzer in the REALEM software package successfully retrieved two sequences differing by an indel for over 30 microsporidian DNA isolates (presented in this thesis). In many cases, cloning of the same PCR products revealed that these indels were valid, reflecting polymorphisms in the data.

## Design of Gap Matrix program

If alignments contain many gaps or if gaps occur in regions with essential phylogenetic information for resolving taxa, gap stripping may remove important phylogenetic signal. The Gap Matrix program encodes gap information as characters that can then be analyzed by maximum parsimony or other optimality criteria. The idea to use gap ends as the only gap information in Gap Matrix arose from the recognition that gap length is too ambiguous, because larger gaps may occur from repeated insertion or deletion (indel) events over time, or may be due to a single indel event. In contrast, shared gap end positions most likely result from shared indel events.

The algorithm first finds positions at the ends of gaps in a block of aligned DNA, and then it replaces these positions with a " 0 ". Next, positions in which no taxa have a gap end (now coded as " 0 ") are removed. Remaining positions have a gap end in some species, nucleotides in other species, and middles of gaps (a dash) in other taxa. Gap ends are then converted to "C", and all nucleotides are converted to " T ". Finally, the data are converted to PAUP* format with the original DNA block followed by the C and T matrix of gap ends. This program also allows a variety of modifications of these steps. For example, the first step can be made to encode the full length of gaps, if desired. Another modification that can be made if desired, is to remove positions in which some species have a middle of a gap (i.e. still have a
dash character), or to encode the dash character as " N " or " T " at this step. The final C and T matrix may also be converted to any other symbol (e.g. zero and one) and may be weighted, as desired. Additionally, the gap matrix can be analyzed without the original nucleotide data.

## Demonstration of Gap Matrix

Figures A1.10 to A1.19 illustrate the application of Gap Matrix in REALEM to microsporidian rDNA. Figure A1.10 and A1.11 show the first menus in which the user selects the Gap Matrix program, and then opens an appropriate input file. Figure A1.12 shows an opened input file, which is any alignment block for a number of taxa (up to 10000 ) with sequence (up to 100000 positions) including gaps denoted with dash characters. Figure A1.13 shows the menu choices for analyses in this program, in which the user may run all steps of the program without further prompting, or may chose to run each step separately and make modifications as suggested in the Materials and Methods. Figure A1.14 shows the result of the first step, in which gap ends are encoded with " 0 ". The next menu choice available will convert the alignment block to a gap matrix, as shown in Fig. A1.16. The resulting gap matrix shown in Fig. A1.17 has had all positions without gap ends removed. Figure A1.18 shows the result when the gap ends in this matrix are converted to Cs and the nucleotides are converted to Ts. These letters are arbitrarily chosen, and can be changed to symbols or other nucleotide letters, as desired. The original data block and the gap matrix can be placed together in PAUP* format by using "File", "Format for PAUP" or F6, as shown in Fig. A1.19.

The gap matrix program successfully produced gap matrices for microsporidian rDNA in this thesis and elsewhere (Brown \& Adamson, 2000; Matthews et al., 2001; Brown, 2002;

Brown \& Kent, 2002).

## Design of Stem State program

Ribosomal DNA sequences are often aligned using rRNA secondary structural models from the rDNA structural database (Van de Peer et al., 2000) prior to phylogenetic analysis. These hypothesized secondary structures must be encoded in the data to be recognized by the Stem State program. Secondary structures were inferred using a comparative approach, by alignment with relatives that have well characterized secondary structures in the database, and also with the help of free energy minimization predictions from software available on-line (e.g. mfold), and basic assumptions about canonical (and sometimes non-canonical) base-pairing rules. While predicted rRNA secondary structures may be incorrect for species that are not closely related to those that are known empirically, many parts of the rRNA structure have been
shown to be highly conserved across all eukaryotes and some parts even across prokaryotes. As a result, some portions of the secondary structure (e.g. several universally conserved stems) can be predicted with great confidence, while others (e.g. some highly variable terminal stem-loops) cannot be predicted with as much certainly. Arguably, those portions of the structure that can only be predicted with lower confidence (i.e. not as universally conserved) should not be included or should be down-weighted, depending on the degree of confidence in the structure during analyses. Not including or down-weighting such regions is easily

For the Stem State program, stems are indicated by using a guide row above the first line of aligned sequence with square brackets to indicate stems. Nucleotides that are to be paired in stems are indicated with upper case letters, while nucleotides that are not paired are indicated with lower case letters. In the guide line, stems are numbered consecutively, with the first and second halves of each stem that will be paired up labelled with "a" and "b", respectively, after the stem number, as follows:

```
Stems [1a ] [2a ] [1b ] [3a 3b]
sp.1 [CAGGT]ga[TTC]t[GCCTG]at[GTaTGCGCGgcctttaagCGCGCAAC]...etc.
sp.2 [CAGGT]ga[TTC]t[ACCTG]aa[GTcTGTGAAttc---gttTTTACAAC]...etc.
```

This example demonstrates the fact that stems may be separated by other stems (e.g. stem 1) or they may form terminal stem-loops (e.g. stem 3), and notation for the latter can be simplified in the program by including both " 3 a " and " 3 b " labels within one set of brackets.

The program then matches up paired nucleotides and lists these, in order, from the proximal to the distal ends of each stem, as follows:

```
STEM 1
    1 2 3 4 5
sp.1 CG AT GC GC TG
sp.2 CG AT GC GC TA
STEM 2
    1 2 3
sp.1 TA TA CG
sp. 2 TG TA CG
STEM 3
    1
sp.1 GC TA TA GC CG GC CG GC
sp.2 GC TA TA GC TA GT AT AT
etc.
```

The program then encodes all 16 possible unique "stem states" using a single-letter symbol (A. through P , except that N is reserved to represent and unknown, so Z is used in its place):

```
encoded: A B C D D E F G H I I J K L L M I Z O
original: GC GT GA TA AT TG AG CG CC GG TT AA CA CT AC TC
```

The program re-sorts the encoded data into an alignment block, listing the letters in order by stem and stem position (proximal to distal end), for example:

```
stem#1 2 3 ...etc.
    1234512312345678...etc.
sp.1 HEAAFDDHADDAHAHA ..etc.
sp.2 HTAADFDHADDADBEE ..etc.
```

A model of rDNA stem constraints can analyze this encoded stem state data; in this case, the model was developed partially based on those of Rousset et al. (1991), Nunn (1992), and Otsuka et al. (1997). Such studies presented models that included only canonical pairs (G with C, A with T) and rarely the most common non-canonical pair that is know to be common in rRNA ( G with T ), whereas the often observed G with A was not included in those models. Those studies also omit five other pairs ( T with C , A with C , two As, two Ts, two Cs or two Gs) that are observed in rRNA, but which will be more disruptive in stems (i.e. will form a bulge rather than a stable bond). These more rare and less stable states both structurally and in evolution are important to the model of stem transitions if they occur at some reasonable frequency in nature. Therefore, in this study, a model was developed to include all 16 stem states (including the disruptive ones) and propose that paths through the model (changes in state) must happen by single substitutional events, as illustrated in Fig. A1.20.

In Figure A1.20 the four canonical and four most common non-canonical stem states are indicated in the center, while the disruptive states are shown outside the central network of transitions. Costs of changes between states can be added along arrows between stem states. In Figure A1.20, only $\alpha=$ transition and $\beta=$ transversion are shown, however, an additional cost of $\delta=$ disruption is applied in the model when passing to the outer states, while its reciprocal, $\rho=$ return from disruption, is added when passing from the states on the outside of the network, to the inside.

The program prompts the user to supply relative transition ( $\alpha$ ) and transversion ( $\beta$ ) frequencies that will be applied to the creation of a weight matrix for phylogenetic analysis
using parsimony in PAUP*. The Stem State program will calculate disruption ( $\delta$ ) and return from disruption ( $\rho$ ) costs from the data. Disruption is calculated as 8 x the total number of letters in the single-letter encoded stem state data block, divided by the sum of each letter representing a disrupted state (e.g. I, J, K, L, M, Z, O, P). Return from disruption is calculated as a simple reciprocal of the previous value. The program also calculates frequencies of each type of canonical and non-canonical pair, for interest, as follows:
$\phi=\mathrm{F}$ and B (TG and GT) - calculate: total $/ \mathrm{F}+\mathrm{B}$
$\eta=H$ and $A(C G$ and GC) - calculate: total / H $+A$
$\varepsilon=\mathrm{E}$ and D (AT and TA) - calculate: total $/ \mathrm{E}+\mathrm{D}$
$\chi=\mathrm{C}$ and $\mathrm{G}(\mathrm{GA}$ and AG$)-$ calculate: total $/ \mathrm{C}+\mathrm{G}$
Figure A1.21 shows the matrix upon which weights can be generated. Letters "a" and "b" refer to transition and transversion frequencies, respectively, as input previously by the user. Letters "d" and " $r$ " refer to disruption and return from disruption costs. These formulae in the matrix of Figure A1.21 were obtained by adding the shortest path in the model shown in Fig. A1.20. The Stem State program then applies the formulae to create a numerical matrix of weights for transitions from stem states on the left to stem states listed in the top row, such that an output weight matrix would appear as follows:

|  |  | A | B | C | D | E | F | G | H | I | J | K | L | M | Z | O | P |
| :--- | :--- | ---: | ---: | ---: | ---: | ---: | :---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
|  |  | GC | GT | GA | TA | AT | TG | AG | CG | CC | GG | TT | AA | CA | CT | AC | TC |
| A | GC | 0 | 4 | 6 | 12 | 8 | 16 | 14 | 20 | 40 | 40 | 44 | 46 | 46 | 44 | 40 | 40 |
| B | GT | 4 | 0 | 6 | 12 | 4 | 16 | 10 | 16 | 44 | 40 | 40 | 44 | 46 | 40 | 44 | 44 |
| C | GA | 6 | 6 | 0 | 6 | 10 | 10 | 16 | 14 | 46 | 40 | 46 | 40 | 40 | 46 | 46 | 46 |
| D | TA | 12 | 12 | 6 | 0 | 16 | 4 | 10 | 8 | 52 | 46 | 40 | 40 | 46 | 48 | 52 | 40 |
| E AT | 8 | 4 | 10 | 16 | 0 | 12 | 6 | 12 | 52 | 46 | 40 | 40 | 52 | 40 | 40 | 48 |  |
| F | TG | 16 | 16 | 10 | 4 | 12 | 0 | 6 | 4 | 44 | 40 | 40 | 44 | 44 | 44 | 46 | 40 |
| G AG | 14 | 10 | 16 | 10 | 6 | 6 | 0 | 6 | 46 | 40 | 46 | 40 | 46 | 46 | 40 | 46 |  |
| H | CG | 20 | 16 | 14 | 8 | 12 | 4 | 6 | 0 | 40 | 40 | 44 | 46 | 45 | 45 | 45 | 45 |
| I | CC | 5 | 9 | 11 | 17 | 13 | 9 | 11 | 5 | 0 | 45 | 49 | 53 | 45 | 45 | 45 | 45 |
| J | GG | 5 | 5 | 5 | 9 | 9 | 5 | 5 | 5 | 45 | 0 | 45 | 45 | 45 | 45 | 45 | 45 |
| K | TT | 9 | 5 | 11 | 5 | 5 | 5 | 11 | 9 | 49 | 45 | 0 | 45 | 45 | 45 | 45 | 45 |
| L AA | 11 | 9 | 5 | 5 | 5 | 5 | 5 | 11 | 53 | 45 | 45 | 0 | 45 | 45 | 45 | 45 |  |
| M | CA | 11 | 11 | 5 | 5 | 13 | 9 | 11 | 11 | 45 | 45 | 45 | 45 | 0 | 45 | 51 | 49 |
| Z CT | 9 | 5 | 11 | 13 | 5 | 9 | 11 | 9 | 45 | 45 | 45 | 45 | 45 | 0 | 49 | 49 |  |
| O AC | 5 | 9 | 11 | 13 | 5 | 11 | 5 | 11 | 45 | 45 | 45 | 45 | 51 | 51 | 0 | 45 |  |
| P | TC | 5 | 9 | 11 | 5 | 13 | 5 | 11 | 9 | 45 | 45 | 45 | 45 | 45 | 49 | 45 | 0 |

To obtain integer values, the program rounds out calculated frequencies to the nearest 0.25 to 0.01 (specified by the user), and applies a factor of 4 to 100 to obtain step matrix values.

Relatives costs of disruptions versus substitutions (transitions or transversions) can be varied by the user by entering lower or higher transition ( $\alpha$ ) and transversion ( $\beta$ ) frequencies, for example, if the calculated disruption level is 8.5 , the user may choose to examine the effects of transitions being $1 / 10$ as costly as disruptions (enter $\alpha=0.85$ ) or $1 / 100$ as costly (enter $\alpha=$ 0.085 ), or $1 / 2$ as costly (enter $\alpha=4.25$ ).

## Demonstration of Stem State

Figures A1.22 through A1.31 show the use of Flip Analyzer in REALEM on microsporidian rDNA data. Figures A1.22 and A1.23 show the input data and menu option for choosing either a full analysis or a step-by-step approach. As with Gap Matrix, it may be desirable to use parts, but not all of the features of this program. Figure A1.24 shows the results of the first step of the program, which trims any characters that are not capitalized. If the number of upper case letters or stems contains errors, appropriate error messages are produced to indicate this before the program proceeds. Figure A1.25 shows the results from the second step in the "Processes" menu, after nucleotides have been paired, and Fig. A1.26 shows the menu choice to convert the paired stem data to encoded single-letter stem state data. Figure A1.27 shows the resulting data block of stem state data, in which " N " may be required for missing stems in some taxa or for missing sequence data. Figures A1.28 and A1.29 show the menu options that appear next, prompting the user to enter transition and transversion ( $\alpha$ and $\beta$ ) values and choose a degree of resolution for the weight matrix calculations. As with the Gap Matrix program, the data may be formatted for PAUP* using the menu shown in Fig. A1.31, and the results will be displayed as in Fig. A1.30, with frequencies shown for various stem states, etc.

The Stem state program was applied to microsporidian DNA examined in the thesis and these results were presented elsewhere (Brown \& Adamson, 2000; Brown, 2002), but were not included in this thesis due to the complexity and considerations of the model, which shall be addressed elsewhere.

## Help files

Figures A1.32 and A1.33 show the available help file menu choices with sample files, and the author and version information for the software package.

## Discussion of program applications

Application of the Flip analyzer and Gap matrix programs were demonstrated by their use in Chapters 2-6 of this thesis and elsewhere (Brown \& Adamson, 2000; Matthews et al.,

2001; Brown \& Kent, 2002) while application of the Stem state program was presented in Brown \& Adamson (2000). Further uses of these programs and their relationships to other problems or other programs are discussed here.

## Further applications of Flip Analyzer

This program can be applied to any situation where two similar copies of a gene in DNA or protein are sequenced simultaneously and produce a double signal, as long as one copy is longer than the other (i.e. there is a shift difference between them). The shift may occur anywhere in the sequence (e.g. even upstream of the sequence seen). It can be applied in cases of simple indels or for genes containing introns of different size. If three or four signals occur at most sites, or if copies having multiple indels exist (e.g. a double signal becomes triple after a given point), the program can only be used if one or two sequences are removed first. This situation would be applied for a case when the 3rd or 4 th sequences are known. Sequences can be removed easily by using a guide sequence that matches the supposedly known sequence and employing the "flip" idea, by swapping any of the three or four stacked letters. This flip idea can be applied to situations where any two genes are sequenced together where one is known, such as for contamination with a known DNA sequence. For example, in a few cases during this study, the double signal arose from two separate, unrelated gene sequences that were shown by Blast (NCBI) searches to be contaminant DNA. In one such case, the contaminant DNA was from a spillage of DNA from an adjacent lane in the sequencing gel (a plant gene sequenced by a different researcher. In these situations the Flip analyzer program will produce high error numbers for all indel sizes, however, if one or more sequences are known, the "flip" process can be performed by hand to extract the two sequences.

In future, the program may be modified to give an option for the user to input a known sequence to use as a guide for extracting a contaminating sequence, where this occurs. The present version of the program suggests whether the double sequence is likely to be due to contamination, but contaminating sequences must be removed by using the flip algorithm manually.

## Further applications of Gap Matrix

Aligned sequences that are different in length as a result of indel differences require the insertion of dashes or other characters to indicate gaps. Phylogenetic analysis programs do not generally have an option to deal with gaps other than by treating them as unknown characters, equivalent to any nucleotide (degenerate code N ). This treatment of gaps may be misleading
(Giribet \& Wheeler, 1999), while the alternative, removal of all positions in the alignment with gaps (gap-stripping), can remove useful phylogenetic information. For example, in an alignment SSU rDNA from 24 species of microsporidians, about 300 gaps were required (Brown \& Adamson, 2000) with both high gap costs and low gap costs using alignment software (ClustalW, Thompson et al., 1994) or using by-eye alignment. The use of Gap matrix in REALEM easily produced gap matrices that were used in phylogenetic analysis with the PAUP* (Swofford, 2000) package. Brown and Adamson (2000) showed that gaps in microsporidian SSU rDNA, when analyzed alone without any nucleotide substitutional data produce a tree with similar topology to others based on nucleotide data, suggesting gaps could be informative. Furthermore, there were sufficient phylogenetically informative sites from gap data alone to produce fairly high bootstrap support for many nodes in that tree.

Supporting documentation for Swofford's (2000) PAUP* phylogenetic analysis package suggests that gaps (dashes) can be encoded as a matrix of zeros and ones, but does not suggest how the length of gaps are to be encoded. For example, it is not recommended (Giribet \& Wheeler, 1999; Swofford, 2000) that each position in a long indel be given a character weight of one, as the long indel likely arises from a single or very few events, rather than a number of events equal to the length of the gap. Swofford's (2000) documentation suggested the alignment could be broken into codon-sized blocks (blocks of three letters per column) and given a gap value of one (equivalent to one nucleotide) for each three positions including a gap. These results may be appropriate for protein sequences, but not rDNA sequence that may produce gaps of any length. The results of Brown and Adamson (2000) using the Gap matrix program suggest that encoding only gap end positions may be suitably realistic for rDNA, as analyses from gap end data alone yielded a tree that was consistent with trees from gap stripped data. This is the first program to encode gap ends for analysis in PAUP*, as far as I know.

After this program was designed, others have published a similar program (Cheynier et al., 2001) to encode gap information. Cheynier et al. (2001) developed a program called Indelstack, which differs from Gap matrix in several ways. The latter program encodes only the $3^{\prime}$ ' end of each gap, and therefore is similar to Gap matrix in that gap positional differences are used but gap length is ignored. Cheynier et al. (2001) demonstrated that gap end positions and nucleotide data produced similar trees for SIV virus data, confirming the results by Brown and Adamson (2000), which suggest gap ends are informative and suitable characters marking evolutionary events. Cheynier et al.'s (2001) program replaced all nucleotide data by the
predominant nucleotide at that site. This approach is different from that used in Gap matrix, and may have merit in that it may enable other phylogenetic optimality criteria (other than maximum parsimony) to be used without severely violating the assumptions of the models, such as equal base frequencies (distance) or models of substitution and base frequency specified in maximum likelihood.

Cheynier et al.'s (2001) program may be more realistic than Gap matrix in that it counts gaps only once rather than twice (for each end), however, it will miss evolutionary events because some gaps differ at the 5' end but not at the $3^{\prime}$ end among species. Thus, the Indelstack program may be more conservative than Gap matrix, but Gap matrix is able to resolving relationships with higher sensitivity.

In future, Gap matrix could be modified in several ways, for example, to convert nucleotide signal to the predominant base (as in Cheynier et al. 2001) or to a single, randomly selected nucleotide at each site, or to allow the user to down weight two-end gaps relative to one-end gaps (gaps of size one).

## Further applications of Stem State

Application of the Stem state program from REALEM produced phylogenetic trees from stem-pairing that had similar topologies to those produced from standard nucleotide analysis (Brown \& Adamson, 2000), suggesting stem-pairing according to secondary structure models may be sufficiently realistic as a model for phylogenetic analyses. Whether stem-pairing and the weighting model of this program is better than treating rDNA data in a standard linear way has not been thoroughly examined, and will have to be tested with both simulated and real data in further studies. However, the Stem state program has advantages over previously presented models (Rousset et al 1991; Nunn, 1992; Winnepenninckx and Backeijau 1996; Otsuka et al. 1997; Otsuka et al 1999) in that it incorporates all character states (e.g. disrupted and noncanonical states) that could occur, and generates an easily used PAUP*-formatted result that can be explored with different models of transition, transversion, and disruption.

Preliminary results with this program (Brown \& Adamson, 2000; Brown, 2002) suggested the treatment of rDNA secondary structure by removing unpaired loops or pairing nucleotides in stems and analyzing these pairs with rate-corrected maximum-likelihood as pairs produce differences in branch lengths, resolution, and sometimes topology of the tree, thereby suggesting stems and stem-pairing may be important to the phylogenetic outcome, as has been observed elsewhere (Otsuka et al. 1999, and references therein).

## Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada strategic grant 582073 to M. L. Adamson. I am sincerely grateful to Michael Coury for his patience and careful attention to detail in writing the programs and making them run efficiently.

## Literature cited

Brown, A. M. V. 2002. Microsporidians from the depths described by genes and other means. Program - 14th Meeting of the Society for Evolutionary Protistology. June 19-24, 2002, Vancouver, Canada. p. 4.

Brown, A. M. V. and Adamson, M. L. 2000. Microsporidian evolution doing cart-wheels: new rDNA phylogenies using gap matrices and stem models. Program guide and abstracts -75 th annual meeting of the American Society of Parasitologists. June 24-28, 2000, San Juan, Puerto Rico. p. 61.

Brown, A. M. V. and Kent M. L. 2002. Molecular diagnostics for Loma salmonae and Nucleospora salmonis (microsporidia) In Molecular diagnostics of salmonid diseases. Cunningham, C. O. (ed.). Kluwer Academic Publishers, Dordrecht p. 267-283.

Buckler, E. S., Ippolito, A. and Holtsford, T. P. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics 145:821-832.

De Rijk, P., Van de Peer, Y., Chapelle, S. and De Wachter, R. 1994. Database on the structure of large ribosomal subunit RNA. Nucleic Acids Research 22(17):3495-3501.

Drouin, G., Moniz de Sa, M. and Zuker, M. 1995. The Giardia lamblia actin gene and the phylogeny of eukaryotes. Journal of Molecular Evolution 41:841-849.

Giribet, G. and Wheeler, W. C. 1999. On gaps. Molecular Phylogenetics and Evolution 13(1):132-143.
Hancock, J. M. 1995. The contribution of DNA slippage to eukaryotic nuclear 18 S rRNA evolution. Journal of molecular evolution 40:629-639.

Hancock, J. M. and Dover, G. A. 1988. Molecular coevolution among cryptically simple expansion segments of eukaryotic 26S/28S rRNAs. Molecular Biology and Evolution 5(4):377-391.

Hancock, J. M. and Dover, G. A. 1990. 'Compensatory slippage' in the evolution of ribosomal RNA genes. Nucleic Acids Research 18(20):5949-5954.

Hillis, D. M. and Dixon, M. T. 1991. Ribosomal DNA: Molecular evolution and phylogenetic inference. The Quarterly Review of Biology 66(4):411-433.

Katiyar, S. K., Visvesvara, G. S. and Edlind, T. D. 1995. Comparisons of ribosomal RNA sequences from amitochondrial protozoa: implications for processing, mRNA binding and paromomycin susceptibility. Gene 152:27-33.

Kjer, K. M. 1995. Use of rRNA secondary structure in phylogenetic studies to identify homologous positions: an example of alignment and data presentation from frogs. Molecular phylogenetics and Evolution 4(3):314330.

Linares, R. A., Hancock, J. M. and Dover, G. A. 1991. Secondary structure constraints on the evolution of Drosophila 28S ribosomal RNA expansion segments. Journal of Molecular Biology 219:381-390.

Matthews, J. L., Brown, A. M. V., Larison, K., Bishop-Stewart, J. K. and Kent, M. L. 2001. Pseudoloma neurophilia n . g., n. sp., a new microsporidium from the central nervous system of the zebrafish (Danio rerio). Journal of Eukaryotic Microbiology 48(2):227-233.

Nunn, G. B. 1992. Nematode molecular evolution. Ph.D. thesis, University of Nottingham, Nottingham, England.
Nunn, G. B., Theisen, B. F., Christensen, B. and Arctander, P. 1996. Simplicity-correlated size growth of the nuclear 28 S ribosomal RNA D3 expansion segment in the crustacean order Isopoda. Journal of Molecular Evolution 42:211-223.

Otsuka, J., Nakano, T. and Terai, G. 1997. A theoretical study on the nucleotide changes under a definite functional constraint of forming stable base-pairs in the stem regions of ribosomal RNAs; its application to the phylogeny of eukaryotes. Journal of Theoretical Biology 184:171-186.

Otsuka, J., Terai, G. and Nakano, T. 1999. Phylogeny of organisms investigated by the base-pair changes in the stem regions of small and large ribosomal subunit RNAs. Journal of Molecular Evolution 48:218-235.

Ragan, M. A. 1988. Ribosomal RNA and the major lines of evolution: a perspective. BioSystems 21:177-188.
Rousset, F., Pelandakis, M. and Solignac, M. 1991. Evolution of compensatory substitutions through G*U intermediate state in Drosophila rRNA. Proceedings of the National Academy of Sciences USA 88:1003210036.

Swofford, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Syvanen, M. 2002. Rates of ribosomal RNA evolution are uniquely accelerated in eukaryotes. Journal of Molecular Evolution 55(1):85-91.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.

Van de Peer, Y., De Rijik, R., Wuyts, J., Winkelmans, T. and De Wachter, R. 2000. The European small subunit ribosomal RNA database. Nucleic Acids Research 28:175-176.

Waters, E. R. and Schaal, B. A. 1996. Heat shock induces a loss of rRNA-encoding DNA repeats in Brassica nigra. Proceedings of the National Academy of Sciences USA 93:1449-1452.

Winnepenninckx, B. and Backeijau, T. 1996. 18S rRNA alignments derived from different secondary structure models can produce alternative phylogenies. Journal of Zoology, Systematics and Evolution Research 34:135-143.

$\underset{240}{24 N G G A A C A C T C T N G G N N G T G N C T T T T T T T C T C T C N N C ~ N A T N N A N N N A A A C N C N T G A G A G ~}$


Figure A1.1: Example of a sequence trace (graphical result from ABI PRISM automated sequencing) that becomes completely doubled after around position 240. The top row shows single peaks throughout, except position 148, which is a double signal, while the bottom row shows double peaks at all, but a few positions (e.g. 250, 267 and 285).


Figures A1.2-A1.3: Screen images showing menus in REALEM to run "Flip analyzer" program. Fig. A1.2 is the first menu, showing three menu choices for different programs. Here Flip analyzer has been chosen. Fig. A1.3 shows the menu choice appearing after "FILE" and "OPEN" were chosen, which allows the user to browse the system for an input file. Here the chosen file is a sample of peak data from top and bottom peaks from a sequence trace of double peaks in text format.


Figures A1.4-A1.5: Screen images showing first steps user takes to run Flip analyzer. Fig. A1.4 shows input data file with top and bottom peaks (nucleotide letters) from the double peak portion of a sequence. Fig. A1.5 shows the "Processes" menu choice. Here "Analyze" will be chosen.


Figures A1.6-A1.7: Screen images showing application of Flip analyzer for an indel (alignment gap, or "shift") of one base pair. Fig. A1.6 shows menu that appears when "Analyze" is chosen. The user types any number, in this case " 1 " was chosen. Fig. A1.7 shows the output results in which the original two rows of nucleotide letters are "flipped" (top and bottom letters swapped) in all positions at which a shift of the top sequence one position right $(\mathrm{Shift}=1)$ or left (Shift = 1) does not make nucleotides on top and bottom rows match. The flipped positions are denoted with "*". If the "flip" does not correct the signal (i.e. the two rows do not match at a position), then the symbol "!" is used. The "SUMMARY" at the bottom can be used to assess the relative number of errors, or "!" symbols representing mismatches between the rows.



Figures A1.8-A1.9: Screen images showing results of Flip analyzer for a larger range of indel sizes (one to 10 shifts in either direction). Fig. A1.8 shows part of a file with "Shifts" of $-10,-9$, -8 , etc. Fig. A1.9 shows the summary, in which for this input, errors are much lower for Shift = 6 , suggesting this is the correct indel size in the original data.


Figures A1.10-A1.11: Screen images showing menus to run "Gap matrix" program. Fig. A1.10 is the first menu, showing three menu choices for different programs. Fig. A1.11 shows the results of choosing "FILE" and "OPEN", which brings up a menu for browsing the system for the input file. The input file in this case is an alignment of DNA sequence data in text format with taxon labels followed by a tab.



Figures A1.12-A1.13: Screen images showing use of Gap matrix program. Fig. A1.12 shows an input file before analysis. Fig. A1.13 shows the first menu choices under "Processes". The user may chose either "Full conversion" to run all the steps of the program or may chose "Convert gap ends" to begin running the program in separate steps.


Figures A1.14-A1.15: Screen images showing further steps in the Gap matrix program. Fig. A1.14 shows the DNA alignment with zero (" 0 ") in each position representing the end of an indel (gap). Fig. A1.15 shows choices under Processes. Here the second step "Create gap matrix" is available.



Figures A1.16-A1.17: Screen images showing results of "Create gap matrix" and the final menu choice in Gap matrix. Fig. A1.16 shows a gap matrix in which all DNA alignment positions without zero (" 0 ") in any taxa were removed, leaving only positions with gap ends (zeros). Fig. A1.17 shows a menu for the final step, "Final conversion".

| [25 HEALEM : DSgapdenol hi |  |
| :---: | :---: |
| Eile Edit Processes Help |  |
| 2i lif | $Q \quad \leqslant \mid$ |
| cLSA4a§ | тTTTCTTCC--ccce-ceccecc |
| clsal cscol |  |
| catleg§Aas | TCTCITITITTTCCTITCT |
| 4_catLe\& | TTTCTITITTITCCTITCTTCCCCTTI |
| O5 catlyas53 | тСССТTTTC-СTCCTTTCTTCCC |
| catlyans53 |  |
| TL¢NF17F | TTTCTITITTTTCTC-ССССССССTTITTTC |
| ATLC $¢ * N \mathrm{NF} 8 \mathrm{~b}$ | TTTCTITITTTTCTC-ССССССССTITTTTCC |
| cPoLb | TTTCTITTC-стете-есессесстTTTTTCC |
| 10_cPOLC\&\&AB4 |  |
| 11_POL@P14758 |  |
| 12_POL®2P2925 | тСССTTTTCCTTCCL-ССССССССТTTTTTC |
|  |  |
| 14_cHaDbas 51 |  |
| 15_CHAD\&H5155 | тTTCTTTTTTTTССТІТСТТССССТСССТТСС |
| 16_HADA¢乡H115 | тСССтTTTTTTTССС-ССсесесетсесттС |
| 17_CTOM¢T5355 | тСССТTTTС-СтССС-ССТТССССТС--СТСС |
| 18_CBLK¢AB28s | TTTTCTTCC--есесетстTC |
| 19 _CLINd\&\&AB8 |  |
| 20_LINC\&QLI 165 |  |
| 21_SHIp $\%$ :S103 | тTCTITITTTCTCесетсесесесте--СтСС |
|  | С-СTCTTCC--ССССТТСССССССТС--CTIT |
| 1.18 | С-СTITTC¢--СССТTTCCTTCCTC---CCC |

Figures A1.18-A1.19: Screen images showing final Gap matrix results. Fig. A1.18 shows a gap matrix in which all nucleotide data has been converted to "T", whereas all gap end data has been converted to "C". Fig. A1.19 shows the results of choosing "FILE" and "Format for PAUP" after gap conversion steps. The program adds the gap matrix to the original data block, with taxon labels, and counts the nucleotides and taxa and inserts PAUP command instructions before and after the data block.


Figure A1.20: Model of rRNA stem state changes used to create a weight matrix. All possible pairs of four nucleotides ( $\mathrm{U}=$ uracil, $\mathrm{A}=$ adenine, $\mathrm{G}=$ guanine, $\mathrm{C}=$ cytosine ) that may be adjacent and hydrogen-bound or in a disrupted unbound state at a single position in an rRNA stem. All evolutionary (nucleotide substitutional) paths that require only one letter to change in the pair are shown (arrows). Transitional changes (purine to purine or pyrimidine to pyrimidine) are shown with " $\alpha$ " and transversional changes (purine to pyrimidine or pyrimidine to purine) are shown with " $\beta$ ". All pairs on the perimeter (outermost letters) of this figure are predicted to be the lowest energy pairs, which are considered as "disruption" states in stems by the model. The model is used to make a matrix assuming the shortest steps following the paths given, in which only one nucleotide may change at a time. A matrix is made by adding $\alpha$ and $\beta$ along the shortest path between two states, as well as adding an additional value "d" (or $\delta$ ) for change from a non-disrupted pair to a disrupted pair, or " $r$ " ( or $\rho$ ) for the reverse, a return from disrupted to non-disrupted state.

|  | A | B | c | D | E | F | G | H | I | J | K | L | M | 2 | 0 | P |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | GC | GT | GA | TA | AT | TG | AG | CG | CC | GG | TT | AA | CA | CT | AC | TC |
| A GC | - | a | $b$ | 2b | 2a | $2 \mathrm{~b}+\mathrm{a}$ | $2 \mathrm{a}+\mathrm{b}$ | $2 \mathrm{a}+2 \mathrm{~b}$ | b+d | $b+d$ | $a+b+d$ | $a+b+d$ | $2 \mathrm{~b}+\mathrm{d}$ | $a+b+d$ | $a+d$ | $b+d$ |
| B GT | a | - | b | 2b | a | $2 \mathrm{~b}+\mathrm{a}$ | $a+b$ | $2 \mathrm{~b}+\mathrm{a}$ | $a+b+d$ | b+d | $b+d$ | $a+b+d$ | $2 \mathrm{~b}+\mathrm{d}$ | $b+d$ | $2 a+d$ | $a+b+d$ |
| C GA | b | b | - | b | $a+b$ | $a+b$ | $2 b+a$ | $2 a+b$ | $2 \mathrm{~b}+\mathrm{d}$ | $a+d$ | $2 \mathrm{~b}+\mathrm{d}$ | $a+d$ | b+d | $2 \mathrm{~b}+\mathrm{d}$ | $a+b+d$ | $2 \mathrm{~b}+\mathrm{d}$ |
| D TA | 2b | 2b | b | - | $2 b+a$ | a | $a+b$ | 2a | $2 a+b+d$ | $a+b+d$ | b+d | b+d | $a+d$ | $2 a+b+d$ | $2 b+a+d$ | $b+d$ |
| E AT | 2a | a | $a+b$ | $2 \mathrm{~b}+\mathrm{a}$ | - | 2b | b | 2b | $2 a+b+d$ | $a+b+d$ | $b+d$ | b+d | $2 \mathrm{~b}+\mathrm{a}+\mathrm{d}$ | $b+d$ | $a+d$ | $2 a+b+d$ |
| F TG | $2 \mathrm{~b}+\mathrm{a}$ | $2 \mathrm{~b}+\mathrm{a}$ | $a+b$ | a | 2b | - | b | a | $a+b+d$ | $b+d$ | $b+d$ | $a+b+d$ | $2 a+d$ | $a+b+d$ | $2 \mathrm{~b}+\mathrm{d}$ | $b+d$ |
| G AG | $2 \mathrm{a}+\mathrm{b}$ | $a+b$ | $2 \mathrm{~b}+\mathrm{a}$ | b+a | b | b | - | b | $2 \mathrm{~b}+\mathrm{d}$ | $a+d$ | $2 \mathrm{~b}+\mathrm{d}$ | $a+d$ | $a+b+d$ | $2 \mathrm{~b}+\mathrm{d}$ | $b+d$ | $2 b+d$ |
| H CG | $2 \mathrm{a}+2 \mathrm{~b}$ | $2 \mathrm{~b}+\mathrm{a}$ | $2 \mathrm{a}+\mathrm{b}$ | 2 a | 2 b | a | b | - | $b+d$ | $b+d$ | $a+b+d$ | $b+a+d$ | $a+d$ | $b+d$ | $2 \mathrm{~b}+\mathrm{d}$ | $a+b+d$ |
| I CC | $\mathrm{b}+\mathrm{r}$ | $b+a+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $2 \mathrm{a}+\mathrm{b}+\mathrm{r}$ | $2 a+b+r$ | $b+a+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $b+r$ | - | $2 b+r+d$ | 2 a | 2b | b | a | b | a |
| J GG | $b+r$ | $b+r$ | $a+r$ | $\mathrm{b}+\mathrm{a}+\mathrm{r}$ | $b+a+r$ | $b+r$ | $a+r$ | $b+r$ | $2 b+r+d$ | - | $2 \mathrm{~b}+\mathrm{r}+\mathrm{d}$ | $2 \mathrm{a}+\mathrm{r}+\mathrm{d}$ | $a+b+r+d$ | $2 \mathrm{~b}+\mathrm{r}+\mathrm{d}$ | $a+b+r+d$ | $2 \mathrm{~b}+\mathrm{r}+\mathrm{d}$ |
| K TT | $b+a+r$ | $b+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $\mathrm{b}+\mathrm{r}$ | $b+r$ | $b+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $b+a+r$ | 2a | $2 \mathrm{~b}+\mathrm{r}+\mathrm{d}$ | - | $2 \mathrm{~b}+\mathrm{r}+\mathrm{d}$ | $a+b+r+d$ | a | $a+b$ | a |
| L AA | $a+b+r$ | $a+b+r$ | $a+r$ | $b+r$ | $b+r$ | $a+b+r$ | $a+r$ | $a+b+r$ | 2 b | $2 a+r+d$ | $2 \mathrm{~b}+\mathrm{r}+\mathrm{d}$ | - | b | 2b | b | 2b |
| M CA | $2 \mathrm{~b}+\mathrm{r}$ | $2 \mathrm{~b}+\mathrm{r}$ | $b+r$ | $a+r$ | $2 b+a+r$ | $2 \mathrm{a}+\mathrm{r}$ | $a+b+r$ | $a+r$ | b | $a+b+r+d$ | $a+b+r+d$ | b | - | b | 2b | $a+b$ |
| Z CT | $b+a+r$ | $\mathrm{b}+\mathrm{r}$ | $2 \mathrm{~b}+\mathrm{r}$ | $2 \mathrm{a}+\mathrm{b}+\mathrm{r}$ | $b+r$ | $a+b+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $b+r$ | a | $2 \mathrm{~b}+\mathrm{r}+\mathrm{d}$ | a | 2b | b | - | $a+b+r+d$ | 2 a |
| $\bigcirc \mathrm{AC}$ | $a+r$ | $2 a+r$ | $a+b+r$ | $2 \mathrm{~b}+\mathrm{a}+\mathrm{r}$ | $a+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $b+r$ | $2 \mathrm{~b}+\mathrm{r}$ | b | $a+b+r+d$ | $a+b$ | b | 2b | $a+b+r+d$ | - | b |
| P TC | b+r | $b+a+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $b+r$ | $2 \mathrm{a}+\mathrm{b}+\mathrm{r}$ | $b+r$ | $2 \mathrm{~b}+\mathrm{r}$ | $a+b+r$ | a | $2 b+r+d$ | a | 2b | $a+b$ | 2a | b | - |

Figure A1.21: Weight matrix for rRNA stem state changes according to shortest paths in the model shown in Fig. 20. (DNA rather than RNA is shown, so uracil U has been replaced by thymidine). The left-most columns represent the starting states of rDNA pairs before the substitution. A single letter (A through P except Z is used instead of N , which is reserved for "unknown") is used to encode the nucleotide pairs as a single-letter character in the new data block that is created as input in PAUP*. The top rows of letters ( A to P ) and nucleotide pairs represent the final state or rDNA pairs after a set of substitutions through states (the shortest path) indicated in the model shown in Fig. 20. The values in the matrix are relative weights that arise from tallying up values along the paths in the model leading from the state in the left column to the state in the top row, where symbol used are: "a" for a transition " $\alpha$ ", "b" for a transversion " $\beta$ ", " $d$ " for a disruption " $\delta$ " and " $r$ " for a return from a disruption " $\rho$ ".


## A1.23

Figure A1.22-A1.23: Screen images showing "Stem state" program. Fig. A1.22 shows input data. Upper case letters represent paired nucleotides in stems. Lower case letters represent nucleotides in loops. Stems are labelled $1,2,3$, etc. with " $a$ " and " $b$ " for facing portion. Fig. A1.23 shows the first menu that gives the choice of a full analysis or partial analysis in steps. Here, "Trim data" was chosen.


Pseudólöna sp ZEBRAFISH Ichthyosporidum sp Let Ichthyosporidium 2 sp L39110 GLugea_atherinae Pleistophorasp Pleistophora anguillarum Microsporidium prosopium Spraguea lophii Trachipleistophora hominis Vavzata oncoperae Micleospora salmonis Endoreticulăcus schubergi Ameson_michaelis Anblyospora_californica Escherichia_coli BACTERTA Tritrichomonas föetus PARABAS Amblyospora californica_2 Eharardia-aedis:
A1.24 Serna_ASCO

Figures A1.24-A1.25: Screen images of processes in Stem state program. Fig. A1. 24 shows data after lower case letters representing loops were removed (using "Trim data"). The next step in the menu choices "Pair nucleotides" was chosen. Fig. A1.25 shows the results of "Pair nucleotides". Each stem is listed and pairs are matched up and listed numerically from the proximal to distal (internal to external) of the rRNA folding structure.


Figures A1.26-A1.27: Screen images in Stem state program. Fig. A1.26 shows stem pairs for a longer stem in this data. Notice that most stem pairs in this block are non-disrupted, while others are disrupted (e.g. AA). The next menu choice "Convert to stem state code" was chosen. Fig. A1.27 shows stem state code results, produced by listing a letter (A to P, but Z replacing N) in place of a nucleotide pair, for each stem and position in order. " N " for unknown is allowed, and here reflects the fact that some sequences were shorter than the rest. N's can be cropped prior to further analysis.


[^6]

Figures A1.30-A1.31: Screen images in showing output weight matrix. Fig. A1.30 shows the result from "Create step matrix", which produces a weight matrix for each stem state transition underneath the stem state data block. Frequencies of each stem condition (e.g. GT and TG, or CG and GC) and disruption $\delta$ and return from disruption $\rho$ are calculated from the data and are shown. Fig. A1.31 shows the menu choice for converting the final results into PAUP format.


Figures A1.32-A1.33: Screen images in showing part of help menus and the version menu of REALEM. Fig. A1.32 shows a portion of the pages in the help files that give sample input files for several steps of the program. Short descriptions of the processes are also included in help files. Fig. A1.33 shows the "About" caption, giving details, dates, authors, etc. for the REALEM programs.

## APPENDIX 2

Example of likelihoods and other tree parameters, calculated in PAUP* (Swofford, 2001) using instructions in Modeltest (Posada \& Crandall, 1998), to compare 64 different models of evolution. These data are used as input in Modeltest.

```
Tree - lnL
1 1659.27540352
Tree -lnL p-inv
1617.86333321 0.69429175
Tree -lnL gamma shape
1614.77317545 0.209972
Tree -lnL p-inv gamma shape
1 1611.71564600 0.49092021 0.709204
Tree -lnL freqA freqC freqG freqT
1 1632.61950530 0.36805452 0.16524064 0.24753362 0.21917121
Tree -lnL freqA freqC freqG freqT p-inv
1 1591.32862349 0.36852888 0.16449961 0.24577896 0.22119255
    0.69221586
Tree -lnL freqA freqC freqG freqT gamma shape
1 1588.13307175 0.36904464 0.16428658 0.24565056 0.22101822
    0.211801
Tree -lnL freqA freqC freqG freqT p-inv gamma shape
1 1585.11810981 0.36889659 0.16414575 0.24531462 0.22164304
    0.48610545 0.699383
Tree -lnL ti/tv ratio
1 1656.73559005 0.73632660
Tree -lnL ti/tv ratio p-inv
1 1615.53734317 0.74049089 0.69357544
Tree -lnL ti/tv ratio gamma shape
1 1612.39071381 0.74244855 0.211310
Tree -lnL ti/tv ratio p-inv gamma shape
1 1609.27867683 0.74781917 0.49245532 0.714040
Tree -lnL freqA freqC freqG freqT ti/tv ratio
1 1631.29986139 0.36623543 0.16705771 0.24555318 0.22115368
    0.70843399
Tree -lnL freqA freqC freqG freqT ti/tv ratio p-inv
1 1589.86241639 1
    0.73024735 0.69287319
Tree -lnL freqA freqC freqG freqT ti/tv ratio gamma shape
1 1586.62283324 0.36767782 0.16556754 0.24318820
    0.73305282 0.211275
Tree -lnL freqA freqC freqG freqT ti/tv ratio p-inv gamma shape
1 1503.46875392 150.36766713 0.16527145 0.24275565 0.22430577
    0.74457875 0.49035207 0.701749
Tree - lnL R(a) R(b) R(c) R(d) R(e)
1 1656.28481070 1.00000 1.60555 1.00000 1.00000
        1.24800
Tree -lnL R(a) R(b) R(c) R(d) R(e) p-inv
1 1614.14877228 1.00000 1.76795 1.00000 1.00000
        1.06937 0.69655361
•
.
etc.
```


## APPENDIX 3

Example of output file from Modeltest 3.06 (Posada \& Crandall, 1998), showing the results of the likelihood ratio test and Akaike criteria used to estimate the best-fit model for a single data set.

Testing models of evolution - Modeltest Version 3.06(c) Copyright, 1998-2000 David Posada (dp47@email.byu.edu) Department of Zoology, Brigham Young University WIDB 574, Provo, UT 84602, USA

```
Sat Dec 27 18:09:14 2003
Input format: Paup matrix file
    ** Log Likelihood scores **
\begin{tabular}{llllll} 
& & \(+I+G\) & & & \\
JC & \(=\) & 1617.8633 & 1617.8633 & 1614.7732 & 1611.7157 \\
F8I & \(=\) & 1591.3286 & 1591.3286 & 1588.1331 & 1585.1182 \\
K80 & \(=\) & 1615.5374 & 1615.5374 & 1612.3907 & 1609.2787 \\
HKY & \(=\) & 1589.8624 & 1589.8624 & 1586.6228 & 1583.4688 \\
TrNef & \(=\) & 1614.1488 & 1614.1488 & 1610.9298 & 1607.5643 \\
TrN & \(=\) & 1589.8329 & 1589.8329 & 1586.5811 & 1583.4545 \\
K81 & \(=\) & 1509.5154 & 1609.5154 & 1606.2714 & 1603.0746 \\
K81uf & \(=\) & 1608.0945 & 1608.0945 & 1604.7798 & 1601.3213 \\
TIMef & \(=\) & 1584.5305 & 1584.5305 & 1581.2087 & 1578.0905 \\
TIM & \(=\) & 1596.6903 & 1596.6903 & 1592.8766 & 1589.6929 \\
TVMef & \(=\) & 1576.6593 & 1576.6593 & 1573.2216 & 1570.2214 \\
TVM & \(=\) & 1595.3271 & 1595.3271 & 1591.3938 & 1588.0809 \\
SYM & \(=\) & 1576.6163 & 1576.6163 & 1573.1611 & 1570.1820
\end{tabular}
** Hierarchical Likelihood Ratio Tests (hLRTs) **
Equal base frequencies
    Null model = JC
    -lnL0 = 1659.2754
    Alternative model = F81
    -lnL1 = 1632.6195
    2(lnL1-lnL0)=53.3118
        df = 3
    P-value = <0.000001
Ti=Tv
    Null model = F81 -lnL0 = 1632.6195
    Alternative model = HKY - - nL1 = 1631.2998
    2(lnL1-lnL0) = 2.6394 }\quad\textrm{df}=
    P-value = 0.104243
Equal rates among sites
    Null model = F81 -lnLO = 1632.6195
    Alternative model = F81+G - nnL1 = 1588.1331
    2(lnL1-lnL0) = 88.9729
        df = 1
    Using mixed chi-square distribution
    P-value = <0.000001
No Invariable sites
    Null model = F81+G -lnL0 = 1588.1331
    Alternative model = F81+I+G -lnL1 = 1585.1182
    2(lnL1-lnLO) = 6.0298
        df = 1
    Using mixed chi-square distribution
    P-value = 0.007033
```

```
Model selected: F81+I+G
    -lnL = 1585.1182
    Base frequencies:
        freqA = 0.3689
        freqC = 0.1641
        freqG = 0.2453
        freqT = 0.2216
    Substitution model:
        All rates equal
    Among-site rate variation
        Proportion of invariable sites (I) = 0.4861
        Variable sites (G)
            Gamma distribution shape parameter = 0.6994
PAUP* Commands Block: If you want to implement the previous estimates as
likelihod settings in PAUP*, attach the next block of commands after the data
in your PAUP file:
[!Likelihood settings from best-fit model (F81+I+G) selected by hLRT in
Modeltest Version 3.06]
BEGIN PAUP;
Lset Base=(0.3689 0.1641 0.2453) Nst=1 Rates=gamma Shape=0.6994
Pinvar=0.4861;
END;
--
    ** Akaike Information Criterion (AIC) **
    Model selected: TVM+I+G
    -lnL = 1570.2214
        AIC = 3158.4429
        Base frequencies:
        freqA = 0.3439
        freqC = 0.1698
        freqG = 0.2645
        freqT = 0.2219
    Substitution model:
        Rate matrix
        R(a) [A-C] = 4.2151
        R(b) [A-G] = 5.5175
        R(c) [A-T] = 7.5058
        R(d) [C-G] = 2.6986
        R(e) [C-T] = 5.5175
        R(f) [G-T] = 1.0000
    Among-site rate variation
        Proportion of invariable sites (I) = 0.4780
        Variable sites (G)
            Gamma distribution shape parameter = 0.6752
PAUP* Commands Block: If you want to implement the previous estimates as
likelihod settings in PAUP*, attach the next block of commands after the data
in your PAUP file:
[!Likelihood settings from best-fit model (TVM+I+G) selected by AIC in Modeltest Version 3.06]
```

```
BEGIN PAUP;
Lset Base=(0.3439 0.1698 0.2645) Nst=6 Rmat=(4.2151 5.5175 7.5058 2.6986
5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
```

END;

Time processing: 0 seconds
If you need help type '-?' in the command line of the program

## APPENDIX 4

Sample PAUP* commands used in a typical analysis, showing maximum parsimony (MP), distance (ME), and maximum likelihood (ML) with parameters and model estimated in Modeltest of Posada \& Crandall (1998), each followed by bootstrap commands.

```
#NEXUS
    BEGIN DATA;
    DIMENSIONS NTAX=240 NCHAR=1869;
    FORMAT MISSING=N GAP=- MATCHCHAR=. INTERLEAVE DATATYPE=DNA;
    MATRIX
Taxon 1 AtgCGGGTCG...
Taxon 2 AtTtGGCTGG...
etc.
begin paup;
log file = logtest.txt start;
set autoclose=yes warntree=no warnreset=no;
set criterion=PARSIMONY;
set maxtrees=1000 increase=no;
hsearch addseq=random nreps=10 swap=TBR nchuck=100 chuckscore=1;
savetrees file=MPtest.tre brlens replace=yes;
bootstrap keepall=yes treefile=MPBtrees.out search=heuristic nreps=1000
bseed=0;
gettrees file=MPBtrees.out StoreTreewts=yes mode=3 storebrlens=yes
duptrees=eliminate;
contree all/strict=no majrule=yes usetreewts=yes le50=yes
treefile=MPBootcontree.tre;
savetrees file=MPBcontree.tre brlens savebootP replace=yes;
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch addseq=random nreps=10 swap=TBR nchuck=100 chuckscore=1;
savedist triangle=lower file=dist;
showdist;
savetrees file=MEtest.tre brlens replace=yes;
bootstrap keepall=yes treefile=MEBtrees.out search=heuristic nreps=1000
bseed=0;
gettrees file=MEBtrees.out StoreTreewts=yes mode=3 storebrlens=yes
duptrees=eliminate;
contree all/strict=no majrule=yes usetreewts=yes le50=yes
treefile=MEBootcontree.tre;
savetrees file=MEBcontree.tre brlens savebootP;
[model (TVM+I) selected by AIC]
set criterion=LIKELIHOOD;
Lset Base=(0.2915 0.1970 0.3099) Nst=6 Rmat=(0.6467 1.7089 2.6068 0.2142
1.7089) Rates=equal Pinvar=0.7741;
Dset distance=ML;
Hsearch start=nj swap=TBR;
lscore 1;
savetrees file=NJtest.tre brlens=yes;
```

```
Lset Base=estimate Nst=6 Rmat=estimate Rates=gamma Shape=estimate
Pinvar=estimate;
hsearch nreps=1 swap=NNI start=1;
savetrees file=MLtest.tre brlens;
bootstrap keepall=yes treefile=MLBtrees.out search=faststep nreps=1000;
gettrees file=MLBtrees.out StoreTreewts=yes mode=3 storebrlens=yes
duptrees=eliminate;
contree all/strict=no majrule=yes usetreewts=yes le50=yes
treefile=MLBootcontree.tre;
savetrees file=MLBcontree.tre brlens savebootP;
log stop;
end;
```


## APPENDIX 5

Two kinds of hypothesis test were performed in Chapter 4 to evaluate species validity. For molecular data, species must form statistically supported groups (or clades) that separate sympatric sister-groups. Genetic intermediates between groups would suggest there is gene flow, whereas reduced intermediates suggests a genetic discontinuity or speciesboundary. Statistical support for species clades was evaluated by creating alternate trees of two kinds (below), then comparing them using Shimodaira \& Hasegawa's (2001) AU test to reject less likely trees:

1. Create a new tree that places the two proposed species-groups together:


Formal hypotheses:
$\mathrm{H}_{0}$ : Left tree is most likely.
$\mathrm{H}_{\mathrm{A}}$ : Right tree is equally likely.
If we Reject $\mathrm{H}_{\mathrm{A}}$, the proposed species-groups may be valid. If we do not reject $H_{A}$, species may or may not be valid (data are insufficient).
2. Create a new tree that separates a proposed sub-group from a larger group:

$\mathrm{H}_{\mathrm{A}}$ : this tree is as likely as the tree on the left.

Formal hypotheses:
$\mathrm{H}_{0}$ : Left tree is most likely.
$\mathrm{H}_{\mathrm{A}}$ : Right tree is equally likely
If we Reject $H_{A}$, the proposed sub-group is NOT a valid species.
If we do not reject $\mathrm{H}_{\mathrm{A}}$, the sub-group may or may not be a valid species (data are insufficient).

## APPENDIX 6

## Justification for use of the "AU test" of Shimodaira \& Hasegawa (2001)

For hypothesis testing in Chapters 4 to 6 , phylogenies were compared using a variety of statistics. Several test statistics have been devised for application to phylogenetic questions where the data suggests several equally good trees, all involving some form of bootstrap resampling of the data. These include the familiar bootstrap probability, the Kishino-Hasegawa $(\mathrm{KH})$ test, the Shimodaira-Hasegawa (SH) test, and the Approximately Unbiased (AU) test (Kishino \& Hasegawa, 1989; Goldman et al., 2000; Shimodaira \& Hasegawa, 2001; Shimodaira, 2002). These tests can help refine the search for best trees, or help provide support for or against a set of hypothetical trees created based on separate (a priori) knowledge or assumptions. All these tests have been shown to have either restrictive assumptions (e.g. KH requires the trees be chosen a priori, rather than from a set of phylogenetic results, Goldman et al., 2000), significant biases (e.g. bootstrap and KH are biased because the trees to be tested come from a set already selected in tree-searches), or limitations (e.g. SH and WSH become inaccurate when large numbers of trees are tested, and AU becomes inaccurate when many of the best trees are almost equally as good, Shimodaira, 2002). Regardless of such potential problems, Shimodaira's AU test (Shimodaira, 2000; Shimodaira \& Hasegawa, 2001; Shimodaira, 2002) has been demonstrated to have important advantages over the above tests in that it does not tend to be too conservative, as the SH was demonstrated to be (i.e. SH does not reject enough wrong trees) or cause type 1 error (e.g. KH will reject correct trees), and performs better in most cases than other tests, including the weighted SH (WSH) test that apparently corrects some but not all of the conservatism of the normal SH test (according to Shimodaira, 2002). The AU test involves a form of bootstrap resampling of the data with a bias correction achieved by changing sequence lengths in each replicate, known as a "multi-scale bootstrap", which came from geometric theory devised in Efron et al. (1996) and expanded on by Shimodaira (2000). The AU test was considered of interest and value in this thesis because it could be applied to test a variety of hypothetical (a priori) trees based on predictions that species from the same hosts should group together and then to compare these trees of interest to one another, and to provide some statistical support for or against such hypotheses. In this manner, species-boundaries corresponding to various host-parasite groups could be critically tested.

Some assumptions of the AU test need to be considered. In this thesis some, but not all, trees tested varied only in ways that were hypothesized by a priori assumptions (i.e.
assumptions not depending on knowledge from ML analyses, or even DNA data). The requirement that trees be chosen strictly a priori applies to the KH and SH tests, according to Goldman et al. (2000), such that one should test either all-possible tree topologies (too many to compute for some of the data here), or should test only a set of trees that were derived from prior knowledge. While Shimodaira (2002) did not specify this restriction in his detailed description and defense of the AU test, and he instead suggested it be used on a set of best trees chosen by maximum likelihood, one might expect that the assumption that applies for KH and SH tests, that all trees be chosen from a priori hypotheses also applies to the AU test, on the same basic principle discussed in Goldman et al. (2000). One alternative way to apply Shimodaira's AU test on a set of trees smaller than the set of all possible trees using only a priori knowledge is to turn the a priori rule around, and use a priori knowledge to reject large numbers of all possible trees based on independent knowledge. In this thesis, a large number of trees with branching patterns not consistent with other data were eliminated - because they do not correspond with what is known about the morphology, geographic locality, transmissibility, or likelihood or a hypothesis of host-parasite co-evolution.

In chapters of this thesis an even smaller set of trees than the set of "all possible trees minus those unlikely given a priori assumptions" were compared, even though the "turned around" a priori rule (i.e. the rule that allows tree topologies to be eliminated, rather than constructed, based on prior knowledge) was inadvertently being applied. The set of chosen trees was similar in all ways to a best distance tree except that they were designed to vary based on $a$ priori hypotheses (e.g. host group). Therefore, if the AU test depends on strict adherence to the a priori rule, its application in this thesis may be flawed, however, it is important to note that in this thesis (1) the trees selected for comparison were not chosen from results of ML analyses which would be the most severe violation of the a priori rule (Goldman et al., 2000) because such tests and best ML trees would then both come from comparison of likelihood scores. Instead, the trees tested here were always obtained using the minimum evolution distance (logDet paralinear) criterion, and so would not so severely violate the rule of independence ( $a$ priori rule), as would trees from an ML search. Of course, these distance criterion trees are, nevertheless, not independent of the DNA data, so a second important point to note is that (2) the trees selected for comparison in this thesis were chosen such that all nodes except those of interest were fixed in the trees to be tested, whereas nodes of interest differed only in ways determined strictly a priori (i.e. by restricting taxa to host groups, gap groups that are
independent of DNA substitutional data used in tree searches, or genus/species groups that were established from independent morphological study). In this way, the AU test was employed to compare trees with a priori topological changes.

These were also compared against a best distance tree not chosen independently of the data, but chosen independently of ML likelihood scores. I justify inclusion of the best distance tree, the "unconstrained" trees in chapters of this thesis, by the fact that Shimodaira (2002) suggested the test requires that the "true" tree (i.e. the tree representing the actual evolutionary history of species) be included among the test set, and so the best distance tree was considered an important contender hypothesis. This "true" tree, of course, may not have been included, and for several reasons such a "true" dichotomously branching tree may not, in fact, exist. For these reasons, the AU test used here should be regarded as a comparison of a set of similar trees that are as close as possible to the set of best (most true) trees, and that they vary in a priori ways (complying with the a priori rule as much as possible), while again, it is important to note that Shimodaira (2002) presented no evidence that the a priori rule was critical for this kind of test (as it is for KH or SH tests, according to Goldman et al., 2000).

Regardless of the strictness with which this application of the AU test complies with its assumptions, the relative statistical support for these hypotheses should be valid, even if the absolute p-values are potentially in error. Thus, even if the conclusion to reject or not reject trees is over or under-conservative, the relative order in which the trees are ranked in probability is correct and informative in these tests of species-boundaries.

## Literature Cited

Efron, B., Halloran, E. and Holmes, S. 1996. Bootstrap confidence levels for phylogenetic trees. Proceedings of the National Academy of Sciences USA 93:13 429-13 434.

Goldman, N., Anderson, J. P. and Rodrigo, A. G. 2000. Likelihood-based tests of topologies in phylogenetics. Systematic Biology 49(4): 652-670.

Kishino, H. and Hasegawa, M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29:170-179.

Shimodaira, H. and Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics. 17(12): 1246-1247.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology. 51(3): 492-508.

Shimodaira, H. 2000. Another calculation of the p-value for the problem of regions using the scaled bootstrap resamplings. Technical Report No. 2000-35. Stanford University.

## APPENDIX 7

Instructions used to force groups into monophyletic clades in PAUP* (Swofford, 2001), and then calculate tree and site log likelihood scores for the constrained and unconstrained trees. The output will later be used for AU and other tests in CONSEL (Shimodaira \& Hasegawa, 2001).

```
#NEXUS
    BEGIN DATA;
    DIMENSIONS NTAX=72 NCHAR=479;
    FORMAT MISSING=N GAP=- MATCHCHAR=. INTERLEAVE DATATYPE=DNA;
    MATRIX
[Positions: ]
    [ 1111...]
    [1234567890123...]
cLSA4&$BA4 ctggatca...
cLSA1&$COL ctggatca...
cLSA2&$COL ctggatca...
.
.
;
end;
begin paup;
log file = logtest.txt start;
set autoclose=yes warntree=no warnreset=no;
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch addseq=random nreps=10 swap=TBR nchuck=100 chuckscore=1;
savetrees file=MEtestl.tre brlens replace=yes;
[TVM+I+G]
set criterion=LIKELIHOOD;
Lset Base=(0.3439 0.1698 0.2645) Nst=6 Rmat=(4.2151 5.5175 7.5058 2.6986
5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
lscores 1-50 /sitelikes=yes scorefile=Polyl.txt;
constraints PoM2 (monophyly)=
(1,2,3,4,8,10,11,12,15,16,17,13,24,14,18,19,20,21,22,23,25,26,27,
28,31,32, (40,39,38,37,36,35,34,33),29,30,41,42,46,48,43,44,45,47,53,51
,50,49,52,72,69,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,70,71
,5,6,7,9);
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch constraints=PoM2 enforce=yes addseq=random nreps=10 swap=TBR
nchuck=100 chuckscore=1;
savetrees.file=MEtest2.tre brlens replace=yes;
set criterion=LIKELIHOOD;
```

Lset $\operatorname{Base}=(0.34390 .16980 .2645) \quad$ Nst=6 $\quad$ Rmat $=(4.21515 .51757 .50582 .6986$ 5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
lscores 1-49/sitelikes=yes scorefile=PoM2.txt;
constraints PaM3 (monophyly) =
$(1,2,3,4,8,10,11,12,15,16,17,13,34,18,19,20,21,22,23,35,36,37,38$, $39,40,(32,31,29,30,28,27,26,25,24), 33,14,41,42,46,48,43,44,45,47,53,51$
$, 50,49,52,72,69,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,70,71$ ,5,6,7,9);
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees $=1000$ increase $=$ no;
hsearch constraints=PaM3 enforce=yes addseq=random nreps=10 swap=TBR
nchuck=100 chuckscore=1;
savetrees file=MEtest3.tre brlens replace=yes;
set criterion=LIKELIHOOD;
Lset $\operatorname{Base}=(0.34390 .1698$ 0.2645) $\quad$ Nst=6 $\quad \operatorname{Rmat}=(4.21515 .51757 .50582 .6986$
5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
lscores 1-49 /sitelikes=yes scorefile=PaM3.txt;
constraints AtM4 (monophyly) =
$(1,2,3,4,8,24,34,25,26,27,28,31,32,35,36,37,38,39,40,29,30$
$, 33,(13,14,10,11,12,15,16,17,18,23,22,21,20,19), 41,42,46,48$,
$43,44,45,47,53,51,50,49,52,72,69,54,55,56,57,58,59,60$,
$61,62,63,64,65,66,67,68,70,71,5,6,7,9)$;
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch constraints=AtM4 enforce=yes addseq=random nreps=10 $s w a p=T B R$
nchuck=100 chuckscore=1;
savetrees file=MEtest4.tre brlens replace=yes;
set criterion=LIKELIHOOD;
Lset $\operatorname{Base}=(0.34390 .1698$ 0.2645) Nst=6 $\quad$ Rmat $=(4.21515 .51757 .50582 .6986$
5.5175) Rates=gamma Shape $=0.6752$ Pinvar $=0.4780$;
lscores 1-49 /sitelikes=yes scorefile=AtM4.txt;
constraints LFLsM5 (monophyly)=
$((1,2,3,4,5,6,7,8,9,70,71), 10,11,12,15,16,17,13,24,34,18,19,20,21,22,23,25,26$ ,27,
$28,31,32,35,36,37,38,39,40,29,30,33,14,41,42,46,48,43,44,45,47,53,51$
, 50, 49, 52, 72,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69);
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch constraints=LFLsM5 enforce=yes addseq=random nreps=10 swap=TBR nchuck=100 chuckscore=1;
savetrees file=MEtest5.tre brlens replace=yes;
set criterion=LIKELIHOOD;
Lset $\operatorname{Base}=(0.34390 .16980 .2645) ~ N s t=6 \quad$ Rmat $=(4.21515 .51757 .50582 .6986$ 5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
lscores 1-49 /sitelikes=yes scorefile=LFLsM5.txt;

```
constraints BlLiSep6 (monophyly)=
(1,2,3,4,8,10,11,12,15,16,17,13,24,34,18,19,20,21,22,23,25,26,27,
28,31,32,35,36,37,38,39,40,29,30,33,14,41,42,46,48,43,44,45,47,53,51
,50,49,52,72,69,(55,56,57,58,59,60,54),(61,62,63,64,65,66,67,68),70,71
,5,6,7,9);
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch constraints=BlLiSep6 enforce=yes addseq=random nreps=10 swap=TBR
nchuck=100 chuckscore=1;
savetrees file=MEtest6.tre brlens replace=yes;
set criterion=LIKELIHOOD;
Lset Base=(0.3439 0.1698 0.2645) Nst=6 Rmat=(4.2151 5.5175 7.5058 2.6986
5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
lscores 1-49 /sitelikes=yes scorefile=BlLiSep6.txt;
constraints TomSep7 (monophyly)=
(1,2,3,4,8, (() (10, 11, 12,15,16,17, 13,24,34,18,19,20,21,22,23,25,26,27,
28,31,32,35,36,37,38,39,40,29,30,33,14,41,42,46,48,43,44,45,47),53,51)
,50,49,52),72,69,55,56,57,58,59,60,54,61,62,63,64,65,66,67,68,70,71
,5,6,7,9);
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch constraints=TomSep7 enforce=yes addseq=random nreps=10 swap=TBR
nchuck=100 chuckscore=1;
savetrees file=MEtest7.tre brlens replace=yes;
set criterion=LIKELIHOOD;
Lset Base=(0.3439 0.1698 0.2645) Nst=6 Rmat=(4.2151 5.5175 7.5058 2.6986
5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
lscores 1-49 /sitelikes=yes scorefile=TomSep7.txt;
constraints AtNFSep8 (monophyly)=(1,2,3,4,8,13,24,18,19,20,21,
22,23,25,26,27,28,31,32,35,36,37, (38,15,16,17),39,40,29,30, (34,33,
10,11,12),14,41,42,46,48,43,44,45,47,53,51,50,49,52,72,69,55,56,57,
58,59,60,54,61,62,63,64,65,66,67,68,70,71,5,6,7,9);
set criterion=DISTANCE;
dset distance=logdet objective=ME;
set maxtrees=1000 increase=no;
hsearch constraints=AtNFSep8 enforce=yes addseq=random nreps=10 swap=TBR
nchuck=100 chuckscore=1;
savetrees file=MEtest8.tre brlens replace=yes;
set criterion=LIKELIHOOD;
Lset Base=(0.3439 0.1698 0.2645) Nst=6 Rmat=(4.2151 5.5175 7.5058 2.6986
5.5175) Rates=gamma Shape=0.6752 Pinvar=0.4780;
lscores 1-49 /sitelikes=yes scorefile=AtNFSep8.txt;
log stop;
end;
```


## APPENDIX 8

Example of tree- and site-likelihood scores, produced by PAUP* (Swofford, 2000), for a series of trees, including the unconstrained and various monophyly-constrained trees. These data are input into CONSEL (Shimodaira \& Hasegawa, 2001), which performs the AU test and other statistical comparisons among trees.

```
Tree -lnL Site -lnL
1 1583.91945729
    1 2.02383548
    2 1.77175895
    1.54083868
    477 5.14081322
    478 9.58336746
    479 9.20252484
2
    1577.83931389
    1 2.02109592
    2 1.76897037
    1.53827310
    . .
    -
    4 7 7 \quad 5 . 1 2 9 6 7 3 5 3
    478 9.57040623
    479 9.19038700
3 1579.44933140
    1 2.02225728
    2 1.77014922
    3 1.53935583
    . .
    •
    4 7 7 \quad 5 . 1 3 6 9 1 1 0 8
    478 9.57813140
    479 9.19681205
4 1582.82849749
    1 2.02144131
    2 1.76929795
    3 1.53857685
    etc.
```


## APPENDIX 9

Example of output generated by CONSEL (Shimodaira \& Hasegawa, 2001), showing results of the AU test and other statistical comparisons among trees.

| \# reading fitconsel.pv |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| \# | rank | item | obs | au | np | bp | pp | kh | sh | wkh | wsh |
| \# | 1 | 2 | -1.6 | 0.761 | 0.518 | 0.518 | 0.823 | 0.612 | 0.956 | 0.612 | 0.959 |
| \# | 2 | 3 | 1.6 | 0.542 | 0.329 | 0.331 | 0.164 | 0.388 | 0.935 | 0.388 | 0.919 |
| \# | 3 | 4 | 5.0 | 0.112 | 0.003 | 0.001 | 0.006 | 0.088 | 0.824 | 0.088 | 0.404 |
| \# | 4 | 7 | 5.9 | 0.187 | 0.070 | 0.074 | 0.002 | 0.170 | 0.783 | 0.170 | 0.620 |
| \# | 5 | 1 | 6.1 | 0.039 | 0.003 | 0.001 | 0.002 | 0.110 | 0.813 | 0.110 | 0.556 |
| \# | 6 | 6 | 6.1 | 0.039 | 0.003 | 0.001 | 0.002 | 0.110 | 0.813 | 0.110 | 0.556 |
| \# | 7 | 5 | 6.7 | 0.168 | 0.073 | 0.070 | 0.001 | 0.147 | 0.738 | 0.147 | 0.588 |
| \# | 8 | 9 | 17.1 | 0.003 | 0.001 | 0.001 | 3e-008 | 0.042 | 0.305 | 0.042 | 0.136 |
| \# | 9 | 10 | 24.4 | 0.007 | 0.003 | 0.003 | $2 e-011$ | 0.023 | 0.168 | 0.023 | 0.056 |
| \# | 10 | 8 | 34.9 | 0.003 | 0.002 | 0.001 | 6e-016 | 0.011 | 0.039 | 0.011 | 0.032 |
| \# | 11 | 11 | 47.2 | 7e-005 | 5e-005 | 0 | 3e-021 | 0.010 | 0.021 | 0.004 | 0.008 |

## APPENDIX 10

Dot-blot results for radio-labelled (GT) ${ }_{15}$ microsatellite probe on DNA from purified spores of Loma salmonae (Chapter 2). Arrows show weak hybridization of the probe to only the largest and second largest amounts of DNA (from left to right, L. salmonae and sockeye salmon dots represent $1.82 \mu \mathrm{~g}, 0.88 \mu \mathrm{~g}, 0.44 \mu \mathrm{~g}$, and $0.11 \mu \mathrm{~g}$ of DNA) applied to the nitrocellulose membrane. Negative control (Neg. Ctr) dots represent $1.6 \mu \mathrm{~g}$ and $4 \mu \mathrm{~g}$ of 1 kb ladder DNA.


Loma salmonae DNA


Salmon +'ve Ctr

## APPENDIX 11

Features of previously described and new Loma species (Chapter 3). PV or $S=$ parasitophorous vacuole formed by coalescence of vesicles from host (PV) or formed by blistering or delamination of material from parasite (S); Gills or not $=$ in gills $(\mathrm{G})$ or not in gills (n); Host O/F/G = letter from order/family/genus (orders: Perciformes, Gadiformes, Salmoniformes, Anguilliformes, Scorpaeniformes; families: Percidae, Sparidae, Gadidae, Cichlidae, Gobiidae, Embiotocidae, Salmonidae, Mugilidae, Ophichthidae, Trichiuridae, Hexagrammidae, Anoplopomatidae, Lotidae; genera: see Table 1.1 Chapter 1); X = xenoma larger or smaller than in new species, this study; $\mathrm{S}=$ spore size distinctly larger ( L ) or different shape (*) or not known or not clearly different (?); $\mathrm{W}=$ host habitat (marine, freshwater); $\mathrm{L}=$ locality ( $\mathrm{p}=$ Pacific Northwest, $\mathrm{a}=$ Atlantic, $\mathrm{f}=$ France, $\mathrm{r}=$ Russia, $\mathrm{af}=$ Africa, $\mathrm{i}=$ India, $\mathrm{m}=$ Mediterranean, $\mathrm{b}=$ Brazil, $\mathrm{fi}=$ Finland); $\mathrm{pf}=$ polar filament turns, lots $(\mathrm{L})$, medium number ( m ) or few (f); $\mathrm{o}=\left({ }^{*}\right)$ other morphological differences, see Chapter 3.

| Species and authority | $\begin{aligned} & \text { PV } \\ & \text { or } \mathrm{S} \end{aligned}$ | Gills or not | Host O/F/G | X | S | W | L | pf | 0 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| L. pacificodae n. sp. (this study) | PV | G | GGG | $s$ | - | m | p | L | * |
| L. wallae n. sp. (this study) | PV | G | GGT | $s$ | - | m | p | L | * |
| L. kenti n. sp. (this study) | PV | G | GGM | $s$ | - | m | p | m | * |
| L. lingcodae n. sp. (this study) | PV | G | CHO | $s$ | - | m | p | m | * |
| L. richardi n . sp. (this study) | PV | G | CAA | $s$ | - | m | p | f | * |
| L. branchialis (Nemeczek, 1911) <br> Morrison \& Sprague, 1981 | PV | G | GGG | L | L | m | a | L |  |
| L. salmonae (Putz, Hoffman \& Dunbar, 1965) Morrison \& Sprague, 1981 | PV | G | SSO | L | L | f | p | m |  |
| L. embiotocia Shaw, Kent, Docker, Brown, Devlin \& Adamson, 1997 | PV | G | PEE | s | ? | m | p | L |  |
| L. fontinalis Morrison \& Sprague, 1983 | PV | G | SSS | L | ? | f | a | m |  |
| Loma sp. studied by E. U. Canning | ? | ? | GLE | ? | $?$ | m | a? | ? |  |
| L. diplodae Bekhti \& Bouix, 1985 | S | G* | PRD | s | L | m | f | L |  |
| L. mugili Ovcharenko, Sarabeev, Wita \& Czaplińska (2000) | ? | G | PMM | L | ? | m | r | m |  |
| L. trichiuri Sandeep \& Kalvati, 1985 | ? | G | PTL | L | * | m | i | ? |  |
| L. dimorpha Loubès, Maurand, Gasc, De Buron \& Barral, 1984 | S | n | PBG | s | ? | m | m | f |  |
| L. boopsi Faye, Toguebaye \& Bouix, 1995 | S | n | PSB | L | ? | m | af | f |  |
| L. myrophis Azevedo \& Matos, 2002 | S | n | AOM | s | ? | f | b | m |  |
| L. camerounensis Fomena, Coste \& Bouix, 1992 | S | n | PCO | s | ? | f | af | f |  |
| Loma sp. of Bekhti (1984) | ? | G* | PCT | ? | ? | f | af | ? |  |
| L. acerinae (Jírovec, 1930) Lom \& Pekkarinen, 1999 | PV | n | PPG | ? | ? | f | fi | m |  |

## APPENDIX 12

Name equivalents for Loma species whose names vary between chapters. Proposed new species names in Chapter 3 were not used in Chapter 4 because Chapters 3 and 4 represent draft manuscripts intended for publication at the same time. * = these five new species were also previously known as "Loma sp. of Kent et al., (1998)". Proposed new species names should be regarded as temporary names only, according to the International Code of Zoological Nomenclature (see DISCLAIMER in Chapter 3).

| Proposed new species names (Chapter 3): | Name equivalent in Chapter 4 | Derivation of name in Chapter 4 from host common name |
| :---: | :---: | :---: |
| Loma pacificodae n . sp . | Loma sp. PAC * | PACific cod |
| Loma wallae n. sp. | Loma sp. POL | walleye POLlock |
| Loma kenti n . sp. | Loma sp. TOM | Pacific TOMcod |
| Loma lingcodae n. sp. | Loma sp. LIN | LINgcod |
| Loma richardin. sp. | Loma sp. BLK | sablefish or BlacK cod |
| Published species names (Chapters 3 \& 4): |  |  |
| Loma morhua of Morrison \& Sprague (1981a; b) and Loma branchialis of Canning \& Lom (1986) | Loma morhua ATL | ATLantic cod |
| Loma branchialis of Morrison \& Sprague (1981a; b) and other records of this species in haddock (see Table 4.3) | Loma branchialis HAD | HADdock |
| Published names (Chapter 2): |  |  |
| Loma salmonae SV of Sánchez et al. (2001a; b) and possibly also Loma fontinalis | Loma sp. BRO | BROok trout |

## APPENDIX 13

Key to reading isolate names used in all figures and tables in Chapter 3. Species names are comprised of several parts, the first indicating the species name (e.g. Loma sp. or $L$. morhua), the second part indicating the host species, the third part indicating the isolate number (corresponding to a single host individual, collection details of which can be found in Table 4.1), and the last part is the sequence number if more that one unique sequences were obtained. For species that begin with "Loma sp.", the host part of the name was included throughout the text (e.g. Loma sp. PAC) in order to distinguish these un-named* species from one another.

For example:


* See species name equivalences between chapters in Appendix 12.
${ }^{\phi}$ Host species name abbreviations were derived from the host common name (as shown in Appendix 12)

APPENDIX 14
Pairwise $\log$ Det/paralinear distance matrix from analysis of the full rDNA cistron of reference sequences for Loma species (Chapter 4). Reference sequences were generated by choosing the character shared by the majority of isolates, clones, or sequenced products. Boxes in bold indicate distances for species that could not be transmitted experimentally between reciprocal hosts (Shaw \& Kent, 1999; R. W. Shaw, personal communication). Boxes in fine lines enclose members of sometimes polyphyletic clades G and B (Chapter 4).


## APPENDIX 15

Partial elongation factor-1 alpha (EF-1 $\alpha$ ) reference DNA sequences for species of Loma obtained from the majority nucleotide state among isolates and sequence clones (Chapter 4). Sequence labels: LsaREF = Loma salmonae; ShiREF = Loma embiotocia; TomREF = Loma sp . from Pacific tomcod; AtlREF = Loma morhua; PacREF = Loma sp . from Pacific cod; PolREF $=$ Loma sp. from walleye pollock; LFREF = Loma sp. from brook trout; and glugeaEF = Glugea plecoglossi sequence from Genbank Accession D32139).




LsaREF acttaaagcaccaacaaaggtaatgctgcacttatcaaaccaggacacgttttcagtgacaataagaatgaaccatgtgaaattgctgaagcagccaag 900 Shiref LinREF TomREF AtlREF PacRef
PolREF LFREF glugeaEE.

LsaREF
Shi REF
LinREE TomREE AtlREF PacREE POIREE
LFREE

LsaREE

ShiREF
LinREE
TomREF AtlREF
PacREF PolREf
LFREF
LFREF
glugeaEf
LsaREF
Shiref
LinREF
TomREE
At1REF
PacREE
PolREF
LFREF
glugeaEF...t..c...............g...gtc...t

## APPENDIX 16

Partial elongation factor-1 alpha (EF-1 $\alpha$ ) amino acid sequences for species of Loma (from reference sequences obtained from the majority nucleotide state among isolates and sequence clones) (Chapter 4). Sequence labels: LsaREF = Loma salmonae; ShiREF = Loma embiotocia; TomREF = Loma sp. from Pacific tomcod; AtlREF = Loma morhua; PacREF = Loma sp. from Pacific cod; PolREF = Loma sp. from walleye pollock; LFREF = Loma sp. from brook trout; and glugeaEF = Glugea plecoglossi sequence from Genbank Accession D32139).


## APPENDIX 17

Partial RNA polymerase largest subunit II (RPB1) DNA and amino acid sequences for Loma salmonae (Chapter 4). First three rows of sequence represent DNA, and last three rows represent amino acids. Sequence labels: L.sal = Loma salmonae; L.ace $=$ Loma acerinae obtained from Genbank Accession AJ278951; G.ano = Glugea anomala from Genbank Accession AJ278952).


## APPENDIX 18

Phylograms for data represented in Figure 4.6 of Chapter 4, for maximum likelihood (single best tree), maximum parsimony (one of 1000 most parsimonious trees), and distance (one of 90 shortest evolution trees) for Loma species, estimated from 479 alignment positions plus a gap matrix of 27 additional characters, from rDNA Region 3.




## APPENDIX 19

Consensus 50\% majority-rule maximum parsimony tree for Loma spp. (Chapter 4) with the inclusion of polymorphic characters (as "AND") generated from 479 alignment positions for rDNA Region 3, showing bootstrap values $>50 \%$ ( 1000 replicates) on branches.


## APPENDIX 20

Technique improvements suggested by results of this thesis.

## Collection of material:

1. Effective geographic breadth of sampling was greatly reduced in this study because many isolates collected failed to amplify due to problems listed in the next 3 sections.
2. Prevalence estimates depended on technique and species. For example, estimates were higher in histology than in wet mounts from Loma species with xenomas in the central venous sinus, while the reverse was true (wet mount prevalence was higher) for species with xenomas primarily in the secondary gill lamellae.

DNA extraction was made difficult in this study by some tissue-preparation techniques:

1. Spores can be quite resistant to ethanol (and also TEM fixatives), so tissue should be extracted fresh or frozen for better yields. Probably the best amplification in this study came from tissue that was infected with a greater percentage of earlier stages (non-spore stages).
2. Bead-beating, which is generally required to get good DNA yields from samples of fully mature spores, appears to produce more sheared DNA, making it difficult to clone large segments or amplify single-copy genes.
3. Green gland tissue of shrimp dissolved in ethanol, and became a significant PCR inhibitor. In future, tissue from shrimp should be extracted fresh, or frozen, and not preserved in ethanol.
4. Humic acid or other soil inhibitors in filtrates of copepods must be completely removed, by placing copepods in fresh distilled water, otherwise they inhibit PCR.
5. Spores may germinate or be digested by tissue lysis if fish are left for some time before fixation or freezing, making it impossible to obtain sequenceable DNA even from a heavily infected host (e.g. this appears to have occurred for a pile perch).

## PCR and sequencing was made more difficult than necessary in this study because:

1. Because of rDNA paralogs, the region should be amplified in long, continuous pieces, rather than partial segments, then these long rDNA regions can be cloned and each clone can be sequenced in smaller segments.
2. Sequencing directly from PCR products without cloning should not be done unless one is certain that paralogs do not differ by indels (or else it will be necessary to use the Flip Analyzer).
3. Purified spore DNA is difficult to clone (for partial genomic library construction) if antibiotics have been used to reduce bacteria accumulation in decaying gill tissue.
4. Single-copy primer design based on distantly related microsporidia was difficult because sequences tend to be highly diverged.
5. General primers risk getting bacteria or host contaminant sequence yet purification of spores for wild-caught specimens (e.g. at sea) can be difficult.

## TEM was made difficult in this study by:

1. Lots of microtome thick-section cutting was required to located xenomas in light infections, especially for smaller xenomas that were often cut-through. Gram stain of thick-sections helped locate small xenomas. Unfortunately, gram stain doesn't work on pre-spore stages.
2. Resin rarely infiltrated spores sufficiently to prevent the posterior vacuole from tearing under the electron beam. Countless specimens were lost this way.
3. Sampling spores for both DNA and TEM from an individual copepod was difficult.

[^0]:    ${ }^{1}$ DISCLAIMER: Species descriptions and the accompanying species-group names that follow are not intended to represent formal, official entries to the zoological record, but are intended as a part of a draft manuscript to be submitted for publication at a later time. Names given to these taxa should be regarded only as temporary, unofficial names, used here for the purpose of clarity. This is in compliance with Articles 7 to 9 of the International Code of Zoological Nomenclature, which states that a university thesis does not constitute a valid, recognized publication. Species names proposed temporarily in this chapter will not be used in Chapter 4, as these chapters are destined for simultaneous submission for publication as independent manuscripts. Please refer to Appendix 12 for species-name equivalents for Chapters 3 and 4.
    ${ }^{2}$ But data in Chapter 4 will show both species are valid, making $L$. morhua, not $L$. branchialis, the type species.

[^1]:    ${ }^{1}$ Species of Loma described in Chapter 3 do not constitute valid species according to Articles 7 to 9 of the International Code of Zoological Nomenclature. See "DISCLAIMER" in Chapter 3. Chapters $3 \& 4$ are intended as draft manuscripts to be submitted for publication simultaneously and independently, therefore temporary, unofficial species names used for clarity in Chapter 3 will not be used in this chapter. A table of speciesname/isolate equivalences across chapters is presented in Appendix 12.

[^2]:    ${ }^{2}$ See Note 1. (Loma species described in Chapter 3 are do not constitute valid species according to the International Code of Zoological Nomenclature, and should be regarded as tentative descriptions only. Name equivalences between Chapters 2, 3 and 4 are shown in Appendix 12).

[^3]:    ' DISCLAIMER: This description does not constitute a valid publication according to the International Code of Zoological Nomenclature, and as a result, the new name should be regarded as temporary and unofficial.

[^4]:    ${ }^{1}$ Species described in Chapters 3 and 6 do not constitute valid species according to Articles 7 to 9 of the International Code of Zoological Nomenclature. See "DISCLAIMER" in Chapter 3. A table of speciesname/isolate equivalences across chapters is presented in Appendix 11.

[^5]:    ${ }^{2}$ see footnote 1 . Species described in Chapters 3 and 6 do not constitute valid species according to Articles 7 to 9 of the International Code of Zoological Nomenclature. See "DISCLAIMER" in Chapter 3. A table of speciesname/isolate equivalences across chapters is presented in Appendix 11.

[^6]:    Figures A1.28-A1.29: Screen images in showing weighting menu. Fig. A1.28 shows menu that results from choosing "Processes" and "Create step matrix". The user is prompted to input any value for transitions $\alpha$ and transversions $\beta$. These values will be relative to empirically will convert the relative values into weights for a step matrix.

