

MOLECULAR PHYLOGENY OF THE OXYMONADS

by

AARON ANDREW HEISS

B.A., Bennington College, 1995

B.S., Portland State University, 2002

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA

August 2006

ABSTRACT

The oxymonads are a group of structurally complex anaerobic flagellates about which we know very little. They are found in association with complex microbial communities in the guts of animals. There are five recognized families of oxymonads; molecular data have been acquired for four of these. Here, I describe the first molecular data from the last remaining group, represented by *Saccinobaculus*, an organism that is found exclusively in the hindgut of the wood-eating cockroach *Cryptocercus*. I sequenced small subunit ribosomal RNA (SSU rRNA) from total gut DNA to describe *Saccinobaculus* SSU rRNA diversity. I also sequenced SSU rRNA from manually isolated cells of the two most abundant and readily identifiable species: the type species *S. ambloaxostylus* and the taxonomically contentious *S. doroaxostylus*. I inferred phylogenetic trees including all five known oxymonad subgroups in order to elucidate the internal phylogeny of this poorly-studied group, to resolve some outstanding issues of the taxonomy and identification of certain *Saccinobaculus* species, and to investigate the evolution of character states within it. My analysis recovered strong support for the existence of the five subgroups of oxymonads, and consistently grouped the subgroups containing *Monocercomonoides* and *Streblomastix*, but was unable to resolve any further higher-order branching patterns. Additionally, I sequenced alpha-tubulin from *S. ambloaxostylus*, which in conjunction with environmentally obtained tubulin sequences established the use of an alternate genetic code by organisms related to *Monocercomonoides* and not by *Saccinobaculus*. This same alternate code is also used by *Streblomastix*, further cementing the relationship between that organism and the group including *Monocercomonoides*.

TABLE OF CONTENTS

Abstract	ii
Table of Contents	iii
List of Figures	v
Acknowledgements	vi
Dedication	vii
Co-Authorship Statement	viii
1 Introduction	1
1.1 The Cockroach <i>Cryptocercus</i> and Its Symbionts	1
1.2 The Oxymonads	3
References	6
2 Small Subunit Ribosomal RNA	9
2.1 Introduction	9
2.2 Materials and Methods	10
2.2.1 Sample Acquisition, Single Cell Isolation, and Molecular Methods	10
2.2.2 Phylogenetic Analyses	11
2.2.3 Light Microscopy	12
2.3 Results and Discussion	13
2.3.1 Surveying Phylogenetic Diversity of Oxymonads in <i>Cryptocercus</i> by Environmental PCR	13
2.3.2 Single Cell Isolation of <i>S. ambloaxostylus</i> and <i>S. doroaxostylus</i>	13
2.3.3 Relationships between Oxymonad Subgroups	15
References	24

3 Protein-Coding Genes	26
3.1 Introduction	26
3.2 Materials and Methods	28
3.3 Results and Discussion	29
3.3.1 Alpha-Tubulin Phylogeny and Identification of the Alternate-Code- Using Lineage	29
3.3.2 Evolution within the Oxymonads	30
References	34
4 Overall Conclusions and Directions for Future Research	35
4.1 Conclusions	35
4.2 Future Directions	36
References	37

LIST OF FIGURES

2.1 SSU Tree of Eukaryotes	19
2.2 SSU Tree of Oxymonads	20
2.3 Light Micrographs of Saccinobaculus	21
2.4 SSU Tree of Subset of Oxymonads	23
3.1 Alpha-Tubulin Tree of Eukaryotes	32

ACKNOWLEDGEMENTS

This work would not have been possible without the financial and academic support of my supervisor, Dr. Patrick J. Keeling. I also received support from the Frances Chave Memorial Scholarship and from the University of British Columbia's University Graduate Fellowship programme. Advice from my committee, comprising Drs. Brian S. Leander and James D. Berger, was also helpful and appreciated. Members of the Keeling, Leander, and Fast labs were essential to the success of my research. In particular, Audrey P. de Koning, Matthew B. Rogers, Dr. Claudio H. Slamovits, and Dr. Bryony A. P. Williams taught me laboratory techniques and computer analysis, and assisted in the interpretation of data. I thank Christine A. Nalepa for her collecting and supplying me with the cockroaches necessary for this work. Finally, my partner in life, Lisa A. Paul, deserves and receives my deepest appreciation for her emotional support, and for never having complained about her partner having turned into a graduate student.

DEDICATION

I dedicate this work to the fond memory of Dr. Richard B. Forbes. His enviable breadth and depth of knowledge, contagious passion for his subject, and untiring attention to his students were in no small part responsible for my decision to pursue a career in biology; he will forever remain a role model for me.

CO-AUTHORSHIP STATEMENT

A version of Chapter 2, on the small subunit ribosomal RNA of *Saccinobaculus* and the associated work done on oxymonad phylogeny, has been accepted for publication in the journal *Protist*. Patrick Keeling is listed as co-author on this paper, and was instrumental in guiding the approach of this research. He isolated all cells of *S. doroaxostylus* (our sample I1) and one set of cells of *S. ambloaxostylus* (our samples I2a and I2b), while I performed the remainder of the isolations (four total). Dr. Keeling sequenced sample I1, while I sequenced all samples of *S. ambloaxostylus*, and I carried out all laboratory work on all of the environmental samples. Dr. Keeling was responsible for all photography. I performed all of the data analysis and wrote the first draught of the paper. This was revised by Dr Keeling, and then further by me.

Research done in Chapter 3 relies heavily on work performed by Audrey de Koning, Geoff Noble, and Jensen Wong, who obtained all of the environmental sequences from the gut of the woodroach used in my analysis. I appear as co-author on a paper describing these results which has been submitted for publication. This, however, is a separate entity from my thesis chapter, the writing of which was solely the product of my own efforts; the analyses presented here are also all mine.

1: INTRODUCTION

1.1: The Cockroach *Cryptocercus* and Its Symbionts

The genus *Cryptocercus* represents a group of populations of damp-wood-eating insects (Cleveland *et al.* 1934). It has traditionally been considered a cockroach based on its morphology, a classification supported by some studies, which go further in placing it basal in that lineage (for instance, (Kambhampati 1995)). Other studies turn this phylogeny on its head, instead placing *Cryptocercus* as basal to termites, and the whole termite-*Cryptocercus* clade as derived cockroaches (Lo *et al.* 2000). This latter possibility, besides being more rigorous in its methodology, is attractive on the grounds of the similarities in natural histories between *Cryptocercus* and the “lower termites”. It also emphasises the fact that members of the two groups each possess a similar diversity of microbial symbionts in their hindguts. It is undermined, however, by the presence of two other groups of xylophagous cockroaches which are unrelated to each other, to *Cryptocercus*, and to termites (Pellens *et al.* 2002). Nevertheless, all groups appear to possess the same groups of protists (Brugerolle *et al.* 2003).

The functions of these protists are not clear. Early studies showed that the entire complement of protists is necessary for the survival of the host insect; when treated to kill its protist symbionts, the host starved, while hosts so treated recovered after reintroduction of the protists (Cleveland *et al.* 1934). These studies also showed that it was specifically the protists that were required for the cockroach’s continued health: the cockroach starved as well when treated to kill the protists while leaving the abundant prokaryotic biota alive. This work also established that it was the protists and not the bacteria that were responsible for the cellulase necessary to breaking down the wood in the cockroach’s diet. However, while these studies did show that the cockroach was dependent on its gut protists for the digestion of its food, they did not determine which of the protists were responsible for that digestion. In other words, it is possible that some of the protists comprising the gut biota are commensals or parasites, and indeed it is common to see the latter term to describe the relationship (as in the overview given in (Brugerolle and Lee 2000a)). At this stage, the matter is still undecided, and indeed has barely been investigated since Cleveland’s pioneering work over seventy years ago.

The protists present in the gut of *Cryptocercus*, and indeed of all of its xylophagous relatives (termites and other wood-eating cockroaches), are members of a number of

lineages within a contentious group of eukaryotes called the excavates (Simpson 2003). The common membership of the xylophagous symbionts in this group appears in this case to be coincidental, however, and not the result of common descent: free-living organisms are basal to at least two of the groups found in such environments, so the ancestor common to these groups is almost certainly to have been itself free-living. Indeed, while some groups of excavates have been unambiguously linked in multiple studies using many different approaches, they tend to form three or four groups in molecularly based studies, and these groups show little tendency to branch together. It is possible that the excavates are a paraphyletic group, representing a variety of lineages basal to other eukaryotic groups. In any event, the groups containing the symbionts of *Cryptocercus* and its kin tend not to branch together in these studies, furthering the impression that any common classification is unrelated to their common environment.

One of the more spectacular of these groups is the hypermastigotes. These organisms are members of a larger group called the parabasalids, none of which have mitochondria (although they do have relic organelles known as hydrogenosomes), and all of which have a characteristically developed Golgi apparatus which is called the parabasal body (from which term the group's name is derived) and generally one or more axostyles per cell. The axostyle is in this case a hollow tube made up from a sheet of microtubules which tends to flare out into a funnel around the nucleus; parabasalids with multiple nuclei (of which there are several genera) tend as well to have multiple axostyles, one per nucleus. A couple of members of this group are free-living, but most are symbionts of the digestive and urogenital tracts of animals (Brugerolle and Lee 2000b).

The diplomonads are another group represented in the xylophagous-insect symbiote, albeit somewhat cryptically. Members of this group are tiny, and therefore not readily identified amongst the gut biota while still alive, but are nevertheless morphologically fascinating. They too lack mitochondria, and although the relics of mitochondria have been identified, they are not as apparent as the hydrogenosomes of the parabasalids. More startlingly, they also lack a Golgi apparatus. The cellular anatomy in most follows a curious biradial pattern in which all organelles are duplicated about the cell's central axis, although this differs from the highly-derived bilateral symmetry

observed in the best-known diplomonad, *Giardia*. Each half of the cell includes its own nucleus and four associated flagella. As with the parabasalids, the numbers of free-living species are few compared with those of symbionts or parasites of animal digestive tracts (Brugerolle 2000).

The oxymonads are the final group of symbionts found in the xylophagous insects, and are quite possibly the most poorly studied group of eukaryotes yet known. Not only do they lack mitochondria, no relic of this organelle has yet been unambiguously identified, either. Like the diplomonads, they lack a Golgi apparatus, and like the parabasalids, they have an axostyle, although the axostyle in this case is a solid bundle of microtubules rather than a hollow tube, and is not homologous. Also like the parabasalids, some (presumably more derived) members of the group are multinucleate, and each nucleus can have its own axostyle. There are no free-living members of this group known; all but a very small number of species are symbionts of xylophagous insects, and the few exceptions are symbionts of vertebrates (Brugerolle and Lee 2000a). It is the oxymonads that will form the focus of my study.

1.2: The Oxymonads

There are five subgroups within the oxymonads; since the group as a whole has been given the rank of order, the subgroups within it have been given the rank of family (Brugerolle and Lee 2000a). Additionally, a sixth group, representing the oxymonads' closest non-oxymonad relative, has been identified (Dacks *et al.* 2001). This is the free-living flagellate *Trimastix*, an organism representative of the excavate cell form (Simpson *et al.* 2002): a feeding groove is found along the ventral surface of the cell. Four flagella are found at the anterior end of the cell in association with the nucleus, one of which follows the feeding groove to the posterior tip of the cell, one of which projects anteriorly, and the remaining two of which project laterally. The posterior flagellum wafts food into the feeding groove, and the anterior flagellum is the primary source of locomotion. It lacks mitochondria, but organelles resembling and undoubtedly descended from mitochondria have been identified (Brugerolle and Patterson 1997). Remarkably, no other organism has ever been shown or indeed suggested to have a close relationship with the oxymonads; the only other proposed relatives are other excavates, and no confident relationships amongst them have been put forth (Simpson 2003; Hampl *et al.* 2005).

Amongst the oxymonads themselves, the simplest subgroup from a morphological perspective is the Polymastigidae. These small (generally on the order of 5 μm long) flagellates have a comparatively thin axostyle and four flagella; ultrastructurally, they are the group that presents the greatest resemblance to the typical excavate cell (Simpson *et al.* 2002). Two members of this group, *Monocercomonoides* and *Paranotila*, are to be found in *Cryptocercus*. A third genus, *Polymastix*, is widely distributed, as is *Monocercomonoides*; *Paranotila* is found only in *Cryptocercus* (Brugerolle and Lee 2000a). One strain of *Monocercomonoides* bears the distinction of being the only oxymonad ever maintained in culture outside of a host animal (Hampl *et al.* 2005).

The Saccinobaculidae comprises two genera, *Saccinobaculus* and *Notila*, both of which are found only in *Cryptocercus*. These organisms can be as small in size as the polymastigids, but some species can be almost 200 μm long. They have four flagella and a prominent axostyle, which writhes and thrashes about spectacularly in the cell body (Brugerolle and Lee 2000a). This latter feature has been investigated as a model of microtubule movement (Grimstone and Cleveland 1965; McIntosh 1973; McIntosh *et al.* 1973). There are an unknown number of species of this group: originally, three were described, but with the understanding that variations amongst the members of each species could suffice to split them into as many as nine (Cleveland *et al.* 1934). Later work split the genus *Saccinobaculus* and as well transferred some members to the genus *Oxymonas* (a member of another family, described below) (Cleveland 1950c, 1950b, 1950a), compounding the initial confusion, which has never been directly addressed in later work.

The Streblomastigidae contains one species, *Streblomastix strix*, which is found in the termite *Zootermopsis*. It is a long, thin cell, typically 20-50 μm long and 6-10 μm across. This organism is remarkable in its use of ectosymbiotic bacteria, which adhere to the plasma membrane in rows running from anterior to posterior. The bacteria appear to have a structural function, since after treatment to remove them, the host *Streblomastix* cell shortens anteroposteriorly and swells laterally, radically changing its appearance (Leander and Keeling 2004). Otherwise, *S. strix* has the usual four flagella, a short and moderately contractile axostyle, and a holdfast, a microfibrillar organelle with which the cell attaches to the lumen of the host insect's gut. The holdfast is not always present, however; the cells have also been seen to be free-living (Brugerolle and Lee 2000a).

Two genera, *Dinenympha* and *Pyronympha*, make up the family Pyronymphidae. This group is characterised by a moderately contractile axostyle and a set of flagella that wrap about the cell body in a spiral fashion, protruding at the posterior end for some distance. *Dinenympha* is motile, has four flagella, is fusiform, and typically measures 12-

30 μm in length, while *Pyrrsonympha* attaches to the host gut lumen with a holdfast, has eight flagella, is pear- or teardrop-shaped with the wider portion of the cell body posterior, and measures 150-200 μm (Brugerolle and Lee 2000a). The taxonomy of this group has a confusing history, as the two genera were for some time thought to be different life-stages of the same organism. However, recent studies showed them to be distinct organisms (Moriya *et al.* 2003), albeit unquestionably related.

The genus for which the whole group is named gives its name as well to the final family, the Oxymonadidae. This family includes, besides *Oxymonas*, the genera *Sauromonas*, *Barroella*, and *Microrhopalodina*, all generally from about 20 to 170 μm in length, and shows a great deal of morphological variation. *Oxymonas* is perhaps the least derived member of the group, with a single nucleus and four flagella. *Sauromonas* has a sessile *Oxymonas*-like stage in its life cycle, and another motile stage, in which the number of flagella is greatly increased. *Barroella* resembles a multinucleate *Oxymonas* with poorly developed or absent flagella and numerous axostyles scattered through the cytoplasm. *Microrhopalodina* is a strawberry-shaped cell with several nuclei arranged in a ring about the anterior constriction of the cell, each with its own axostyle and its own complement of four flagella. All of these genera also exhibit a further organelle, the rostellum, which is a bundle of microtubules that extends anteriorly between the nucleus and the holdfast (always one per cell), and develops separately from the axostyle (Brugerolle and Lee 2000a).

The relations amongst these groups have been tentative, based primarily on light-microscopical studies. Cellular simplicity has been taken to indicate a basal character, and complexity to be derived; the presence or absence of characteristic organelles has been used to produce simple parsimony-based hypotheses of relationships. Using such criteria, the Polymastigidae are obviously basal to the other groups, on the basis of its members' lack of both axostylar contractility and of any attachment organelle. Those groups featuring a holdfast are then grouped together, with the Oxymonadidae regarded as the most highly derived (Moriya *et al.* 2003), this group being distinguished by the presence of the rostellum in all and of the multinucleate condition in some of its members.

However, it has not been until very recently that any of these assumptions were tested using molecular methods. The first such studies emphasised the Pyrrsonymphidae (Moriya *et al.* 1998; Dacks *et al.* 2001; Moriya *et al.* 2001) and attempted more to locate the oxymonads within the greater scheme of the eukaryotes than to sort out the relations within the oxymonads themselves. Indeed, it was not until 2003 that the long-standing debate over the separate identity of *Dinenympha* from *Pyrrsonympha* was addressed using

molecular data (Moriya *et al.* 2003). Further work continued to emphasise the Pyrsonymphidae (Stingl and Brune 2003) but began to incorporate other groups as well (Keeling and Leander 2003), culminating in a study using a wide selection of genes from the first oxymonad kept in culture (Hampl *et al.* 2005). This last study was also the first to investigate the relationships within the group using purely molecular methods. It confirmed the identity of the Pyrsonymphidae and found that the Streblomastigidae and the Polymastigidae are related to the exclusion of all other groups, but could assert no further relationships.

Such is the state of our knowledge of the phylogenetic relationships amongst the oxymonads at the beginning of my work. Here I present the first molecular data to cover all oxymonad groups, and my findings based on those data. As well, I interpret these data in the context of morphological evolution within the oxymonads, and evaluate the pertinence of such data. I also address some of the taxonomic uncertainties that have arisen in the long but spotty history of the scientific awareness of this group. These uncertainties, and those regarding the phylogeny of the group as a whole, are far from completely dispelled, but nevertheless I have made some progress towards that end.

References

- Brugerolle, G. (2000). Order Diplomonadida. *Illustrated Guide to the Protozoa*. J. J. Lee, G. F. Leedale and P. Bradbury. Lawrence, KA, Society of Protozoologists: 1125-1135.
- Brugerolle, G. and Lee, J. J. (2000a). Order Oxymonadida. *Illustrated Guide to the Protozoa*. J. J. Lee, G. F. Leedale and P. Bradbury. Lawrence, KS, Society of Protozoologists: 1186-1195.
- Brugerolle, G. and Lee, J. J. (2000b). Phylum Parabasalia. *Illustrated Guide to the Protozoa*. J. J. Lee, G. F. Leedale and P. Bradbury. Lawrence, KS, Society of Protozoologists: 1196-1250.
- Brugerolle, G. and Patterson, D. (1997). "Ultrastructure of *Trimastix convexa* Hollande, an amitochondriate anaerobic flagellate with a previously undescribed organization." *European Journal of Protistology* **33**(2): 121-130.
- Brugerolle, G., Silva-Neto, I. D., Pellens, R. and Grandcolas, P. (2003). "Electron microscopic identification of the intestinal protozoan flagellates of the xylophagous cockroach *Parasphaeria boleiriana* from Brazil." *Parasitology Research* **90**(3): 249-256.

Cleveland, L. R. (1950a). "Hormone-induced sexual cycles of flagellates: II. Gametogenesis, fertilization, and one-division meiosis in *Oxymonas*." *J. Morphol.* **86**: 185-214.

Cleveland, L. R. (1950b). "Hormone-induced sexual cycles of flagellates: III. Gametogenesis, fertilization, and one-division meiosis in *Saccinobaculus*." *J. Morphol.* **86**: 215-228.

Cleveland, L. R. (1950c). "Hormone-induced sexual cycles of flagellates: IV. Meiosis after syngamy and before nuclear fusion in *Notila*." *J. Morphol.* **87**: 317-348.

Cleveland, L. R., Hall, S. R., Sanders, E. P. and Collier, J. (1934). "The wood-feeding roach *Cryptocercus*, its protozoa, and the symbiosis between protozoa and roach." *Mem. Am. Acad. Arts Sci.* **17**(2): 185-342.

Dacks, J. B., Silberman, J. D., Simpson, A. G., Moriya, S., Kudo, T., Ohkuma, M. and Redfield, R. J. (2001). "Oxymonads are closely related to the excavate taxon *Trimastix*." *Mol Biol Evol* **18**(6): 1034-1044.

Grimstone, A. V. and Cleveland, L. R. (1965). "The Fine Structure and Function of the Contractile Axostyles of Certain Flagellates." *J Cell Biol* **24**: 387-400.

Hampl, V., Horner, D. S., Dyal, P., Kulda, J., Flegr, J., Foster, P. G. and Embley, T. M. (2005). "Inference of the phylogenetic position of oxymonads based on nine genes: support for metamonada and excavata." *Mol Biol Evol* **22**(12): 2508-2518.

Kambhampati, S. (1995). "A phylogeny of cockroaches and related insects based on DNA sequence of mitochondrial ribosomal RNA genes." *Proc Natl Acad Sci U S A* **92**(6): 2017-2020.

Keeling, P. J. and Leander, B. S. (2003). "Characterisation of a non-canonical genetic code in the oxymonad *Streblomastix strix*." *J. Mol. Biol.* **326**: 1337-1349.

Leander, B. S. and Keeling, P. J. (2004). "Symbiotic innovation in the oxymonad *Streblomastix strix*." *J Eukaryot Microbiol* **51**(3): 291-300.

Lo, N., Tokuda, G., Watanabe, H., Rose, H., Slaytor, M., Maekawa, K., Bandi, C. and Noda, H. (2000). "Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches." *Curr Biol* **10**(13): 801-804.

McIntosh, J. R. (1973). "The axostyle of *Saccinobaculus*. II. Motion of the microtubule bundle and a structural comparison of straight and bent axostyles." *J Cell Biol* **56**(2): 324-339.

McIntosh, J. R., Ogata, E. S. and Landis, S. C. (1973). "The axostyle of *Saccinobaculus*. I. Structure of the organism and its microtubule bundle." *J Cell Biol* **56**(2): 304-323.

Moriya, S., Dacks, J. B., Takagi, A., Noda, S., Ohkuma, M., Doolittle, W. F. and Kudo, T. (2003). "Molecular phylogeny of three oxymonad genera: *Pyrrsonympha*, *Dinenympha* and *Oxymonas*." *J Eukaryot Microbiol* **50**(3): 190-197.

Moriya, S., Ohkuma, M. and Kudo, T. (1998). "Phylogenetic position of symbiotic protist *Dinenympha* [correction of *Dinemympha*] exilis in the hindgut of the termite *Reticulitermes speratus* inferred from the protein phylogeny of elongation factor 1 alpha." *Gene* **210**(2): 221-227.

Moriya, S., Tanaka, K., Ohkuma, M., Sugano, S. and Kudo, T. (2001). "Diversification of the microtubule system in the early stage of eukaryote evolution: elongation factor 1 alpha and alpha-tubulin protein phylogeny of termite symbiotic oxymonad and hypermastigote protists." *J Mol Evol* **52**(1): 6-16.

Pellens, R., Grandcolas, P. and da Silva-Neto, I. D. (2002). "A new and independently evolved case of xylophagy and the presence of intestinal flagellates in the cockroach *Parasphaeria boleiriana* (Dictyoptera, Blaberidae, Zetoborinae) from the remnants of the Brazilian Atlantic forest." *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **80**(2): 350-359.

Simpson, A. G. (2003). "Cytoskeletal organization, phylogenetic affinities and systematics in the contentious taxon Excavata (Eukaryota)." *Int. J. Syst. Evol. Microbiol.* **53**(Pt 6): 1759-1777.

Simpson, A. G., Radek, R., Dacks, J. B. and O'Kelly, C. J. (2002). "How oxymonads lost their groove: an ultrastructural comparison of *Monocercomonoides* and excavate taxa." *J. Eukaryot. Microbiol.* **49**(3): 239-248.

Stingl, U. and Brune, A. (2003). "Phylogenetic diversity and whole-cell hybridization of oxymonad flagellates from the hindgut of the wood-feeding lower termite *Reticulitermes flavipes*." *Protist* **154**(1): 147-155.

2: SMALL SUBUNIT RIBOSOMAL RNA*

2.1: Introduction

The members of the oxymonad family Saccinobaculidae can be quite spectacular. They are found in a single natural community within the hindgut of the wood-eating cockroach *Cryptocercus*, although in that community they are abundant and diverse. Members of the family are morphologically so plastic that they defy ready description (Cleveland *et al.* 1934): like all oxymonads, they have a prominent axostyle running along the length of the cell, which in the case of *Saccinobaculus* twists, coils, writhes, and straightens in a variety of changing patterns within the cell. This movement creates the appearance of a snake (the axostyle) attempting to escape from a bag (the plasma membrane), from which the name is derived (Cleveland *et al.* 1934).

All members of the Saccinobaculidae were originally described by Cleveland *et al.* (Cleveland *et al.* 1934), who at first classified all of the oxymonads in *C. punctulatus* in the genus *Saccinobaculus*, with the single exception of the small flagellate *Monocercomonoides globus*. Three distinct species were described, *S. ambloaxostylus* (the type species, and the most abundant), *S. doroaxostylus* (the largest species, which is in addition morphologically and behaviorally distinctive from *S. ambloaxostylus*), and *S. minor*, a looser collection of small flagellates. Beyond these three species, Cleveland *et al.* initially suggested that as many as six more might exist, but the extent of the diversity was difficult to determine at the time (Cleveland *et al.* 1934). In subsequent years, subtle details of staining properties, observations of granules on or within the axostyle, and studies of reproductive life histories led to a whittling away of *Saccinobaculus* diversity as species were reassigned or new groups created. Most significantly, the type species *S. ambloaxostylus* was split to create the new genus *Notila* (Cleveland 1950c), *S. doroaxostylus* was transferred to *Oxymonas* as *Oxymonas doroaxostylus* (Cleveland 1950a), and the smaller *Saccinobaculus* species were transferred to *Oxymonas* as *O. nana* (Cleveland 1950a).

* A version of this chapter has been accepted for publication as Heiss, A. A. and Keeling, P. J. (2006): "The Phylogenetic Position of the Oxymonad *Saccinobaculus* Based on SSU rRNA". *Protist*, Volume 157, Issue 3.

Here, we provide the first molecular data, small subunit rRNA (SSU rRNA), from the Saccinobaculidae, and by doing so allow the first molecular-based analysis of global oxymonad phylogeny. We combined an environmental-PCR strategy (sequencing SSU rRNA from the whole gut) to examine overall diversity of the Saccinobaculidae, with the manual isolation and characterization of two easily identifiable and key species: the type species *S. ambloaxostylus* and the taxonomically ambiguous *S. doroaxostylus* (*O. doroaxostylus*).

2.2: Materials and Methods

2.2.1: Sample acquisition, single cell isolation, and molecular methods.

Specimens of the insect host *Cryptocercus punctulatus* were collected by C. Nalepa from Mountain Lake Biological Station, Giles Co., Virginia; Black Rock Mountain, Rabun Co., Georgia; and Bear Trap Gap, Haywood Co., North Carolina (Nalepa *et al.* 2002). Gut contents of one cockroach (from Mountain Lake Biological Station) were obtained by massaging fluid out of the insect's abdomen and suspending it in Trager's medium U (Trager 1934), and whole gut DNA was extracted from this suspension using the DNEasy kit (QIAGEN). For single-cell isolations, gut contents of several cockroaches (from Black Rock Mountain and Bear Trap Gap) were obtained either as above or by dissecting the cockroach and suspending the gut contents in either Trager's medium U or 0.4% sodium chloride. Cells were isolated and DNA was extracted as described (Keeling and Leander 2003).

Oxymonad SSU rRNA was amplified from environmental samples using an oxymonad-specific forward primer 5'-GCCACTACTCATATGCTTGTCTC-3' and a eukaryote-specific reverse primer 5'-AACCTTGTTACGACTTCTCCTTCC-3' with an annealing temperature of 52° and an extension time of 2 minutes. SSU rRNA was amplified from manually isolated cells using eukaryote-specific primers 5'-GCGCTACC-TGGTTGATCCTGCC-3' and 5'-TGATCCTTCTGCAGGTTACCTAC-3' with an annealing temperature of 45°C and an extension time of 1.5 minutes. In all cases products of approximately 2,100 bp were obtained, which were cloned using TOPO TA cloning (Invitrogen), and several clones were sequenced on both strands.

The environmental samples yielded 14 sequences, here called E1-E14. The single-cell isolates yielded a total of 12 sequences: one sequence (here called I1) from 40 large pigmented cells of *S. doroaxostylus* from the Black Rock Mountain population; two sequences each from isolates of 27 (I2a and I2b), 10 (I3a and I3b), approximately 12 (I4a and I4b), and four (I5a and I5b) cells of *S. ambloaxostylus*, all taken from the Bear Trap Gap population; and three sequences (I6a, I6b, and I6c) from eight cells of *S. ambloaxostylus* taken from the Black Rock Mountain population. New sequences were deposited in GenBank as accessions DQ525690-DQ525715.

2.2.2: Phylogenetic analyses. We generated an alignment of the small subunit ribosomal RNA (SSU rRNA) from inspection of secondary structure of a number of eukaryotes for which such data is available. To this, we included SSU sequences from selected other taxa, all previously published oxymonad sequences, and our experimentally generated sequences (alignment available upon request). To ensure that all new sequences were derived from oxymonads, this alignment included 83 taxa representing all major groups of eukaryotes, including the host insect *Cryptocercus* and parabasalian relatives of other symbionts known to be present. Global analysis of eukaryotes and two oxymonad-specific analyses were performed using 1089 positions and 83 taxa. Oxymonad-specific analyses allowed the use of 1479 alignable sites: one including 50 taxa with all publicly available full-length oxymonad and *Trimastix* sequences, and a second including 13 taxa (two to three representatives of each of the six groups, except for *Monocercomonoides*, for which only one sequence was available).

We generated trees from these datasets using PHYML 2.4.4 (Guindon and Gascuel 2003) using the GTR substitution model (except for the 83-taxon dataset, for which we used the HKY model). Site-to-site rate variation was modeled on a gamma distribution with 8 variable rate categories and invariable sites, with the shape parameter alpha and proportion of invariable sites estimated from the data. Alpha was estimated at 0.538101, 0.378790, and 0.282216 for the 83-, 50-, and 13-taxon datasets, respectively, and proportions of invariants were calculated to be zero for the first two and 0.067 for the last. We also carried out Bayesian analyses on the two smaller data sets using MrBayes

3.0B4 (Huelsenbeck and Ronquist 2001) with these same parameters. 1,000,000 generations with one cold and three hot chains were sampled every 1,000 generations; the first 33 and 2 trees, respectively, from the 50- and 13-taxon datasets were discarded as burn-in, and consensus trees were generated from the remainder. The branch lengths of the consensus topology were calculated by maximum likelihood using PHYML. Bootstrap resampling of the 83-taxon dataset was based on 100 replicates using both maximum likelihood and the BioNJ distance methods provided by PAUP* 4.0b10 (Swofford 2002). We analysed 1,000 bootstrap replicates of both the 50- and 13-taxon datasets with both maximum likelihood (using PHYML) and neighbor joining (using the BioNJ method as implemented by PAUP*). Additionally, we calculated 1,000 bootstrap replicates for the 13-taxon data set using maximum likelihood with stepwise addition (using PAUP*).

The relations between the various groups of oxymonads were investigated using Approximately Unbiased (AU) tests (Shimodaira 2002). Eleven taxa were selected such that each group of oxymonads was represented by two taxa, except for *Monocercomonoides* (for which, again, only a single sequence is known); these were constrained to provide six groups. All 105 possible topologies for these six groups were generated using PAUP*, and site likelihoods for each calculated using TREE-PUZZLE 5.1 (Schmidt *et al.* 2002) with the parameters described above. AU tests were performed using CONSEL 1.19 (Shimodaira and Hasegawa 2001). Parallel tests were also carried out with the nine-taxon dataset obtained by excluding the divergent pyrsonymphid group.

2.2.3: Light Microscopy. Cockroaches were dissected and protist fauna suspended in Trager's medium U as described for DNA isolation. Live cells were observed by Nomarski differential interference contrast microscopy using a Zeiss Axioplan 2 microscope with 63X and 100X Plan-Apochromat lenses. Digital images were taken using a QImaging MicroImager II monochrome camera. Behaviour was observed on live cells soon after dissection, while some structural features were observed on cells once they had ceased to move, so that structural features could be observed more clearly.

2.3: Results and Discussion

2.3.1: Surveying phylogenetic diversity of oxymonads in *Cryptocercus* by environmental PCR. We surveyed SSU rRNA sequence diversity from the *Cryptocercus* gut in order to assess the phylogenetic breadth of oxymonads in this environment. Amplification with oxymonad-specific primers yielded a single detectable band of approximately 2,100 bp. This fragment was cloned and sequenced, resulting in 14 unique sequences, all most similar to other oxymonad rRNAs. Phylogenetic analysis with the global eukaryotic alignment confirmed this by placing all new sequences within the oxymonad/*Trimastix* clade with 96% support (Figure 2.1). Oxymonads generally have larger-than-average SSU rRNA sequences, ranging from about 2,000 to 2,900 bp (by contrast, the average eukaryotic sequence is 1,800 bp). The oxymonad genes all have insertions at various points, many of which are shared amongst the groups of oxymonads (Hampl *et al.* 2005).

We narrowed our focus to oxymonads and *Trimastix* to allow a greater number of alignable sites to be used, a phylogeny of which is shown in Figure 2.2. Here, as in the global analysis, the environmental sequences (our *Saccinobaculus* spp. E1-E14) represent an appreciable degree of sequence diversity, and yet form a single clade with 93-100% support to the exclusion of all other oxymonad groups. Overall, the environmental sequences are all derived from a single oxymonad lineage, and there is no evidence of close relatives of other sampled oxymonad groups that have been proposed to exist in *Cryptocercus*, such as *Monocercomonoides* and *Oxymonas*.

2.3.2: Single cell isolation of *S. ambloaxostylus* and *S. doroaxostylus*. To identify the source of the oxymonad rRNA genes amplified in environmental sample, we manually isolated samples of the two most abundant and conspicuous species of oxymonad in the *Cryptocercus* gut: *S. ambloaxostylus* and *S. doroaxostylus*.

Saccinobaculus ambloaxostylus is the most abundant oxymonad in *Cryptocercus* and is the type species for the genus *Saccinobaculus*. It is intermediate in size (65-110 μm) amongst *Cryptocercus* symbionts and extremely plastic in form (Cleveland 1950b). This is demonstrated in Figure 2.3E-J, which shows a variety of the major features that

allowed us to identify the species, and also shows the plasticity of the cell due to the motile axostyle, resulting in several distinct appearances. Because of this great degree of plasticity, we repeated our isolations five times from different samples (four times from Bear Trap Gap cockroaches and once from Black Rock Mountain cockroaches) to confirm that they all correspond to a single sequence type. Phylogenetic analysis shows that all isolated specimens (sequences *S. ambloaxostylus* I2-I6) are closely related, and indeed form a group within the environmental sequences (Figure 2.2). As expected, the isolated *S. ambloaxostylus* sequences were nearly identical to one of the environmental sequence groups (*S. ambloaxostylus* E4-E8), but interestingly the sequences obtained from single-cell isolations formed distinct subgroups within the tree generated from environmental sequences. The cockroaches collected from Mountain Lake Biological Station, from which all environmental data were generated, are of a different karyotype ($2n=43$) from those of Bear Trap Gap ($2n=45$) and Black Rock Mountain ($2n=39$), suggesting that distinct populations of cockroaches also contain distinct populations of symbionts. Whereas the Black Rock Mountain group (I6) branched within the Bear Trap Gap group (I2-I5), all three of its sequences formed a clade within that group (Figure 2.2). This is the first evidence for such population variation in these protist endosymbiont communities, suggesting the symbionts and hosts may be co-evolving.

The other conspicuous oxymonad in *Cryptocercus* is *S. doroaxostylus*, which is larger than *S. ambloaxostylus* (its size range is reported to be 150-170 μm), and is also distinguished by brown-pigmented cytoplasmic granules that we infer to be a carbohydrate storage product. The cell is less plastic than *S. ambloaxostylus* but does have a motile axostyle with a characteristic protruding tip and a distinctive pyriform nucleus (Figure 2.3A-D). Originally described as a member of *Saccinobaculus* (Cleveland *et al.* 1934), this species was subsequently reassigned to the genus *Oxymonas* (as *O. doroaxostylus*), primarily based on characteristics of its nuclear division (Cleveland 1950a). However, in most respects the species does not closely resemble *Oxymonas*, and we never saw it presenting anything that could be taken to be a rostellum, one of the distinguishing features of the family Oxymonadidae (Brugerolle and Lee 2000). The SSU rRNA sequence from manually isolated *S. doroaxostylus* (*S. doroaxostylus* I1 in Figure

2.2) strongly supports its original placement in the genus *Saccinobaculus*, as it branched within the *Saccinobaculus* clade and formed a group with 100% support in the more-divergent clade of environmental sequences (*S. doroaxostylus* E11-E14 in Figure 2.2).

The third described species of *Saccinobaculus* is *S. minor*, which was also present in all cockroaches examined (e.g. Figure 2.3F, lower right). This taxon was erected to collect all *Saccinobaculus* cells too small to permit reliable identification, and Cleveland suggested that they represent several distinct species that could not easily be distinguished (Cleveland *et al.* 1934). After additional study, Cleveland subsequently renamed this species *Oxymonas nana* (Cleveland 1950a). Since only about half of the environmental sequences are unambiguously attributable to either *S. ambloaxostylus* or *S. doroaxostylus*, we hypothesise that at least some of the remaining sequences (*Saccinobaculus* sp. E1-E3 and E9-E10) are derived from these smaller oxymonads. The high degree of variability in these sequences, together with the fact that they do not form a single cluster in the phylogeny, lead us to propose that these indeed represent several species. Their inclusion in the clade including *S. ambloaxostylus*, the type species of *Saccinobaculus*, also leads us to propose that their original generic classification was correct, although they are not all one species of *Saccinobaculus*, but rather several.

2.3.3: Relationships between oxymonad subgroups. Our characterisation of molecular data from *Saccinobaculus* allows a molecular phylogeny including all five subgroups of oxymonad to be inferred for the first time. Each of the subgroups represented by more than one sequence was strongly recovered, but the relationships between the groups were not well supported by bootstrap analysis, and were different between analyses using different methods (Figure 2.2). The only consistent relationship that emerged from these trees was the placement of *Monocercomonoides* and *Streblomastix* in a single clade.

We narrowed the taxon selection to short-branch representatives of each group, allowing a more thorough analysis (Figure 2.4), but the support for the inter-group relationships did not improve: the only one with any support was once again that uniting *Monocercomonoides* and *Streblomastix*. We therefore tested all possible relationships

among the six groups (the five oxymonad groups and *Trimastix*) using Approximately Unbiased (AU) tests. We generated a data set where each group was represented by two sequences (except for *Monocercomonoides*, for which only a single sequence is available): this constituted the same dataset as in Figure 2.4, but with *Saccinobaculus* sp. E7 and E14, *Dinenympha* sp. 29565528, and *Trimastix pyriformis* 14348576 removed, and *Streblomastix strix* 28784183 and *Saccinobaculus doroaxostylus* I1 added. Pairs of taxa representing known groups were constrained to branch together to avoid intra-subgroup topologies. We generated all 105 trees possible for this dataset and carried out AU tests on them. 80 trees were rejected at the 95% level, and from a consensus of the remaining 25 trees, no clear pattern emerged. The grouping of *Monocercomonoides* and *Streblomastix* was found in the three most favoured trees, but not in the majority of trees that were not rejected. No other grouping was prevalent amongst the accepted trees, and no particular grouping was clearly over-represented by a majority of the rejected trees. The highly divergent nature of the pyrsonymphid sequences is clearly a potential problem with the tests, so we repeated the analysis while excluding this subgroup. In this test, we found six trees that were not rejected, and three of these (including the two most likely trees) did indeed group *Monocercomonoides* and *Streblomastix*.

2.3.4: Conclusions. The characterisation of the first molecular data from *Saccinobaculus* has brought the taxonomy of several species full-circle. With the exception of *M. globus*, Cleveland originally regarded all of the oxymonads in the hindgut of *Cryptocercus* to be members of the genus *Saccinobaculus* (Cleveland *et al.* 1934). Sixteen years later he revisited the problem and reassigned two groups to *Oxymonas* based on their different life cycles and their behaviour under the influence of the host insect's moulting hormone (Cleveland 1950a). Molecular data support the original classification, since the only oxymonad sequences that we were able to recover from the gut of *Cryptocercus* formed a well-supported group not related to the previously published sequences from *Oxymonas*, and the molecular phylogeny of rRNA from manually isolated *S. doroaxostylus* is specifically and clearly in disagreement with its reassignment to *O. doroaxostylus*. Cleveland also reassigned certain morphotypes of *S. ambloaxostylus* to the

new genus *Notila* on the basis of granules embedded in the axostyle, and of differences in the timing of meiosis (Cleveland 1950c). We are unable to determine where *Notila* fits in our phylogeny, nor whether its recognition would make *Saccinobaculus* paraphyletic. It bears mentioning that the differences between the sexual cycles of *Saccinobaculus*, *Notila*, and *Oxymonas* appear to be attributable to heterochrony (Cleveland 1950c, 1950b, 1950a).

Our data therefore suggest that, while *Saccinobaculus* is restricted in distribution to the gut of *Cryptocercus*, it is both abundant and diverse in that environment. Indeed, our data are most consistent with the original suggestion that most oxymonads in *Cryptocercus* belong to the genus *Saccinobaculus*. No data clearly attributable to *M. globus* was present in our survey, but this is perhaps not surprising since the sequenced *Monocercomonoides* SSU rRNA is 2,900 bp long, about 800 bp longer than *Saccinobaculus*, and might therefore be selected against in environmental PCR.

The molecular tree of oxymonads, while still poorly resolved, is nevertheless beginning to shed some light on the evolution of the some characters in the group. The relationships between oxymonad families have primarily been hypothesized based on characteristics of the axostyle and on putative homologies between attachment organelles, specifically the rostellum unique to the family Oxymonadidae and the microfibrillar holdfast seen in several families. The rostellum is thought to be a highly derived trait that arose after the development of the holdfast (Moriya *et al.* 2003). The holdfast, on the other hand, is reported in the Pyrsonymphidae, *Streblomastix*, and some members of the Oxymonadidae (Bruggerolle and König 1997). Assuming that all oxymonad holdfasts are homologous led Moriya *et al.* to propose a single acquisition, and to place the Polymastigidae alone at the base of the oxymonad tree (Moriya *et al.* 2003), a view also held by Simpson *et al.* (Simpson *et al.* 2002). However, our molecular analysis, as well as that of others (Hampl *et al.* 2005), suggest that the presence of the holdfast is not a reliable character trait; indeed, there is no dispute that it has been lost in the pyrsonymphid *Dinenympha*. Resolving this issue will depend on acquiring additional molecular and morphological data, and in particular on characterising the taxonomic diversity of additional genes.

Another feature of the molecular tree of oxymonads is the resolution of *Monocercomonoides* and *Streblomastix* as a single clade, which has been noted before in SSU rRNA analyses (Hampl *et al.* 2005). This was an unexpected development: previous (morphologically-based) analyses of the relationships between the oxymonad groups tended not to group these two organisms together (Moriya *et al.* 2003). However, after being treated to remove its surface bacteria, the morphology of *Streblomastix* resembles that of *Monocercomonoides* (Leander and Keeling 2004). In addition, Hampl *et al.* pointed out (Hampl *et al.* 2005) that both *Monocercomonoides* and *Streblomastix* have unusually large small-subunit rDNA genes. We can confirm that the *Saccinobaculus* gene, while also large for a eukaryote, is within the range of already-published *Oxymonas*, *Dinenympha*, and *Pyrsonympha* genes, making the increased size of the *Monocercomonoides* and *Streblomastix* SSU a potential synapomorphy for the group. Confirmation of this hypothesis, as well as further support for the grouping of the two families as a clade, will depend on the sequencing of the SSU rRNA genes from additional members of the Polymastigidae, the group to which *Monocercomonoides* belongs. This hypothesis has additional implications for the evolution of the protein-translation machinery of the cell, as discussed in the next chapter.

Figure 2.1. Phylogeny of eukaryotic SSU rRNA using opisthokonts (animals and fungi) as an outgroup. The tree shown is from Bayesian analysis with maximum likelihood branch lengths. Numbers at nodes correspond to percentage of bootstrap support from 100 replicates using ML (above) and distance (below).

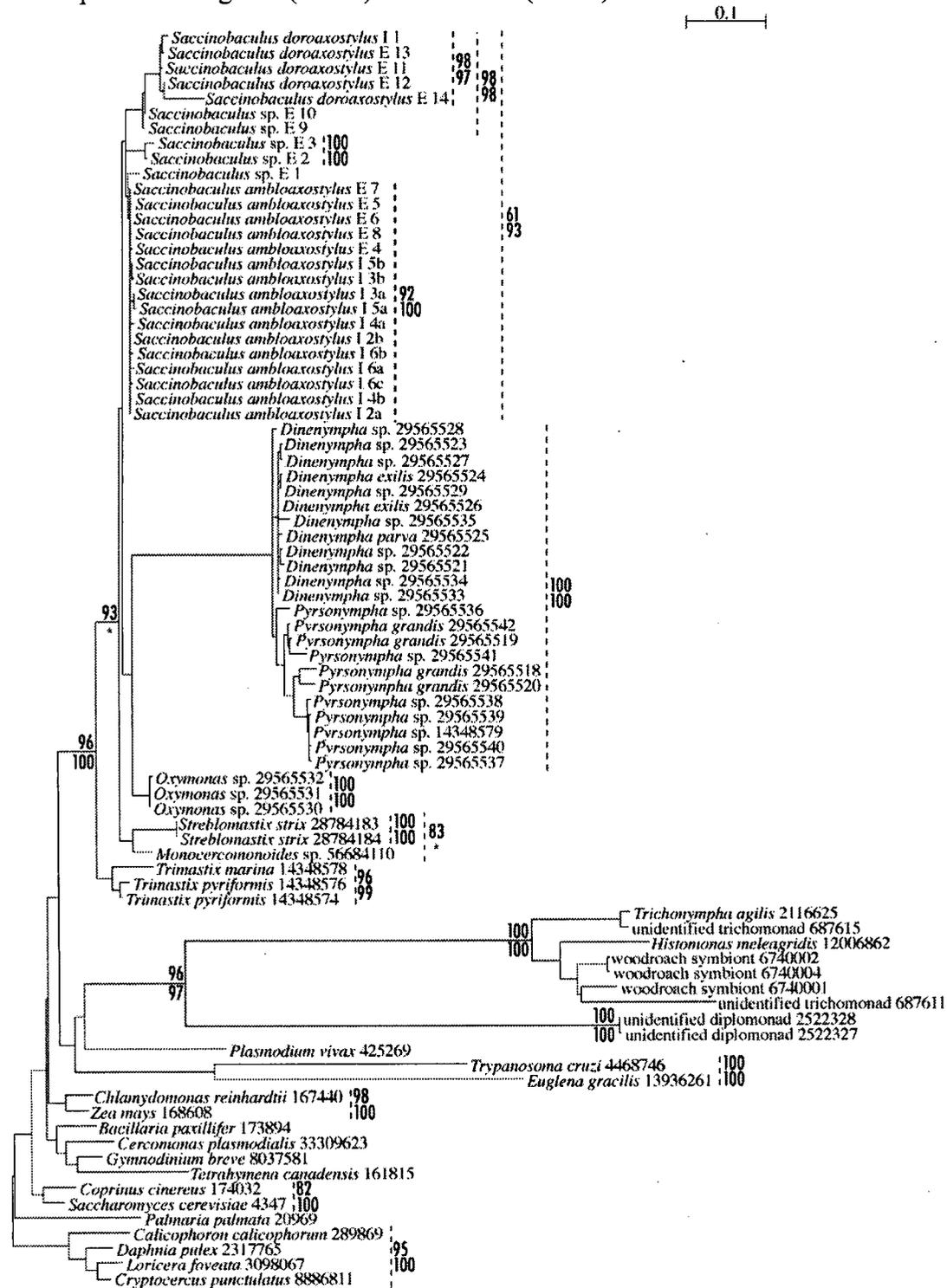


Figure 2.2. Phylogeny of oxymonad SSU rRNA using *Trimastix* as an outgroup. The tree shown is from Bayesian analysis with maximum likelihood branch lengths. Numbers at nodes correspond to percentage of bootstrap support from 1000 replicates using ML (above) and distance (below).

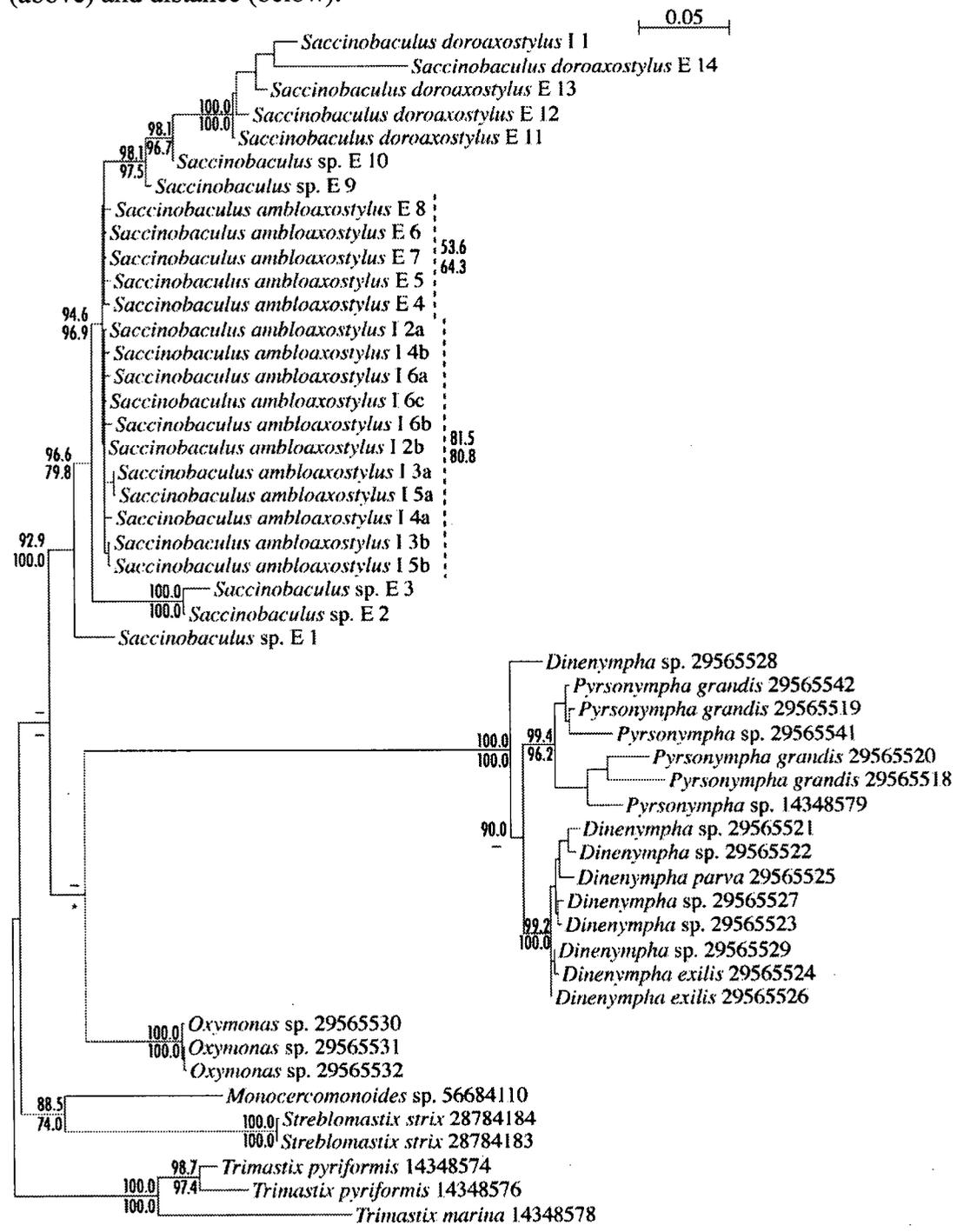


Figure 2.3. Light micrographs showing overall morphology of *S. ambloaxostylus* and *S. doroaxostylus*. These are representative cells that show the characteristics used in the identification of cells for single-cell isolations. (A-D): *S. doroaxostylus*. Two cells are shown (A & B and C & D), both about 160 μm in size. The cytoplasm is granular, with a brown colour and wood particles visible in all individuals. The axostyle (a) is relatively thin and extends through the length of the cell. The nucleus (n) is typically pyriform and associated with the anterior end of the axostyle. In detail of the posterior end, a pointed axostylar projection (ap) can be seen in many individuals. The axostyle is motile, but moves slowly, and the cells maintain a roughly ovoid shape in our observations. (E-J): *S. ambloaxostylus*. Five individual cells are shown, all between 65-100 μm in size. Cells were mostly rounded (e.g., E, G, H) and changed shape rapidly due to axostylar motion. Highly elongated forms (F), in which the axostyle contracted and extended, were also abundant. The nucleus (n) is circular and associated with the anterior end of the axostyle (a) where the flagella (f) emerge. The axostyle is visibly a flexible flattened band (see G & H), which sometimes projects from the posterior end (I). (J): time-lapse sequence of a single cell of *S. ambloaxostylus*, showing the remarkable plasticity of this cell, characteristic of all members of the genus. All scale bars correspond to 50 μm .

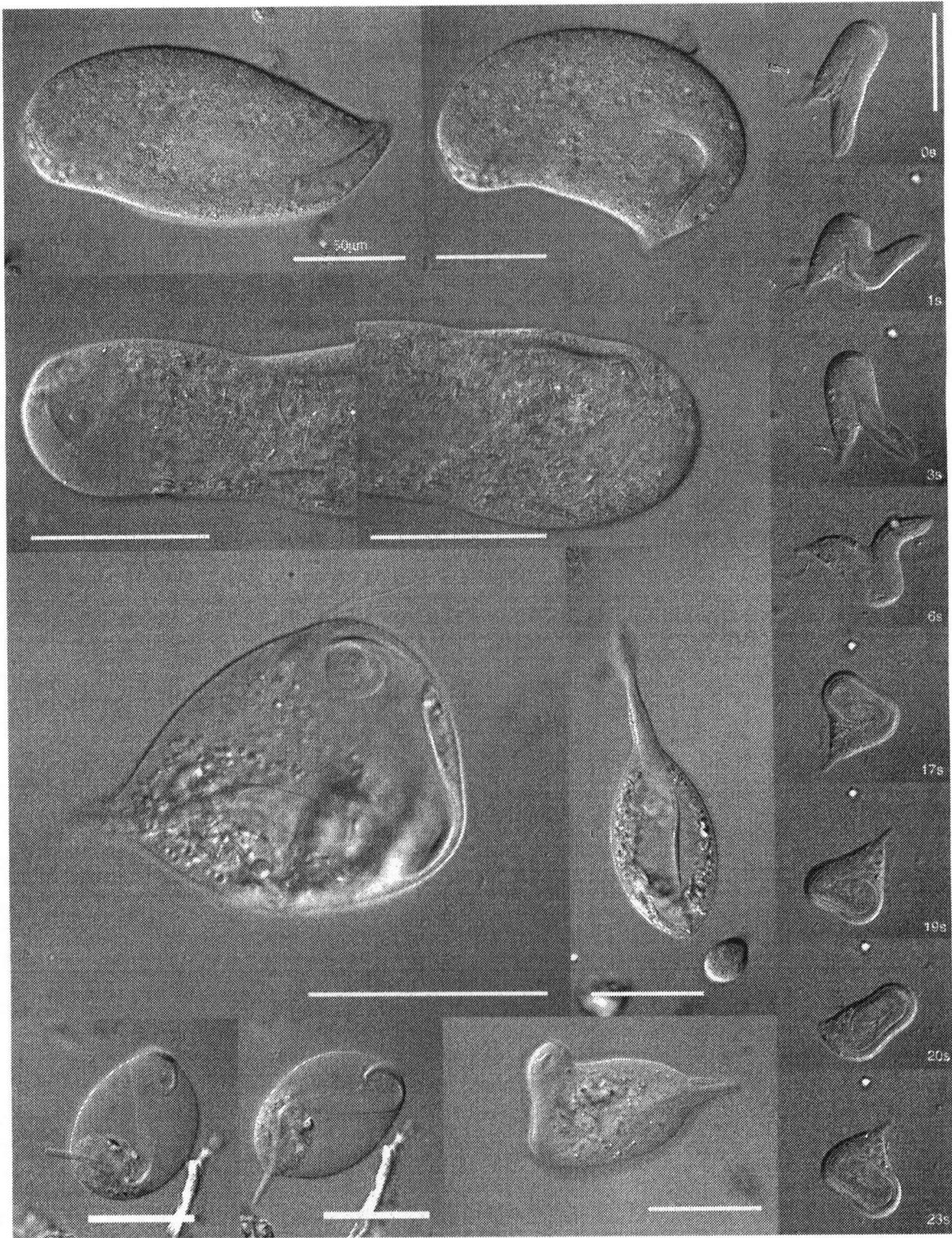
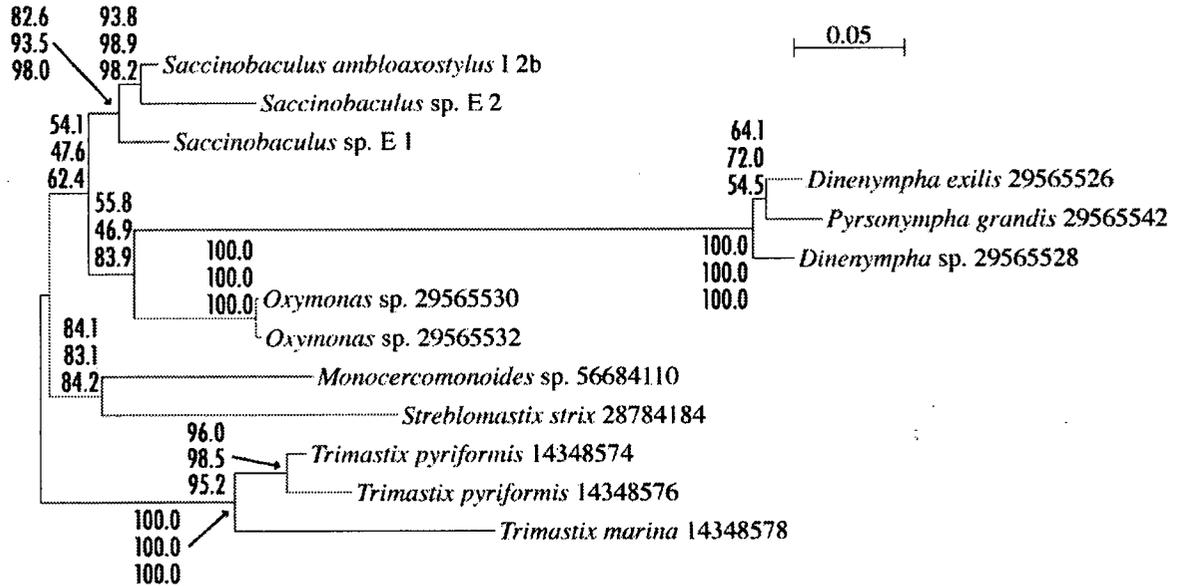


Figure 2.4. Phylogeny of oxymonad SSU rRNA using *Trimastix* as an outgroup. Tree generated as for Figure 2.2. Numbers at nodes correspond to percentage of bootstrap support from ML trees generated with stepwise addition (top), ML trees generated with neighbor-joining starting trees (middle), and distance (bottom).



References

- Brugerolle, G. and König, H. (1997). "Ultrastructure and organization of the cytoskeleton in *Oxymonas*, an intestinal flagellate of termites." *J Eukaryot Microbiol* **44**(4): 305-313.
- Brugerolle, G. and Lee, J. J. (2000). Order Oxymonadida. *Illustrated Guide to the Protozoa*. J. J. Lee, G. F. Leedale and P. Bradbury. Lawrence, KS, Society of Protozoologists: 1186-1195.
- Cleveland, L. R. (1950a). "Hormone-induced sexual cycles of flagellates: II. Gametogenesis, fertilization, and one-division meiosis in *Oxymonas*." *J. Morphol.* **86**: 185-214.
- Cleveland, L. R. (1950b). "Hormone-induced sexual cycles of flagellates: III. Gametogenesis, fertilization, and one-division meiosis in *Saccinobaculus*." *J. Morphol.* **86**: 215-228.
- Cleveland, L. R. (1950c). "Hormone-induced sexual cycles of flagellates: IV. Meiosis after syngamy and before nuclear fusion in *Notila*." *J. Morphol.* **87**: 317-348.
- Cleveland, L. R., Hall, S. R., Sanders, E. P. and Collier, J. (1934). "The wood-feeding roach *Cryptocercus*, its protozoa, and the symbiosis between protozoa and roach." *Mem. Am. Acad. Arts Sci.* **17**(2): 185-342.
- Guindon, S. and Gascuel, O. (2003). "A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood." *Syst Biol* **52**(5): 696-704.
- Hapl, V., Horner, D. S., Dyal, P., Kulda, J., Flegr, J., Foster, P. G. and Embley, T. M. (2005). "Inference of the phylogenetic position of oxymonads based on nine genes: support for metamonada and excavata." *Mol Biol Evol* **22**(12): 2508-2518.
- Huelsenbeck, J. P. and Ronquist, F. (2001). "MRBAYES: Bayesian inference of phylogenetic trees." *Bioinformatics* **17**(8): 754-755.
- Keeling, P. J. and Leander, B. S. (2003). "Characterisation of a non-canonical genetic code in the oxymonad *Streblomastix strix*." *J. Mol. Biol.* **326**: 1337-1349.
- Leander, B. S. and Keeling, P. J. (2004). "Symbiotic innovation in the oxymonad *Streblomastix strix*." *J Eukaryot Microbiol* **51**(3): 291-300.

Moriya, S., Dacks, J. B., Takagi, A., Noda, S., Ohkuma, M., Doolittle, W. F. and Kudo, T. (2003). "Molecular phylogeny of three oxymonad genera: *Pyrrsonympha*, *Dinenympha* and *Oxymonas*." *J Eukaryot Microbiol* **50**(3): 190-197.

Nalepa, C. A., Luykx, P., Deitz, L. L. and Klass, K.-D. (2002). "Distribution of karyotypes of the *Cryptocercus punctulatus* species complex (Dictyoptera: Cryptocerciidae) in the Southern Appalachians: relation to habitat and history." *Ann Entomol Soc Amer* **95**: 276-287.

Schmidt, H. A., Strimmer, K., Vingron, M. and von Haeseler, A. (2002). "TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing." *Bioinformatics* **18**(3): 502-504.

Shimodaira, H. (2002). "An approximately unbiased test of phylogenetic tree selection." *Syst. Biol.* **51**(3): 492-508.

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

Simpson, A. G., Radek, R., Dacks, J. B. and O'Kelly, C. J. (2002). "How oxymonads lost their groove: an ultrastructural comparison of *Monocercomonoides* and excavate taxa." *J. Eukaryot. Microbiol.* **49**(3): 239-248.

Swofford, D. L. (2002). PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Sunderland, Massachusetts, Sinauer Associates.

Trager, W. (1934). "The cultivation of a cellulose-digesting flagellate, *Trichomonas termopsidis*, and of certain other termite protozoa." *Biol. Bull.* **66**: 182-190.

3: PROTEIN-CODING GENES*

3.1: Introduction

The oxymonad *Streblomastix strix* is an organism sufficiently unusual to have been placed in its own family. Like all oxymonads, it is a symbiont living within the guts of animals, and like many gut symbionts it has a remarkably reduced anatomy (Brugerolle and Lee 2000). It is unusual in that it is totally covered in ectosymbiotic bacteria which have been hypothesised to function in a sensory capacity (Dyer and Khalsa 1993) or as a sort of exoskeleton (Leander and Keeling 2004). In addition, it is among the few eukaryotes known to use an alternate genetic code (Keeling and Leander 2003). In this code, two of the standard termination codons (UAA and UAG, collectively referred to as UAR) are instead translated as glutamine (Q: thus, the code is referred to as UAR=Q); this particular alternate code is of special interest because it has arisen independently in several other unrelated eukaryotic lineages (reviewed in (Knight *et al.* 2001)). Investigating the evolution of such things, however, requires a solid understanding of the relationships within the relevant groups, such understanding generally being based upon molecular data.

As discussed in Chapter 2, *Streblomastix* and *Monocercomonoides* (a member of the Polymastigidae) have been hypothesized to be close relatives. The Polymastigidae is perhaps most remarkable in its apparent (and putatively primitive) lack of derived oxymonad features, and in its retention of characters interpretable as ancestral (Simpson *et al.* 2002), making it an obvious candidate for a basal lineage amongst the oxymonads. One implication of this relationship between the highly-derived *Streblomastix* and the putatively-basal *Monocercomonoides* is that the Polymastigidae may be less ancestral than commonly thought: the morphological simplicity of its members could be a secondary, derived state. It is also possible that the Polymastigidae, or indeed the genus *Monocercomonoides*, is not a natural group, being instead an assemblage of superficially similar but unrelated organisms. Such superficial similarity is exaggerated by the small

* This chapter is based on research that is in preparation for publication as de Koning, A. P., Noble, G. P., Wong, J., Heiss, A. A., and Keeling, P. J.: Use of environmental PCR to determine the distribution of a non-canonical genetic code in uncultivable oxymonads.

size of all members of the group, making phenotypic characters difficult to distinguish in living organisms. Indeed, oxymonads can be an awkward group to characterise using only light microscopy, and observations of cell cycle, dependent upon light-microscopy-based techniques, led to the erroneous reassignment of members of one oxymonad group (*Saccinobaculus*) into another (*Oxymonas*), as described in the previous chapter. Similar observations led to a less radical splitting of *Saccinobaculus*, to introduce the new genus *Notila* (Cleveland 1950), as well as the splitting of the genus *Monocercomonoides*, to introduce the new genus *Paranotila* (Cleveland 1966).

The use of the alternate genetic code may be an important character for resolving relationships between *Streblomastix* and other oxymonads, but only if *Streblomastix* is not unique in possessing it. It is therefore of considerable interest that de Koning and collaborators, using an environmental-PCR survey of the gut of *Cryptocercus punctulatus*, found alpha-tubulin genes which use the UAR=Q code and which branch with oxymonads in phylogenetic analyses (de Koning *et al.*, personal communication). The *Cryptocercus* gut biota includes several species of oxymonads, but not *Streblomastix*; the alternate-code-using organism must therefore belong to one of the groups that is represented therein: either the Saccinobaculidae or the Polymastigidae. Given the sister relationship that I have demonstrated between *Monocercomonoides* and *Streblomastix* (a relationship exclusive of *Saccinobaculus*), and given that the adoption of an alternate genetic code is not a character that is prone to lineage sorting, I would hypothesise that the organism using the alternate code in the gut of *Cryptocercus* is most likely a member of the Polymastigidae, namely *Monocercomonoides* or *Paranotila*. Because the identification of the non-canonical-code-using genes was based on environmental PCR, however, the exact identity of the corresponding organism is impossible to determine from these initial data. Not only that, the presence of two groups of oxymonads in the gut of *Cryptocercus* clouds the interpretation of the phylogenies obtained from any environmentally-based assay. Again, interpretation of such phylogenies remains necessarily tentative without identification connecting gene sequences to individual phenotypes.

To determine the identity of the non-canonical-code-using organism in *Cryptocercus*, I sequenced the alpha-tubulin sequence from single-cell isolates of

Saccinobaculus ambloaxostylus, the type species of the family Saccinobaculidae. These nucleotide sequences, all of which translate to the same single amino acid sequence, allow for the unambiguous identification of the lineages represented by the environmental data, and in turn allow for the confirmation of the use of the alternate code by some, but not all, of the members of the Polymastigidae, as well as a more full interpretation of the environmentally-based phylogenies.

3.2: Materials and Methods

A single cockroach was sufficient for single-cell isolation, collected from Black Rock Mountain, Rabun County, Georgia (with a chromosome number of $2n=39$) in September, 2006 by C. A. Nalepa. The cockroach was dissected, and the contents of its gut were suspended in Trager's medium U (Trager 1934). I isolated 47 cells of *Saccinobaculus ambloaxostylus* from this suspension, from which I extracted DNA as previously described (Keeling and Leander 2003). Initially, I amplified DNA using a combined approach, in which primers for four protein-coding genes were combined into a single PCR tube with the total DNA from all 47 cells. The genes to be amplified were alpha-tubulin (for which the primers are 5'-GGGCCCCAGGTCGGCAAYGCNTGYTGG-3' and 5'-GGGCCCCGAGAACTCSCCYTCYTCCAT-3'), beta-tubulin (5'-GCCTGCA-GGNCARTGYGGNAAAYCA-3' and 5'-TCCTCGAGTRAAYTCCATYTCRTCCAT-3'), elongation factor (EF) 1 alpha (5'-CAACATCGTCGTCATCGGNCAYGTTGA-3' and 5'-GCCGCGCACGTTGAANCCNACRTRRTC-3'), and heat shock protein (HSP) 90 (5'-GGAGCCTGATHATHAAYACNTTYTA-3' and 5'-CGCCTTCATDATNCKYTCCATRRTTNGC-3'). The cycling parameters included a 45° annealing temperature, a 90-second extension time, and 40 cycles. The PCR products from this first amplification were reamplified using the same parameters as before except this time with only the alpha-tubulin primers. These products were then reamplified using internal primers (5'-CGCGGCCTCARGTNGGNAAAYGCNTGYTGGGA-3' and 5'-CGCGCCATNCCYTCNCCNACRTACCA-3'), which produced a weak band at about 1100 bp. This was excised and reamplified using a 40° annealing temperature, which produced a strong band at the same weight. This was cloned into competent *E. coli* cells; three of the resulting colonies were sequenced.

The nucleotide and amino acid sequences were inserted into appropriate alignments provided by Audrey de Koning and Geoff Noble. Maximum-likelihood analysis was performed on the protein alignment using PHYML 2.4.4 (Guindon and Gascuel 2003) with the Dayhoff substitution matrix. This analysis used a gamma distribution of variable sites with 8 variable rate categories and invariable sites; the shape parameter alpha was estimated from the data to be 1.305885, and the proportion of invariable sites was estimated from the data to be zero. Bootstrap resampling of 100 replicates was performed using the same procedures and parameters, and the resulting confidence values were mapped onto the corresponding single-run analyses.

3.3: Results and Discussion

3.3.1: Alpha-tubulin phylogeny and identification of the alternate-code-using lineage. All three of the alpha-tubulin sequences amplified from isolated *S. ambyloaxostylus* cells translated to the same amino acid sequence, which was already represented amongst the environmental samples. All sequences obtained from both the environmental study and from my single-cell work outlined here branch together with all published oxymonad sequences with 100% bootstrap support, to the exclusion of all other eukaryotes included in the study (Figure 3.1). Some of the oxymonad groups are well-supported as well: the Strebloplastidae and Pyronomastixidae are recovered with 100% and 80% support, respectively (although the former is to be expected, since all of the genes representing the Strebloplastidae came from a single source). Other groups are not apparent in the phylogeny, as no representatives for the Oxymonadidae have been published, and from the environmental data alone, it is impossible to make a clear distinction between members of the Polymastixidae and those of the Saccinobaculidae. It is reasonable to assume that both of these groups are represented in the environmental data, however, as the size of the alpha-tubulin gene is both more consistent and smaller than that of the SSU rRNA; indeed, there is no substantial variation in the gene size amongst both published data and the environmental data presented here. In any case, though, the phylogeny is not resolved with as much support as is that of SSU rRNA (Chapter 2).

Within the oxymonads, the Streblomastigidae alone appear as basal, a relationship incompatible with the sister relationship to the Polymastigidae which is supported by the SSU data presented in the previous chapter. Not surprisingly, this discrepancy is reflected in the bootstrap analysis, which failed to recover basal placement for this taxon. The basal placement of the Streblomastigidae is congruent with the basal placement of a combined Streblomastigidae/Polymastigidae clade in some of the SSU trees, although such placement was not consistent in all analyses.

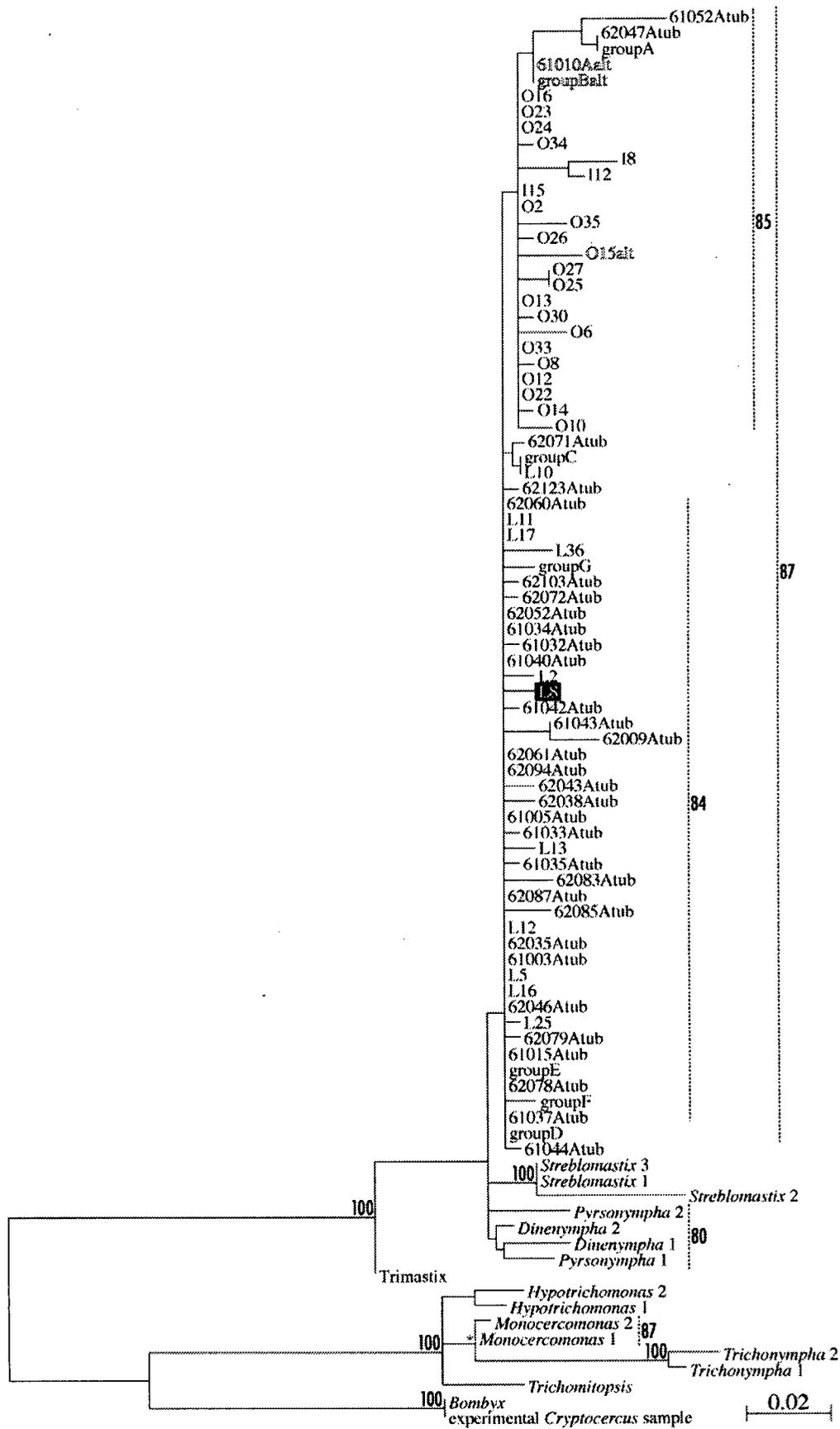
Within the crown taxa, the environmental sequences make up two clades, one which uses only the universal code, and one which occasionally evidences usage of the non-canonical (“UAR=Q”) code. The addition of the single-cell isolates allows us to infer that the clade using the standard code corresponds to the Saccinobaculidae; the other must then represent the Polymastigidae, since these are the only two clades of oxymonads that are present in *Cryptocercus*. From the observable abundance of small oxymonads in extracts of *Cryptocercus* gut, we can expect that both the Saccinobaculidae and the Polymastigidae are in fact represented in this analysis. Again, while support for the integrity of each group is strong, the relationship between the groups is not well supported. The inevitable conclusion is that alpha-tubulin sequences do not provide significant new information about oxymonad phylogeny.

3.3.2: Evolution within the oxymonads. Since the phylogenies presented here are useful primarily only for identification of the lineage using the alternate genetic code, we must refer to the study using the SSU gene in the previous chapter for any analysis of evolution within the oxymonads. Within such a framework, the use of an alternate genetic code is an ideal binary character state, particularly because the extreme unlikelihood of it undergoing lineage sorting makes its acquisition a unidirectional trait. Given these parameters, the environmental *Monocercomonoides* samples using the alternate code must be sister to *Streblomastix*, which must then branch within the Polymastigidae. In other words, we must conclude that the Polymastigidae, and possibly also the genus *Monocercomonoides* as presently understood, are polyphyletic.

This latter finding begs for further investigation. If the genus *Monocercomonoides* is to be maintained as a natural (*i.e.*, monophyletic) taxon, it will be necessary to

reclassify those members using the alternate code. A reliable internal phylogeny of the Polymastigidae *sensu lato* is therefore necessary before further conclusions can be reached. Ideally, a single clade of alternate-code-using oxymonads can be identified, which would then include both *Streblomastix* and the alternate-code-using *Cryptocercus* symbionts. Such a clade would be compatible with an expansion of the family Streblomastigidae; the alternate-code-using organisms would then be reassigned to that family. They would require renaming at the generic level as well, unless they can be identified as belonging solely to the genus *Paranotila*. Cleveland was quite explicit that his identification of *Paranotila* did not encompass all polymastigids within *Cryptocercus* (Cleveland 1966), however, so such identification is not guaranteed: one cannot simply move all polymastigids inhabiting *Cryptocercus* to the genus *Paranotila*. Likewise, the use of the alternate code was not demonstrated in all members of the polymastigid clade inhabiting *Cryptocercus*, although it was not possible to rule such use out, either (de Koning *et al.*, personal communication). It is therefore possible that *Paranotila* uses the standard code, and that some or all of the smaller organisms originally described as *Monocercomonoides* use the alternate code, in which case they would need to be placed in a new genus. Such decisions await further data.

Figure 3.1. Phylogeny of alpha-tubulin from all available oxymonad sequences, representatives of lineages of protist symbionts of *Cryptocercus*, and two insects, including one sample of *Cryptocercus* obtained in the course of amplifying environmental material. The tree is rooted to show a basal dichotomy between oxymonads and all other eukaryotes, and was generated using maximum likelihood. Numbers at nodes show bootstrap values of 100 replicates; values below 80, and those for groups of smaller than four taxa, are not shown. Sequences demonstrating use of the alternate genetic code are highlighted, and the sequence corresponding to the single-cell isolates is shown over a black background.



References

- Brugerolle, G. and Lee, J. J. (2000). Order Oxymonadida. *Illustrated Guide to the Protozoa*. J. J. Lee, G. F. Leedale and P. Bradbury. Lawrence, KS, Society of Protozoologists: 1186-1195.
- Cleveland, L. R. (1950). "Hormone-induced sexual cycles of flagellates: IV. Meiosis after syngamy and before nuclear fusion in *Notila*." *J. Morphol.* **87**: 317-348.
- Cleveland, L. R. (1966). "Nuclear division without cytokinesis followed by fusion of pronuclei in *Paranotila lata* gen. et sp. nov." *J Protozool* **13**(1): 132-136.
- Dyer, B. D. and Khalsa, O. (1993). "Surface bacteria of *Streblomastix strix* are sensory symbionts." *Biosystems* **31**(2-3): 169-180.
- Guindon, S. and Gascuel, O. (2003). "A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood." *Syst Biol* **52**(5): 696-704.
- Keeling, P. J. and Leander, B. S. (2003). "Characterisation of a non-canonical genetic code in the oxymonad *Streblomastix strix*." *J. Mol. Biol.* **326**: 1337-1349.
- Knight, R. D., Freeland, S. J. and Landweber, L. F. (2001). "Rewiring the keyboard: evolvability of the genetic code." *Nat Rev Genet* **2**(1): 49-58.
- Leander, B. S. and Keeling, P. J. (2004). "Symbiotic innovation in the oxymonad *Streblomastix strix*." *J Eukaryot Microbiol* **51**(3): 291-300.
- Simpson, A. G., Radek, R., Dacks, J. B. and O'Kelly, C. J. (2002). "How oxymonads lost their groove: an ultrastructural comparison of *Monocercomonoides* and excavate taxa." *J. Eukaryot. Microbiol.* **49**(3): 239-248.
- Trager, W. (1934). "The cultivation of a cellulose-digesting flagellate, *Trichomonas termopsidis*, and of certain other termite protozoa." *Biol. Bull.* **66**: 182-190.

4: OVERALL CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

4.1: Conclusions

Until the studies outlined in this thesis were completed, the relationships among all five groups of the oxymonads could not be analysed using molecular methods because of a lack of data from *Saccinobaculus*. Such analysis is limited in its results: only one relationship is supported unambiguously, that between the Streblomastigidae and the Polymastigidae, a relationship that contradicts traditional, morphologically-based expectations. Specifically, the holdfast presumed to have arisen only once in oxymonad evolution has instead arisen at least twice, once in *Streblomastix* and at least once in the remainder of the oxymonads; alternatively, it has been lost several times. Likewise, the contractile axostyle appears to be a poorly conserved character state, since it has been reported as being rigid in the Polymastigidae but somewhat flexible in the Streblomastigidae (Brugerolle and Lee 2000). The exceptional motility of the axostyle seen in the Saccinobaculidae is therefore not necessarily homologous with this flexibility.

The exact branching order of all five families of oxymonads is at this point a matter only of speculation. It would be surprising to find further data supporting a basal position of the Oxymonadidae, as this group contains the most morphologically derived oxymonads. Beyond the presence of the rostellum (most likely a development of the preaxostylar lamina characteristic of not only the oxymonads but of *Trimastix* as well), this group contains multinucleate organisms that appear to converge on the state achieved by the hypermastigotes, a state that has culminated in a similarly multiflagellated arrangement. However, such a state could easily be explained by a delayed onset of cytokinesis in the organisms' evolutionary history. This situation is not without parallel in the plasticity of the sexual cycles of the symbionts of *Cryptocercus*, although to a lesser extent (Cleveland 1950c, 1950b, 1950a). Therefore, such surprise would not be impossible to reconcile with our present understanding of oxymonad diversity.

The genetic code is a far better character state to use for phylogenetic analysis. It is central to all cellular information processes, and intermediate states in its development are highly unlikely to persist for long enough for the development of multiple lineages. In other words, with regard to evolutionary processes, the adoption of an alternate code is

likely to occur quickly and irreversibly. The discovery of such a code in the Streblomastigidae and the Polymastigidae is therefore entirely consistent with their being related in SSU phylogenies. Unfortunately, the discreteness of the character does not assist in determining further relations.

4.2: Future Directions

These studies answer some questions, but pose many more. The complete evolutionary history of the oxymonads is very far from completely told. The flexibility of these organisms' sexual cycles, combined with the convergence of some members of the Oxymonadidae to a state resembling some members of the hypermastigotes, suggests that comparisons between the development of these groups may yield insights into the evolution of other highly polyflagellated organisms, such as the opalinids and the ciliates. Likewise, an organelle similar enough to the axostyle of the oxymonads to bear the same name is known in the parabasalids, a group which coincidentally contains the hypermastigotes. This organelle is almost certainly nonhomologous. However, the relationships of each group to other eukaryotes are not clear, and it cannot now be concluded that the parabasalid and oxymonad axostyles did not arise from a common primordium – whether cytoskeletal or genetic – in some distant ancestor.

Clarification of the relationship amongst the oxymonads, beyond the one connection made between the Streblomastigidae and the Polymastigidae, can be achieved through analysis of more genes. Outside of the Saccinobaculidae, elongation factor 1-alpha has been sampled from all oxymonad families except the Oxymonadidae. Beta-tubulin and heat shock protein 90 have been sequenced from members of the Streblomastigidae and the Polymastigidae, as well as from *Trimastix*. Obtaining sequences for these genes from the remaining groups would enable more detailed phylogenies to be inferred, and that would in turn allow for more detailed and less speculative analyses of oxymonad evolutionary history.

One interesting phenomenon in phylogeny is the fact that the addition of taxa can alter the topology of the phylogeny. This is true not only in cases where the added taxa branch within an existing group but even in cases where no relationship can clearly be made between newly and previously sampled taxa (Page and Holmes 1998). Given this

situation, it is clear that more organisms must be added to our existing phylogenetic analyses, both to confirm the relationship between the Streblomastigidae and the Polymastigidae, and to stabilise the phylogenetic topology with respect to other groups. In particular, the Oxymonadidae are represented by only a small number of sequences from just the type genus; given the morphological diversity of this group, further sampling of it is a high priority. Indeed, it is possible that it may not be monophyletic: assertions of monophyly are predominantly based on the presence of the rostellum, a character that Cleveland believed to be observable in some *Saccinobaculus doroaxostylus* cells (Cleveland 1950a). Since the presence of the microfibrillar holdfast appears to be a comparatively evolutionarily labile character, caution should be encouraged when asserting the primacy of the presence of the rostellum as a taxonomically significant character, especially as it appears to be a development of the preaxostyle present in all oxymonads (Brugerolle and König 1997).

Likewise, the composition of the Polymastigidae has been called into question by this study. Surveys of polymastigid diversity are essential to clarifying this issue. Furthermore, such surveys can help to clarify our understanding of the development of the alternate code within the oxymonads, and may shed some light upon the acquisition of the holdfast as well as of the evolution of axostylar motility. Of course, such surveys should also provide further insight into the relations amongst the oxymonads in general.

The oxymonads have proven to be as difficult to characterise with molecular methods as they are fascinating in their morphology, behaviour, and natural history. There is without question a great deal still to be done in improving our understanding of this group, but we are finally beginning to make progress. It is my hope that such progress will continue.

References

Brugerolle, G. and König, H. (1997). "Ultrastructure and organization of the cytoskeleton in *Oxymonas*, an intestinal flagellate of termites." *J Eukaryot Microbiol* 44(4): 305-313.

Brugerolle, G. and Lee, J. J. (2000). Order Oxymonadida. *Illustrated Guide to the Protozoa*. J. J. Lee, G. F. Leedale and P. Bradbury. Lawrence, KS, Society of Protozoologists: 1186-1195.

Cleveland, L. R. (1950a). "Hormone-induced sexual cycles of flagellates: II. Gametogenesis, fertilization, and one-division meiosis in *Oxymonas*." *J. Morphol.* **86**: 185-214.

Cleveland, L. R. (1950b). "Hormone-induced sexual cycles of flagellates: III. Gametogenesis, fertilization, and one-division meiosis in *Saccinobaculus*." *J. Morphol.* **86**: 215-228.

Cleveland, L. R. (1950c). "Hormone-induced sexual cycles of flagellates: IV. Meiosis after syngamy and before nuclear fusion in *Notila*." *J. Morphol.* **87**: 317-348.

Page, R. D. M. and Holmes, E. C. (1998). *Molecular Evolution: A Phylogenetic Approach*. Oxford, Oxford University Press.