The systematics, zoogeography and evolution of Dolly Varden and bull trout in British Columbia.

By

Gordon Robert Haas

B.Sc. (hons.), The University of British Columbia, 1984

A thesis submitted in partial fulfillment of

the requirements for the degree of

MASTER OF SCIENCE

in

The faculty of graduate studies

(Department of Zoology)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September 1988

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Zoology

The University of British Columbia
Vancouver, Canada

Date October 12, 1988

Abstract

An analysis of the systematics, zoogeography and evolution of the Dolly Varden char species complex in British Columbia is presented. These features of this species complex and the morphometric statistical procedures used in these analyses have both long been the subjects of strong debate and also have recently seen much renewed interest and work. This thesis assesses both these areas and is divided into those two parts. The first section deals with these three biological topics, and the second section contains a synthesis and exploratory data assessment of the commonly used morphometric techniques and provides some new methodology for understanding their requirements and interpreting their results.

PART I

1. The systematics of the Dolly Varden char species complex is examined by using principal component analysis (PCA) to designate typological species groupings and then employing linear discriminant function analysis on a reduced set of significant characters to classify the remaining specimens. This typological distinction is verified with distributional information that reveals no interbreeding of the species in areas of parapatry and sympatry, and with preliminary information regarding intra- and inter- specific crosses, spawning colouration, skull osteology, cytology and embryology. This data is also suggestive of competitive exclusion and character displacement. All these results indicate that the Dolly Varden char species complex in B.C. is composed of two species, Dolly Varden (*Salvelinus malma*) and bull trout (*Salvelinus confluentus*).

2. The zoogeography of these two species is analyzed using canonical trend surface analysis (CTS). CTS can potentially separate confounding non-geographic morphometric information from the data and thus could allow historical zoogeographic patterns to be inferred from that data which corresponds to geography. Such a reconstruction reveals the possible glacial refuge origins and post-glacial recolonization patterns of these two species for each of the major river drainages in B.C..

3. The evolution of these two species is assessed through the implementation of PCA to fit the cross-sectional morphometric data to an ontogenetic model. The resultant PCA size and shape vectors effectively portray allometric trends which indicate that Dolly Varden could have evolved from bull trout through neotenic paedomorphosis. This result is supported with data on growth rates and developmental homeostasis.

PART II

4. A synthesis of the available but widely scattered and disparate information on the data and statistical requirements for morphometric statistics reveals the analytical problems that can result from not approximating underlying test assumptions. These assumptions are important, but are not appreciated or often assessed. Simple recommendations and rarely used tests for dealing with these requirements are provided.

5. The effectiveness and compatability of four bivariate morphometric techniques (ratios, $\log_{10}$ ratios, allometric regression, regression residuals) are assessed. All methods provide similar but ineffective individual ordination and group separation. Their effects on characters differ greatly and are often unrealistic. None of these methods effectively removes all the confounding allometric size information, but allometric regression will usually be the best bivariate procedure.

6. A similar assessment of four multivariate morphometric procedures (covariance matrix PCA, correlation matrix PCA, shear matrix PCA, size-constrained matrix PCA) is undertaken. Size-constrained PCA results in non-orthogonal vectors that also do not represent the traditional

multivariate morphometric size and shape vectors. As well, the character and individual information it provides is unrealistic. The other three techniques result in similar and effective individual ordination, group separation and removal of confounding allometric size information. PCA on a covariance matrix is likely the best multivariate method since it provides the most realistic size adjustment and character information.

7. PCA is often carried out on data which has been previously adjusted through bivariate procedures. An examination of this method demonstrates that it results in no benefits since the multivariate morphometric size and shape vectors are lost, and the data variation is no longer synthesized into only two or three resultant significant vectors.

8. PCA is also performed on mixed character data sets (continuous and discontinuous data). An assessment of this procedure shows that it provides improved group separation, but the representation of characters, individuals and multivariate morphometric size and shape relationships is confounded and unrealistic. There also is a slight reduction in data synthesis.

9. A methodology for back-transforming PCA output into the original and more intuitively comprehensible data scale, format and dimensions is given. This back-transformation also verifies the traditional belief that the first resultant PCA morphometric vector is size and that the second is shape. Separate unconfounded matrices for size and shape information in which only the significant data variation is accounted for can thus be independently back-transformed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgements

I would especially like to thank Dr. J.D. McPhail for providing me with the support and freedom to do this thesis, and Dr. Dolph Schluter for kindly taking over my supervision during his sabbatical leave. This same measure of support, tolerance and much more was also always available to me at home and thus at least equivalent gratitude and credit is due Molly Nevin. As well, my family offered me their support in spite of the absence of any post-secondary education tradition and not always really understanding why anyone would want to do this anyway.

Many people helped me in making the extensive samples of these difficult to collect fish, but Mike St. John and Tommy Suzuki deserve special mention for their considerable efforts. Other nonetheless important field assistants were Gary Birch, Peter Haas, Debbie McLennan, Molly Nevin, Rob Sampson and Arlene Tompkins. Fish were also collected and gratefully given to me by Mike Galesloot, Gordon Hartman, Greg Hoyer and Don Webb. Gary Birch also helped me greatly with my laboratory crosses, and Debbie McLennan always fed and tended the fish when I was out of town.

The unfortunate but appreciated first readers of my thesis were Gayle Brown, Chris Foote, Don Hall, Debbie McLennan and Arlene Tompkins. I thank them for their often difficult to give (and take) advice, their continued friendship and belief in me, and also for saving me thousands of dollars in counselling fees. While they did not read my thesis, I must similarly thank Brad Anholt and John Eadie for their mutual thesis suffering and support. Alistair Blachford also helped me with some of my computer programs especially those pertaining to the digitizing of fish shapes. Drs. Don Ludwig and Jack Maze similarly provided me with statistical inspirations when my own ideas had been exhausted. As for my other unmentioned friends, colleagues and professors, you know who you are and that I thank all of you.

GENERAL INTRODUCTION

The salmonid genus *Salvelinus* has long been recognized as a difficult taxonomic group (Berg 1948, Jordan et al. 1930, Vladykov 1954). In the north Pacific region, this difficulty is exemplified by the Dolly Varden char (*Salvelinus malma*) species complex (DeLacy and Morton 1943, McPhail 1961). This complex is composed of an unknown number of species whose validity is hotly debated (Behnke 1980, 1984, Cavender 1978, 1980, Frohne 1973, McCart 1980, Morrow, 1973, 1980a–b, Morton 1970). This long and still unresolved lack of agreement stems largely from difficulties in morphometric analyses of the various morphotypes (McPhail 1961, Mednikov et al. 1980, Savvaitova 1980a–b, Vladykov 1964) that comprise this complex. In an attempt to meet this shortcoming, the Dolly Varden complex in British Columbia is reanalyzed using new morphometric data and techniques. This reanalysis, however, proved to be a problem itself rather than an immediate solution to the taxonomic difficulties in this species complex.

A problem with most morphometric procedures is that no guidelines or consistent recommendations exist for their application or to test their utility. In addition, many of the procedures are poorly understood and recently there has been a strong revival of interest in morphometrics and many new, complex methods are now available. Unfortunately, information on these procedures is scattered throughout the literature, and even worse much of it is contradictory, overwhelming, unsubstantiated and unorganized. Furthermore, the effectiveness and compatability of different morphometric analyses has rarely been examined (for partial exceptions see Leamy and Bradley 1982, Reist 1985, Rohlf and Bookstein 1987, Shea 1985) and studies using different approaches frequently are deemed comparable even though their comparability is unknown. Modifications of traditional morphometric procedures also can be dangerous if their applicable conditions and statistical properties are poorly represented (Corruccini 1978, Sjøvold 1975). Too often, these modified procedures are accepted without question because they are soon assumed to be standard methods (Reyment et al. 1984).

In addition, the necessity of and the insights gained through the more complex morphometric procedures are often questioned (Corruccini 1975, 1978, Reist 1985). Indeed, if these techniques are better than other methods this worth must be demonstrated. Otherwise, the effort required to

learn and use these more difficult methods will not seem worthwhile. As well, the impetus, time and effort involved in learning some of these procedures make them prohibitive to many people, yet computer programs for running them are readily available (Blackith and Reyment 1971, Corruccini 1975, Edwards 1971, Neeley 1972, Rao 1972, Reyment et al. 1984, Yates 1966, Yates and Healey 1964). The computer programs, however, do not provide directions for the analysis and can be easily misapplied. Consequently, these complex techniques are either shunned or sometimes misused.

This thesis is divided into two parts. The second part reviews and compares bivariate and multivariate morphometric procedures, and introduces some new methodology for understanding their requirements and interpreting their results. The examinations are based on a single data set obtained from the Dolly Varden complex in British Columbia. This data set allows for a complete analysis of individual specimens, their hypothesized group relationships, their characters and the allometry coefficients of these characters. A general set of guidelines for morphometric analyses are also presented in **part II**, and they attempt to synthesize the available information on morphometric techniques into a single, comprehensive format. The first part uses the most appropriate morphometric procedures to interpret the systematics, zoogeography and evolution of the Dolly Varden complex in British Columbia. The questions addressed are whether this complex is composed of one or more species, what their zoogeographic patterns are and how did the species evolve. New multivariate morphometric procedures are introduced in **part I** to attempt to quantitatively establish large-scale biogeographic patterns and to provide possibilities concerning the potential evolutionary steps that gave rise to the existing species.

PART I — Biology

CHAPTER ONE

Systematics of the Dolly Varden Char Species Complex in British Columbia

Introduction

The Dolly Varden char (McPhail 1986, Morton 1955, 1980, Nyman 1984) (*Salvelinus malma*) inhabits most north Pacific drainages on both the Asian and North American coasts. It has been recognized as a distinct species, at least in North America (Armstrong and Morrow 1980, Behnke 1980, 1984, Cavender 1978, 1980, Frohne 1973, Johnson 1980, Morrow 1980a–b, Morton 1970, Nyman 1972, Ouellette and Qadri 1968, Rounsefell 1962, Scott and Crossman 1973, Vladykov 1964) and Japan (Behnke 1980, 1984, Behnke and Shimizu 1962, Cavender 1978, 1980, Ishigaki 1969, Maekawa 1977, Nakamura 1963, Oshima 1961), since its formal separation as a species distinct from the Arctic char (*Salvelinus alpinus*) species complex (DeLacy and Morton 1943, McPhail 1961). Recently most Soviet ichthyologists have also recognized it as a distinct species (Chereshnev 1982, Glubokovsky and Chereshnev 1981; but see Savvaitova 1980a–b, 1983). This specific separation resulted in Dolly Varden inheriting a major portion of the notorious variability that is part of their aspect of the Arctic char complex and consequently receiving considerable taxonomic attention as a distinct char species complex (Behnke 1972, 1980, 1984, Frohne 1973, McCart 1980, Morrow 1973, 1980a–b, Morton 1970) with its own similar difficult taxonomic problems.

A recent analysis of this complex by Cavender (1978) suggests that North American Dolly Varden should be divided into two distinct species, Dolly Varden (*S. malma*) and bull trout (*Salvelinus confluentus*). While his study is persuasive, his results have been inconsistently recognized and applied. This is largely the consequence of shortcomings which unfortunately are often characteristic of char systematic studies (McPhail 1961, Mednikov et al. 1980, Savvaitova 1980a–b, Vladykov 1964). The work is based solely on museum specimens, uses a typological species concept and analyzes only single characters and not in any statistical manner. Furthermore, Cavender (1978) did not provide a diagnosis that would consistently identify bull trout.

Individual characters cannot often completely and properly separate species within char complexes (Behnke 1980, 1984, Cavender 1978, McPhail 1961, Mednikov et al. 1980) and thus their

morphometric analysis requires a multivariate statistical approach (Corruccini 1975, 1978, Lubischew 1962; eg. Dempson 1984, Frohne 1973, Henricson and Nyman 1976, McCart 1980, McPhail 1961, Morrow 1980, Ouellette and Qadri 1968, Viktorovsky and Glubokovsky 1977). Such char populations can be characterized by a multivariate combination of variables that will distinguish them from other populations. However, the taxonomic level of such statistical distinctions is unknown unless additional distributional and ecological information is available. Museum work is indispensable in char systematics because of the difficulty in obtaining enough specimens, but it is also often insufficient to decipher any ecological variability (Hammar 1984, McPhail 1961, Savvaitova 1980b). The museum specimen work should be supplemented by field studies which enable the variability to be sorted out.

My study consists of several multivariate analyses of the Dolly Varden complex in British Columbia (B.C.). I interpet these statistical analyses through the use of the biological species concept (Mayr 1963, 1969), the necessity of which has been emphasized in char taxonomy (Chereshnev 1982, McPhail 1961, Savvaitova 1980a–b, 1983). While the biological species concept is not universal (Hull 1970, Wiley 1978), always testable (Ehrlich 1961, Key 1981, Sokal 1970, Sokal and Crovello 1974) or completely ascertainable (Holsinger 1984, Hull 1978), it is generally applicable in sexually reproducing sympatric populations (Paterson and McNamara 1984). Since char are sexually reproducing and I have located regions of sympatry for Dolly Varden and bull trout, I use the biological species concept. It also does not preclude other species concepts (Cracraft 1983, 1987, McKitrick and Zink 1988, Paterson 1985, Scudder 1974, Wiley 1978) but is emphasized here because of its operational nature.

Materials and Methods

Study Approach

Potential regions of allopatry for both Dolly Varden and bull trout in B.C. were identified from Cavender's (1978) study and are verified as allopatric by this study. These areas are the Queen Charlotte Islands for Dolly Varden and the extreme south-eastern area of B.C. (Kootenays) for bull trout. Thirty museum specimens were selected from each of these specific regions and one hundred

eighteen measurements, twenty counts and twenty-six truss measurements (fig. 1) (Bookstein et al. 1985, Strauss and Bookstein 1982) were made on each char.

No single character is found that completely defines the char from these areas, and this is consistent with Cavender (1978). Thus, their typological validity based on these specimens' morphology is multivariately verified using principal component analysis (PCA) and then a linear discriminant function (LDF) is derived that completely separates them. The four LDF characters and the twenty-six truss measurements are then measured on our remaining museum char collection. The distribution of each species in B.C. is established from this, and possible parapatric and sympatric areas are identified. These areas are then sampled to determine if the two nominal char species are merely ecophenotypes or if there is evidence of introgression. The same four LDF and twenty-six truss characters are measured on all these char in areas of sympatry as well. Inter- and intra-specific laboratory crosses and some electrophoresis is also undertaken.

Morphometrics and Meristics

Body measurements are made with Helios vernier calipers accurate to 0.10 mm. Where necessary, measurements and counts are made under a binocular dissecting microscope. All bilateral measurements and counts are made on both sides of the body, but by convention (Hubbs and Lagler 1958) only those on the left side are used in the analyses. Accuracy is further verified by repeating all measurements and counts until the same number is obtained twice. My measurement error is statistically insignificant (see **chp. 4**).

The twenty-six truss measurements (fig. 1) are computed from digitized landmark points taken from projected slides. A slide photograph was taken of every char specimen alongside a scale ruler. The slides are projected onto paper taped to a wall. The twelve landmark points and the scale reference are marked onto the paper. These sheets are digitized on the digitizing tablet available at the Biological Data Centre (B.D.C.), University of British Columbia (U.B.C.). The resulting Cartesian coordinates are then converted with an AWK program (Aho et al. 1988) into the twenty-six truss measurements. Both the digitizing (in BASIC) and AWK computer programs

are available from me, and are adaptable to any truss analysis. The sixty original allopatric char are used to verify my computer programs since their truss measurements are also taken with calipers.

The measurements and counts are presented and defined, with their methodology explained, in **appendix A**. All linear measurements, conventional (Hubbs and Lagler 1958) or devised by me, are straight line distances.

U.B.C. Ichthyological Museum Char Collection

All museum specimens of char used are in the U.B.C. Ichthyological Museum, and my study of them is in two parts. In the first section, thirty char of each species from their aforementioned allopatric areas are chosen for the initial delineation. Previous taxonomic work suggests this number is sufficient for such analyses (Neff and Marcus 1980, Mardia 1971, Reist 1985) and my data indicate 30 specimens give statistically reliable results (see **chp. 4**). Char from each allopatric region are selected to represent both sexes, similar broad size ranges (8.5–49.3 cm) and individuals from all typical habitats (lakes, rivers and streams). The thirty Dolly Varden come from seven localities and the thirty bull trout from ten. The second part of the museum study which only uses the LDF characters and the truss measurements involves three hundred thirty char from ninety-three B.C. sites.

Further Char Collections

I made detailed collections in three of the four areas I had identified as possible regions of sympatry. These are in the Skeena and Nass River drainages, and especially in the Lower Fraser River watershed. Where possible, I also collected in B.C. regions where there are no U.B.C. museum specimens. All fish collected are now deposited in the U.B.C. Ichthyological Museum.

Collections are made by gill-netting, electroshocking, seining, trapping, and angling. All the fish are preserved in 10 % formalin for about 1 month, and later placed in 37 % isopropyl alcohol. Tissue samples of eye, heart, liver and caudal muscle are removed from each fresh specimen before it was preserved and are individually tagged and immediately frozen in liquid nitrogen at the field site.

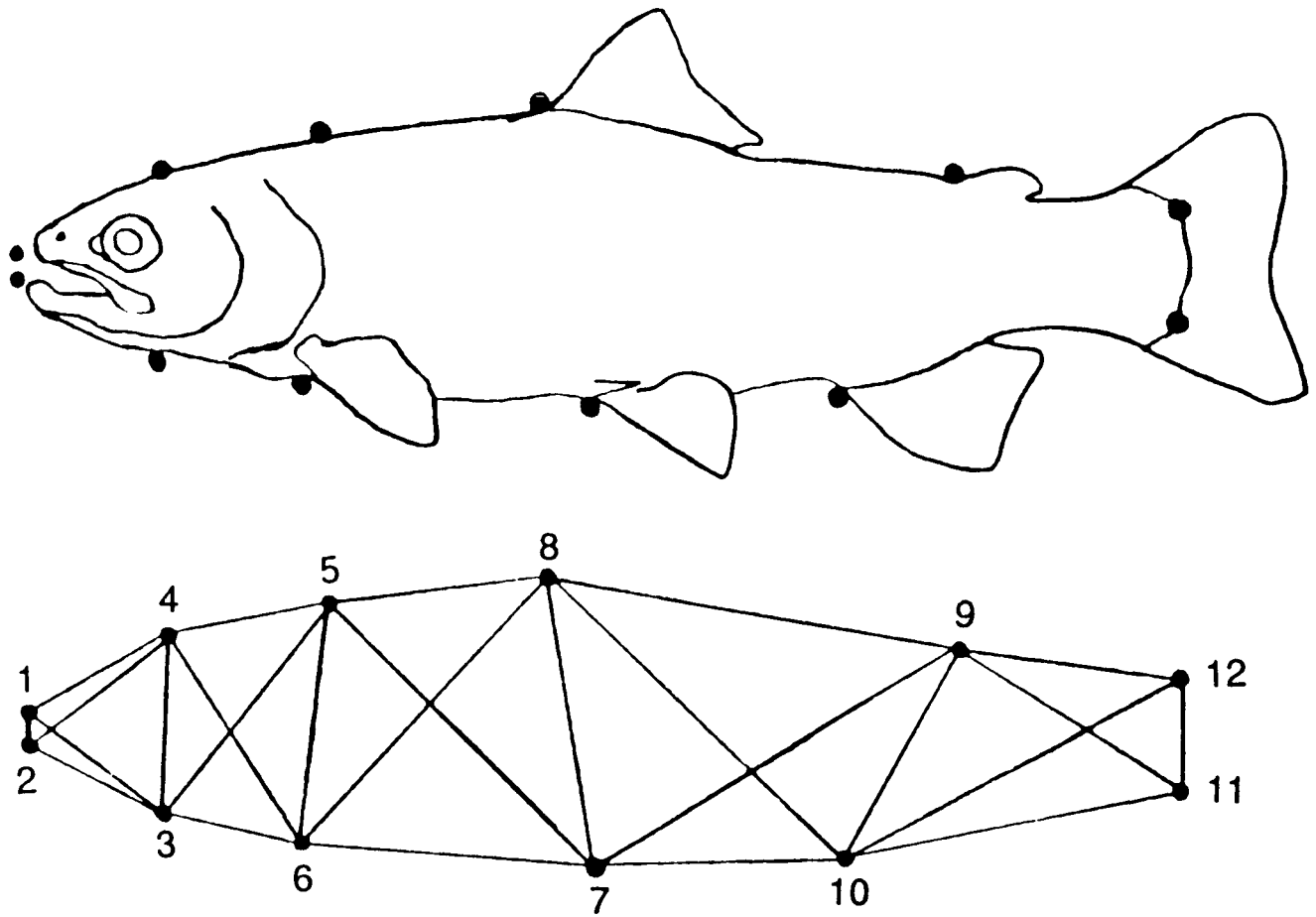


FIGURE 1. Truss series diagram (12 landmark points; 26 measurements).

Collections are made exclusively in B.C. because this is the only geographic region in the range of the Dolly Varden complex where there is almost universal agreement that the char are indeed Dolly Varden (Armstrong and Morrow 1980, Behnke 1984, McPhail pers. comm., Morrow 1980). Most of the U.B.C. Ichthyological Museum specimens are also only from B.C. as much of the collection of McPhail (1961) and other regional char collections had previously been sent to the Canadian National Museum in Ottawa.

Electrophoresis

The small amount of electrophoresis undertaken follows the procedure outlined in Clayton and Ihssen (1980). The enzymes analyzed are isocitrate dehydrogenase (IDH), lactate dehydrogenase (LDH), malate dehydrogenase (MDH), phosphoglucoisomerase (PGI) and phosphoglucomutase (PGM) all from caudal muscle tissue.

Inter – and Intra – Specific Crosses

Inter- and intra-specific crosses were made between Dolly Varden and bull trout collected from separate, but nearby, drainages in the lower Fraser River system . The Dolly Varden are from Katherine Lake ($49°25'N$, $122°35'W$) and the bull trout from Foley Lake ($49°8'N$, $121°30'W$). The parental and hybrid char all developed normally and were reared under identical conditions in the laboratory (see Frost 1965, Nordeng 1983, Savvaitova 1980a, Skreslet 1973). A chlorine pulse in the water system killed the crosses before they reached maturity but all individuals were kept and frozen. An analysis of a subsample of these char was undertaken and more are in progress. No electrophoresis was done on these samples, but the same twenty-six truss measurements and four LDF characters are measured.

Data Analysis

All analyses and graphics are based on computer macros I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the B.D.C.. They are available from me. See **part II** of this thesis for a more detailed discussion.

All morphometric characters were adjusted for body size both by division with and regression against standard length, but neither approach produced any single morphometric variable that

would define the species. Since the meristic characters are not size-confounded, they are directly compared between the species. Again, no single meristic character completely separates the species. For a more detailed discussion of these bivariate procedures and a discussion of the relationship of size in the meristic variables see **chapters 5–6.**

It is necessary to reduce my character set for multivariate analyses. Average-linkage cluster analysis on covariance and correlation (Best 1978, Joliffe 1972, 1973, Power 1971, Thorpe 1976) matrices are used to identify character groupings. Characters are then chosen from each of these clusters on the basis of previous knowledge (Cavender 1978), repeatability, reliability, non-redundancy, usefulness and interest (see **chp. 4**). In this way the character set is reduced to the fifty-one morphometric and ten meristic characters used in the **part II** analysis. Scatter plots of the first two principal components of an R-mode PCA of covariance and correlation matrices of $\log_{10}$ transformed data (Gould et al. 1974, Johnston 1973, Thomas 1968, Thorpe 1976) are used to verify this character selection. The character groupings obtained on these scatter plots confirm those of the dendograms.

PCA is used to objectively verify and define the two typological species assumed for the sixty original allopatric char (for scatter plots see fig. 23; **chp. 6**). This analysis and their general nature is explained in detail in **chapter 6.** After having established the typological validity of the two char types, linear discriminant function analysis (LDFA) with equal group assignment probabilities is used on the sixty allopatric char to create an unweighted LDF (Lachenbruch 1975) that completely separates them (100% correct classification, 0% error rate). A jackknife procedure based on individuals (see **chp. 4**) is used in conjunction with the LDFA to verify the classification and error rate. The LDFA also confirmed the PCA.

This LDF is based on branchiostegal ray number (meristic no. 6), anal fin ray number (meristic no. 2), maxillary length (morphometry nos. $38 + 94$) and standard length (morphometry no. 51). These four variables are chosen because they constitute the minimum number sufficient for separation, and because they load most strongly on the second eigenvector derived in the PCA and the first canonical vector from the LDFA. The first three characters on their own also partially separate the two species, and all four can be made in the field without killing fish. For simplicity of further calculation, the LDF is calculated using untransformed data , and both the covariance and

correlation matrices result in the first canonical vector (and hence discriminant function) accounting for 100 % of the variance.

The LDF based on a covariance matrix of raw data is:

$$-0.62 \times \textit{branchiostegal number} - 0.78 \times \textit{anal fin rays} - 1.42 \times \textit{maxillary length} + 0.17 \times \textit{standard length}$$

Dolly Varden: for untransformed data < -23.

Bull trout: for untransformed data > -23.

The LDFA and PCA are verified in both sympatric regions and over the entire B.C. range to ensure that the LDF derived from the original sixty allopatric char works for all populations. PCA performed on the truss characters on the samples from each of the sympatric areas is used to identify Dolly Varden and bull trout. The PCA parameters are essentially the same as those for the original sixty allopatric char (when only their truss measurements are analyzed) but the principal component scores derived for each species and some of the character eigenvector loadings are more divergent in sympatry perhaps suggesting competitive interaction or character displacement (Baker 1980, Reyment et al. 1984). Such competition or displacement could be trophic as Cavender (1978, 1980) found substantial gill-raker morphological differences between Dolly Varden and bull trout and there are also slight but consistent differences in the overall means of gill-raker and pyloric caeca number for both species. There also is evidence for competitive effects in Arctic char populations (Barbour 1984, Fraser and Power 1984, Henricson and Nyman 1976, Hindar and Jonsson 1982, Klemetsen and Grotnes 1975, 1980, Nilsson and Filipsson 1971, Nordeng 1983, Skreslet 1973, Sparholt 1985), and in Dolly Varden/cutthroat trout (*Salmo clarki*) interactions (Andrusak and Northcote 1971, Henderson and Northocte 1985, Hindar et al. 1988, Hume and Northcote 1985, Jonsson et al. 1984, Nilsson 1954, 1960, 1963, Schutz and Northcote 1972). The LDF obtained using PCA-identified specimens is virtually identical to that derived from the original sixty allopatric char, and again the char are even more dissimilar in sympatry. The PCA and LDF based on the entire sample also successfully and similarly identifies the specimens.

Morphometric and Meristic Description of the Two Char Species

Dolly Varden and bull trout differ only slightly in shape and thus are difficult to quantitatively describe. This similarity is the main reason why they were not recognized as distinct species until recently. There are, however, one morphometric, two meristic and several qualitative features that provide a useful general description of the two char. The meristic and morphometric characters noted are those used in the LDFA and are measured on the entire char sample. Maxillary length is divided by standard length to crudely but simply adjust for size (see **chps. 5–6** for further discussion on size-adjustment). The complete and reduced character sets of fifty-one variables are summarized in **appendix A.**

The qualitative distinctions involve undefined head features (Cavender 1978, 1980). Bull trout have a larger, broader and flatter head than Dolly Varden, and also have more slender and ventrally flatter bodies (see fig. 2). Dolly Varden bodies are more oval and "snake-like", with the head not dominating the profile (fig. 3).

Dolly Varden (*Salvelinus malma* (Walbaum))

branchiostegal number: range=16–24; mean=21.2; median=21.

anal fin rays: range=9–12; mean=10.6; median=11.

maxillary length ratio: range=0.07–0.13; mean=0.10; median=0.10

Bull Trout (*Salvelinus confluentus* (Suckley))

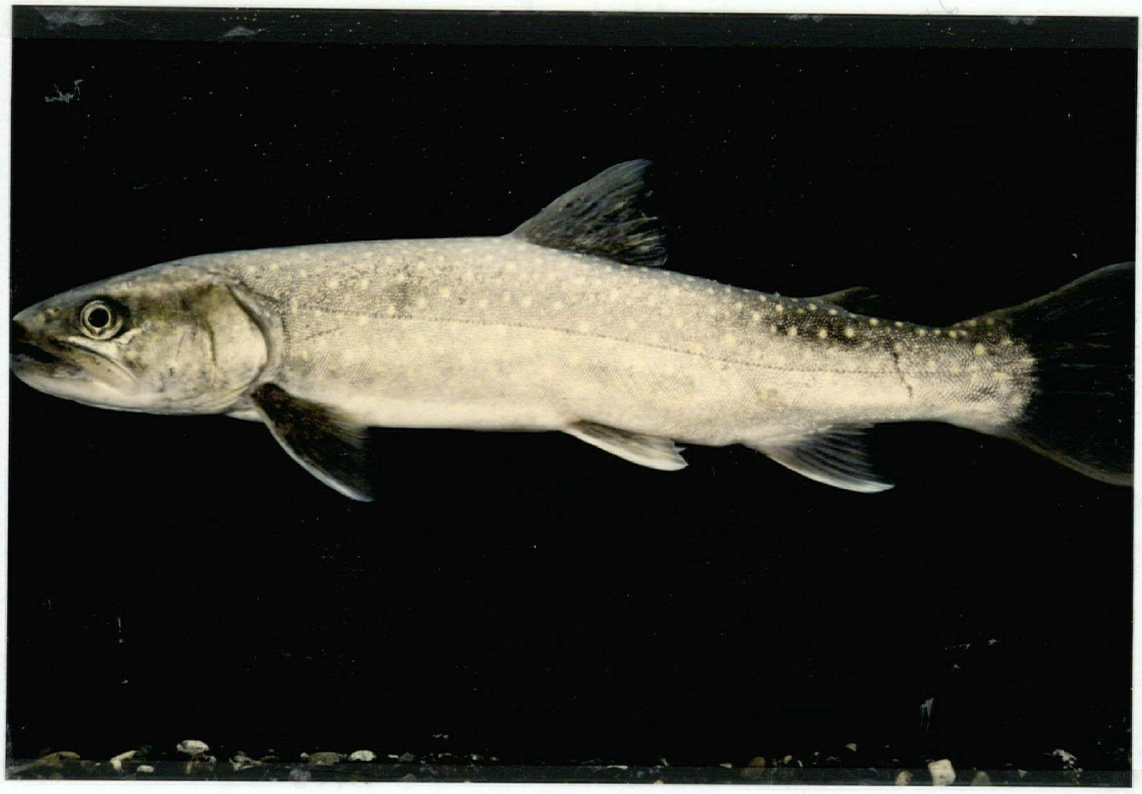
branchiostegal number: range=(rarely 22-24)25–30; mean=26.6; median=27

anal fin rays: range=9–12; mean=11.4; median=11.

maxillary length ratio: range=0.08–0.16; mean=0.11; median=0.11.

Additional Descriptive Features

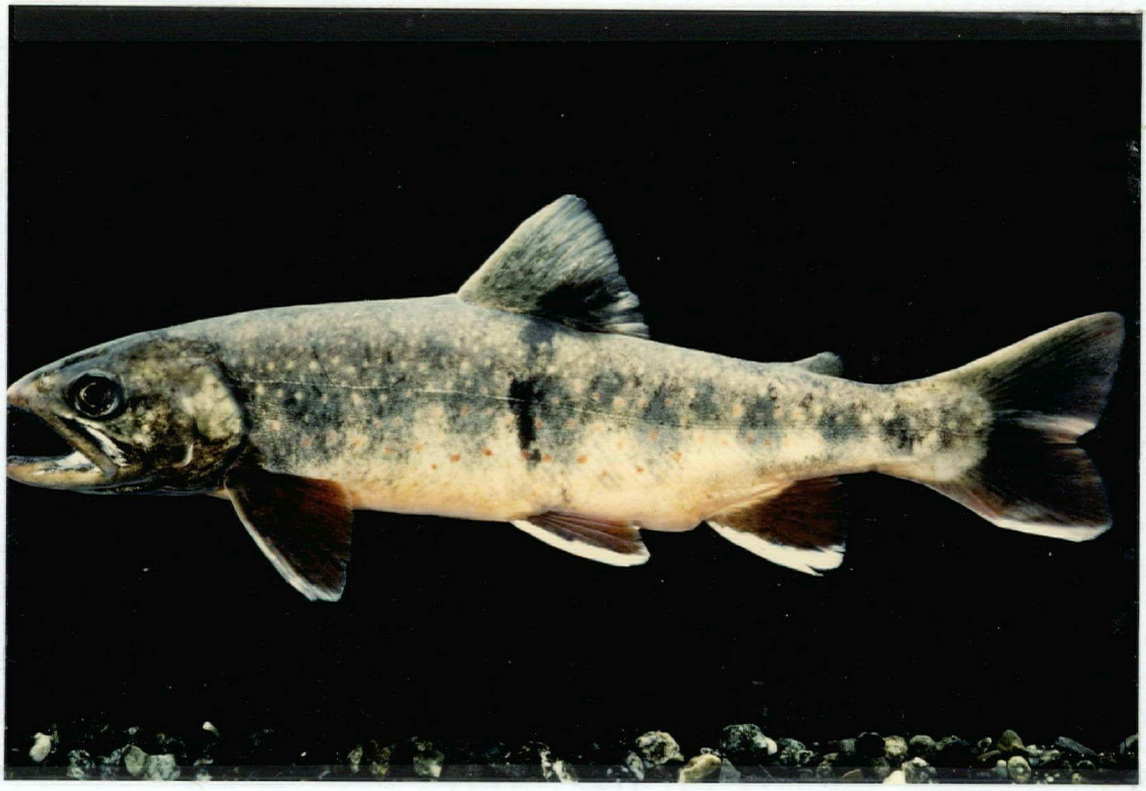
The preliminary morphometric and meristic assessment of the char crosses indicate that Dolly Varden and bull trout are distinct typological species. The characteristics that differentiate them remain the same even when the two species are reared under similar environmental conditions. A detailed analysis of the artificial hybrids is not yet completed, but the initial data indicates that



MALE

FEMALE

FIGURE 2. Typical lower Fraser River bull trout in spawning condition.



MALE

FEMALE

FIGURE 3. Typical lower Fraser River Dolly Varden in spawning condition.

hybrids are distinguishable from pure specimens and that nothing resembling hybrids are present in any of my samples. In addition, no hybrids are revealed by my PCA or LDFA either.

A preliminary electrophoretic assessment of some lower Fraser, Skeena and Nass River watershed char reveals no species-specific variability. No attempt has yet been made to assess differences in allele frequencies because of the small sample size. Previous electrophoretic work (McPhail, unpubl. data) on char from the lower Fraser, Skeena and Kootenays produced similar results. This is true of other data published on the Dolly Varden complex as well (Clayton and Ihssen 1980, Clayton pers. comm.). In fact, the genus *Salvelinus* is characterized by a low level of electrophoretic differentiation relative to other salmonid fishes (Allendorf and Utter 1979, Nyman et al. 1981), and what variability exists is almost never taxonomically characteristic in either Arctic char or Dolly Varden (Andersson et al. 1983, Armstrong and Morton 1969, Ferguson 1981, Hindar et al. 1986, McCart 1980, Mednikov et al. 1980, Omelchenko 1975, Nyman 1967, Nyman et al. 1981, Tsuyuki et al. 1966, Yoshiyasu 1973, Zakharova et al. 1971)

Dolly Varden collected for crosses in geographically proximate lower Fraser River watersheds develop spectacular, sexually dimorphic spawning colours (fig. 3). Bull trout collected nearby, but not in the same tributary, do not possess any typical char spawning characteristics (fig. 2). However, bull trout in allopatric areas, such as in south-eastern B.C., do show typical char spawning colours (Leggett 1980, McPhail pers. comm.).

Leggett (1980) examined spawning behaviour in bull trout but was unaware of their taxonomic distinctiveness. He nevertheless noted that there are differences in his interior B.C. bull trout population's spawning behaviour compared to that known for more coastal populations (Blackett 1968, Needham and Vaughan 1952). The more coastal populations are from Alaska and almost certainly are Dolly Varden. Armstrong and Morrow (1980) were aware of the Dolly Varden/bull trout distinction and describe fish from the same Alaskan coastal watershed as Dolly Varden. They also note some differences in spawing behaviour between these same coastal and interior populations.

Gould (1987) examined the development of bull trout eggs from organogenesis through to yolk sac absorption. He notes several minor and one unique difference between bull trout ontogeny and that published for Dolly Varden (Armstrong and Blackett 1980, Blackett 1968). He emphasizes

that these differences are as large as those between other char species, and Soin (1980) stresses the importance of ontogenetic information to salmonid taxonomy. Subtle ontogenetic changes have been thoroughly documented for subspecies of the Arctic char complex and other char species as well (Balon 1980a–e, 1984, Savvaitova 1973, 1980a). The ontogenetic differences between Dolly Varden and bull trout are further discussed in **chapter 3**.

Additional differences in skull osteology and gill raker morphology between Dolly Varden and bull trout are presented in Cavender (1978, 1980; also see Kolyushev 1971, Medvedeva and Savvaitova 1980). In later work, Cavender (1984; also see Abe and Muramoto 1974, Behnke 1984, Chernenko and Viktorovsky 1971, Hartley 1987, Muramoto et al. 1974, Ueda and Ojima 1984, Vasilyev 1975, Viktorovsky 1975a–b, 1978) presents and summarizes cytological evidence for differences in ploidy and karyotype arrangement between Dolly Varden and bull trout. For a general discussion on the interpretation and value of such cytological work see Arkhipchuk and Berdyshev (1987) and Sites and Moritz (1987).

Distribution

Dolly Varden are largely coastal char and bull trout are mostly interior (fig. 4; also see **chp. 2**). Ironically, it appears that the majority of fish originally described as Dolly Varden are in fact bull trout (Cavender 1980). Cavender (1978) lists distributional information for both species that extends beyond my study region. Others provide more distributional information (Crossman and McAllister 1986, Lee et al. 1980, Lindsey and McPhail 1986, McPhail and Lindsey 1986, Minckley et al. 1986) but their identification of bull trout may be suspect. The interior and coastal separation of Dolly Varden and bull trout is not complete, however, and the two species occur together in four drainage systems in B.C.. Bull trout are apparently not stenohaline (McPhail and Lindsey 1986), as I collected specimens in salt water but near-shore and close to the Fraser River estuary (Roberts Bank). Cavender (1978) also lists a near-shore marine sample of bull trout from Puget Sound. In addition, some bull trout I collected in freshwater had all the characteristics of fresh-run anadromous fish. However, bull trout have not been collected in freshwater on any offshore islands or very far out to sea, and thus although they can enter saltwater they appear not to have dispersed through the sea.

My B.C. regions of parapatry and sympatry for the two species are the lower Fraser River, the Skeena River, the Nass River and the Stikine River. The lower Fraser River drainages usually have only one of the char species present, but adjacent systems can vary in which one it is. This suggests a checker-board distribution pattern (Brown and Gibson 1983, MacArthur 1972) and the possibility of competitive exclusion. In this area actual sympatry is tentatively found only in the Capilano River, Lynn Creek, Seymour River, Dickson Lake and McConnel (Cascade) Creek watersheds. This sympatry is termed tentative because in all these samples the two char species were never caught in precisely the same place and thus the sympatry is only broad or perhaps the distribution is parapatric or syntopic.

The only Skeena River tributary that contains both species is the Tahtsa River and it is represented by a U.B.C. museum sample in poor condition. However, both char are present in the same jar in this sample and thus in this system intimate sympatry is hypothesized. I could not verify this sample because of time and its isolated location. Many regions of parapatry are present in the Skeena system as well. Cavender (1978) also recognizes the Skeena drainage as a sympatric area.

The Nass River contains several geographically adjacent tributaries that contain parapatric populations of both species. I also made true sympatric collections of both char species in single electroshocking and trap samples in tributaries to Meziadin Lake. More systems in this drainage may have similar situations but their isolation prevented more detailed exploration.

The Stikine River contains both Dolly Varden and bull trout again in close but separate drainages. No further collections were made there, but areas of sympatry and parapatry likely exist in this system as well.

Cavender (1978) identified three other regions, the Taku River, Puget Sound and formerly in the Sacramento River, where both species occur together. He also speculates on the presence of hybrids in the Skeena River drainage. I could not verify his Skeena River hybrids as the specimens were unavailable, and the lakes where they occur are inaccessible except by plane. I found no evidence for hybrids in my collections.

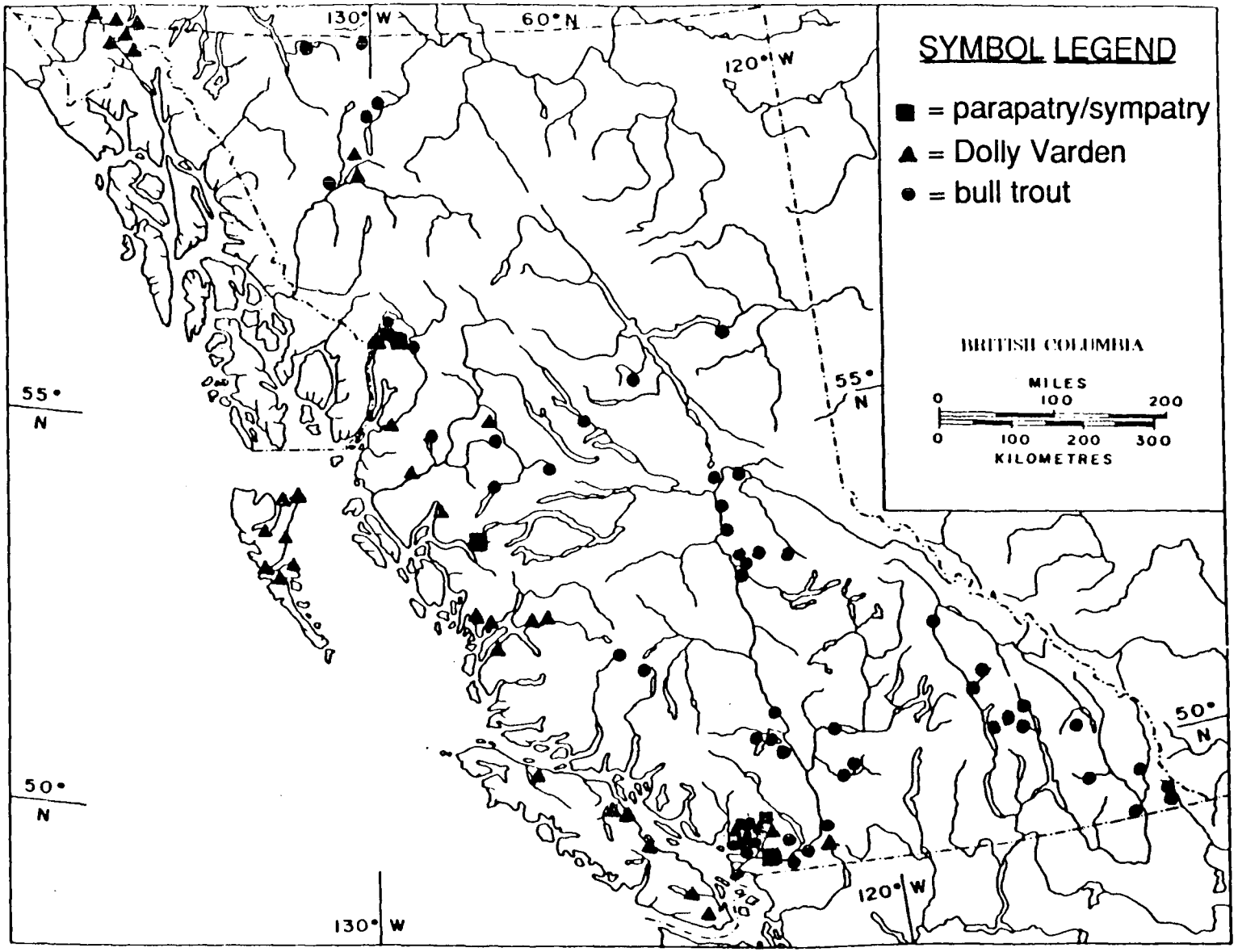


FIGURE 4. Distribution of Dolly Varden and bull trout in British Columbia.

Allopatric regions for Dolly Varden are Vancouver Island and the Queen Charlotte Islands, whereas all interior areas contain only bull trout. The interior regions with sufficient samples to specify them as allopatric for bull trout are the Peace River, Cariboo and Kootenay regions. Although I do not have enough specimens to provide an overall picture, the char from both Alberta and longitudinally similar areas in the United States appear to be exclusively bull trout as well. This latter assessment is based on a few specimens (my unpubl. data) and on Cavender (1978).

Two regions in B.C. are not represented by any char specimens in the U.B.C. museum. I attempted to collect char in the Okanagan and on the Sunshine Coast (Gibsons–Powell River) but was unsuccessful. Carl et al. (1977) and McPhail (1961) also mention the absence of char in the Okanagan and this is further confirmed by personal communication from C. Bull (B.C. Fish and Wildlife biologist for the Okanagan). There is, however, anecdotal evidence for char on the Sunshine Coast (Facchin and King 1980, Straight 1982) but I did not obtain any despite reasonable effort. Cavender (1978) had no collections from these two regions either. The only other B.C. region not represented in my study is the extreme north-eastern part of the province. This is not the result of the absence of char there (Carl et al. 1977) but rather the lack of U.B.C. museum specimens. McPhail (pers. comm.) has previously collected char in this north-eastern region and now believes them to be bull trout.

Bull Trout Taxonomic History and Etymology

Since I was unable to examine or obtain specimens outside B.C. or the U.B.C. Ichthyological Museum my analysis of the nomenclature and taxonomic history of bull trout is limited to a literature review. I, therefore, tentatively agree with Cavender (1978) on the scientific details of naming bull trout, but this opinion could change upon examination of type specimens. Inspection of Japanese and Asiatic Dolly Varden complex specimens could also affect this nomenclature. All subspecific names should be suppressed (Brown and Wilson 1954, Cracraft 1983, Wiley 1981) or at least held back until these Asiatic fish can be examined (Hubbs 1943, Lindsey and McPhail 1970, McPhail 1961, Morton 1970, Utter 1981) and until a more thorough investigation of the so-called northern form of Dolly Varden (Behnke 1980, 1984, McCart 1980, McPhail 1961, Morrow 1980)

is undertaken. The etymology for Dolly Varden is adequately described in DeLacy and Morton (1943), McPhail (1961), Morton (1970) and Scott and Crossman (1973).

The common name, bull trout, is appropriate for the new species because it is often used by local fisherman to describe large Dolly Varden in the Kootenays, Montana and Alberta (Brown 1971, Dymond 1932, Cavender 1978). The name bull trout is also listed as an alternate to (McPhail and Lindsey 1970, Scott and Crossman 1973) and was offered as a possible name for Dolly Varden in one of the original works separating it from the Arctic char complex (DeLacy and Morton 1943). This common name has also been attached to several of the precedent scientific names for this species (Cavender 1978). The only difficulty with the name is the use of "trout" to describe a char. However, other char such as the brook (*Salvelinus fontinalis*) and lake trout (*Salvelinus namaycush*) defer to this difficulty, and the use of bull trout is in accordance with the American Fisheries Society's attempt to stabilize fish nomenclature (Robins et al. 1980). Furthermore, the name bull trout is now already established (Balon 1980, Gould 1987, Johnson and Burns 1984, Leary et al. 1983, 1985, MacDonald 1985).

Bull trout were first described as *Salmo spectabilis* (Girard 1856). This holotype was collected by Suckley in 1854 (Cavender 1978, Morton 1970), who later redescribed it and corrected its collection locality information (Suckley 1860). It came from Fort Dalles on the lower Columbia River, and is now a "mutilated, half-rotted individual" (Cavender 1978) in the United States National Museum (U.S.N.M.). Suckley (1861) realized that *spectabilis* was preoccupied and substituted *Salmo campbelli*. In the same paper, he also described *Salmo bairdii* and *Salmo parkei*. No holotypes for these latter two descriptions are now available (Cavender 1978, Jordan 1879).

Suckley (1858) also described a char from Fort Steilacoom near the Puyallup River as *Salmo confluentus*. This specimen consists of a dried head and skin in the U.S.N.M. (Cavender 1978). Suckley's description suggests the head is definitely a bull trout, but that the fins were covered in dark spots. This latter characteristic is not found in char (Cavender 1978), and thus Cavender (1978) re-examined the type skin and felt that Suckley had actually described and typed *Salmo confluentus* from two different individuals, a bull trout and a Pacific salmon (*Oncorhynchus* spp.). Jordan and Evermann (1896) had placed this type specimen in synonymy with chinook salmon

(*Oncorhynchus tshawytscha*), but Cavender's (1978) analysis concludes that the head is actually a bull trout. In fact, Cavender (1978) believes that all these specific descriptions are of bull trout.

Support for this belief comes from Jordan (1879) where he realistically describes bull trout under the name *Salvelinus spectabilis*. This is apparently based on Clackamas River specimens (Cavender 1978) and on a re-examination of Girard's *Salmo spectabilis*. Here he notes that the *parkei* holotype is lost and was unquestionably the same as *spectabilis*. Jordan and Gilbert (1882) synonymized bull trout with *Salvelinus malma*. This precedent was followed by Jordan and Evermann (1896), and these, and later workers, never correctly distinguished *malma* and *confluentus*.

Five possible scientific species names, *bairdii*, *campbelli*, *confluentus*, *parkei* and *spectabilis*, thus exist for bull trout. The first four were proposed by Suckley (1858, 1861), and the last by Girard (1856). This last and original name *spectabilis* (Girard 1856) is a secondary homonym (Cavender 1978, Morton 1970, Suckley 1861) and thus cannot be used. Of the remaining four names, *confluentus* is chosen because it has publication date precedence (Suckley 1858) and the type specimen for it still exists and is in the best condition (Cavender 1978). The proposed scientific name for bull trout therefore is *Salvelinus confluentus* (Suckley) and its type specimen is U.S.N.M. 1135 (Cavender 1978).

Summary and Conclusions

The Dolly Varden char species complex in B.C. is composed of two species, Dolly Varden (*Salvelinus malma* (Walbaum)) and bull trout (*Salvelinus confluentus* (Suckley)). The species do not appear to interbreed in several regions of parapatry and of broad and intimate sympatry. There is no evidence of introgression or hybridization. The high number and overall pattern of parapatric occurrences of these two species is suggestive of competitive exclusion.

The morphometric and meristic characters that distinguish the two species are consistent throughout their range as studied herein and by Cavender (1978). Any further variation present is not related to taxonomic distinction but rather to ecology, and as Cavender (1978) points out this local variation is at a much lower level than that local variation in other recognized species of *Salvelinus*.

CHAPTER TWO

Quantitative Zoogeography of Dolly Varden and Bull Trout in British Columbia

Introduction

Biogeography is the study of the distribution of organisms in space and time (Cox and Moore 1985). The discipline can be split into two approaches which have different aims and work on different time-scales (Ball 1975, de Candolle 1820 in Nelson 1978, Endler 1982, Patterson 1981, Wiley 1981). The first school is the ecological (MacArthur 1972, MacArthur and Wilson 1967, Simberloff 1974) and studies the dispersion of organisms and the mechanisms and environmental interactions which maintain or change this dispersion. This work is usually done at the population or community level, and often involves direct experimental investigation (eg. Schoener 1974, Simberloff and Wilson 1969). The second school is the historical (Craw and Weston 1984, Croizat 1962, Croizat et al. 1974, Humphries and Parenti 1986, Nelson and Platnick 1981, Nelson and Rosen 1981, Platnick and Nelson 1978, Simberloff 1986, Seberg 1986, Wiley 1981, Wiley and Mayden 1985) and studies the spatial and temporal distribution patterns of organisms. This historical work is conducted at the taxonomic level and attempts to explain these distributions based on past events. Direct experimentation is thus often not possible and explanations rely on inference.

Neither school should be mutually exclusive (Crovello 1981, Endler 1982a–b), and this is especially true of biogeography in regions like Canada. Areas of Canada such as British Columbia (B.C.) were repeatedly glaciated until approximately ten thousand years ago (McKee 1972) and their fauna and flora were therefore eradicated or survived in various refugia (Lindsey and McPhail 1986, McPhail and Lindsey 1970, 1986). The glacial refugia acted as centres of origin for the post-glacial recolonization of this area and represent part of the historical component of the biogeographic picture for B.C.. The other historical aspects are the recolonizing organisms' phylogenies and the patterns of deglaciation which effected the recolonization. The actual recolonization of particular places, once they became inhabitable, and how each place ultimately affects the organisms is the result of ecology and phylogeny. A complete biogeographic analysis requires some integration of this ecological and historical information but must also discern which aspects belong to each of these categories. Many other fields of biology have recently seen or discussed combined studies

of history and ecology or other disciplines (Brooks 1985, Brown 1983, Brundin 1972, Cheverud et al. 1985, Clutton-Brock and Harvey 1984, Dobson 1985, Duellman 1985, Dunham and Miles 1985, Felsenstein 1985, Funk 1985, Lauder 1982, McLennan et al. 1988 (in press), Miller 1987, Mitter and Brooks 1983, Ricklefs 1987, Ridley 1983, Ross 1972, Sillén-Tullberg 1988, Stearns 1983, Wanntorp 1985).

Biogeography in short-term historical areas such as B.C. suffers from the lack of congruence between these two approaches. The ecological school is justifiably fascinated by local populations of organisms whose diversity often has developed within a short evolutionary time-scale. They are usually not interested in broad biogeographic or species-specific patterns, and they often ignore the potential influence of historical events and phylogeny. The historical school finds such small-scale differentiation problematic. The natural and legitimate tendency for them is to work at a higher taxonomic level and look for general biogeographic patterns. This search can involve a vicariance biogeographic analysis, synthetic multivariate analyses or can simply look for the most parsimonious explanations based on whatever distributional data is available (so-called descriptive biogeography (Ball 1975, Cain 1944, Wiley 1981)). None of these historical analyses usually deal with ecology, with detailed subspecific or localized variation or with the biogeographic patterns of single species.

The vicariance method uses phylogenetic systematics (Hennig 1966, Wiley 1981) and attempts to match organisms' phylogenies with geological history. It works best at a large-scale geographic and taxonomic level. Theoretically, it can work at any level, but it is limited by the precision of the phylogenetic and geographic information available (Brooks 1985, Simberloff 1986, Wiley 1981, Wiley and Mayden 1985). It has not been used to assess the biogeographic patterns of single species because their "phylogenies" are difficult to derive because of the confounding variation at that scale (Brooks 1985, Simberloff 1986; for an attempt see Parenti 1984). Even at the species level, short-history biogeographic regions present the problem of properly estimating vicariance (Simberloff 1986, Wiley 1981) if much of the speciation occurred before the last glaciation.

Multivariate analyses have the potential to look at large numbers of characters at any level. However, those tried have tended to analyze and ordinate broad geographic species groupings and patterns and/or have looked for environmental correlates to such patterns (Baker 1980, Bortone

et al. 1982, Chang and Gauch 1986, Chernoff 1982, Fisher 1968, Grady et al. 1983, Green 1971, Hill and Gauch 1980, Hughes et al. 1987, Huheey 1966, Imbrie and Kipp 1971, Larsen et al. 1986, Leland et al. 1986, Legendre 1986, Legendre and Legendre 1983, Matthews 1985, Matthews and Robison 1988, Schnell et al. 1977, Smith and Fisher 1970, Sneath and McKenzie 1973, Stevenson et al. 1974, Wilson 1974). Again, the biogeographic patterns of single species have not often been multivariately assessed, and those which have are also confounded by non-geographic subspecific variation (Baker et al. 1978, Chernoff 1982, Jolicoeur 1959, 1963b, Pauken and Metter 1971, Reyment 1961, Sokal 1965, Sokal and Rinkel 1963, Thomas 1968, Thorpe 1975b, 1976, 1983a) or have analyzed presumably less confounded genetic data (Gould and Johnston 1972, Menozzi et al. 1978, Morton and Lalouel 1973, Piazza et al. 1981a–b, Sokal 1979, Zanardi et al. 1977).

Descriptive biogeography is possible at all levels and does not require precise information, although its accuracy will be improved by it. It does, however, lack statistical rigour and suffer from personal interpretations (Ball 1975, Crovello 1981). Single species patterns are inferred, but usually only through the breakdown of a larger general taxonomic picture. As in vicariance biogeography, common distribution patterns between groups are determined before looking for any causal factors affecting the distribution of any one group. But while this "best scenario" in descriptive biogeography does qualitatively account for historical information before hypothesizing ecological or other effects it remains similarly speculative and is often not detailed. Furthermore, if morphometric or meristic characters are analyzed they are often looked at in terms of clines in one or a few variables that may not represent the overall biogeographic pattern (Thorpe 1985a–c). Descriptive analyses based on parsimony also are not always necessarily correct (Farris 1973, Felsenstein 1978, 1983, Felsenstein and Sober 1986, Sober 1983) and they can still be confounded by ecological or other non-geographic variation especially if the analysis is attempted at a subspecific level.

This lack of congruence between the ecological and historical approaches result in an inability to rigourously analyze unconfounded and detailed biogeographic patterns especially in short-history regions and for individual species. Canonical trend surface analysis (Gittins 1979, Lee 1969, Monmonier 1972, Wartenberg 1985a) can be used to analyze such biogeographic patterns. It separates

the confounding non-geographic information, can utilize data sets based on large numbers of characters and can operate at the specific and higher levels. It potentially allows broad geographic patterns of a single, or more, species to be established. Hypotheses regarding ecological and other geographically unpatterned data can also then be erected to account for this information.

Canonical Trend Surface Analysis

Canonical trend surface analysis (CTS) is based on canonical correlation analysis (CCA) (Green 1978, Hotelling 1936) and was developed for geology (Lee 1969) as an extension of trend surface analysis (Gittins 1968, Krumbein 1959, Marcus and Vandermeer 1968). Two matrices of data are required. For biogeographic analyses, one matrix is a morphometric and/or meristic data set and the other consists of locality coordinates for each individual or sample mean in that character matrix. Essentially, CCA is applied to these two matrices and it simultaneously quantifies and compares them (for details and formulae see Wartenberg 1985a). It results in two new sets of linear composites of variables, one for each original data matrix, that maximize the correlations between the morphometric data matrix and the locality coordinates. Only that morphometric variation which corresponds to the large-scale geographic pattern specified by the locality coordinates is initially accounted for and therefore the unconfounded biogeographic (historical) pattern can be inferred from the first new linear variable set that corresponds to the original character matrix.

The leftover variation which does not correspond to geography can now also be analyzed. A residual correlation matrix that contains this ecological and other non-geographic (non-historical) information can be calculated (Wartenberg 1985a) and studied further. A more localized CTS of small-scale patterns may also be helpful. Wartenberg (1985a) further suggests comparing CTS results to those from principal component analysis (PCA) to see what technique best accounts for what variables. PCA accurately summarizes morphology without accounting for geography or other features (for details see **chp. 6**). This can help resolve which characters are important in terms of overall variation, which variables do not have a spatial variation pattern, and which characters have a variation pattern on a scale too small to be resolved by CTS. This can be even further enhanced by looking at the spatial autocorrelation of the CTS residuals of each variable

(Oden and Sokal 1986, Sokal and Menozzi 1982, Wartenberg 1985a–b). I did not attempt these non-geographic analyses.

The two main limitations of CTS are that this highest correlation between morphometrics and geography may not explain much of the variability within a data set and that the morphometric and meristic character variation may not actually represent the biogeographic patterns (Crain and Bhattacharyya 1967, Norcliffe 1969, Ripley 1981, Wartenberg 1985a). The former is assessed by seeing how much of the variation is accounted for. The CCA eigenvalues give the overall percentage variance explained (for multivariate statistics explanation and terminology see **chp. 6**). It can be further and better monitored through redundancy coefficients (Cooley and Lohnes 1971, Green 1978, Stewart and Love 1968, Wartenberg 1985a; alternatively Glahn 1968) which more specifically quantify the amount of variance of one data matrix explained by the new CCA linear character composite derived for the other variable set. The "redundancy" here equates to explanatory power and thus only when it is sufficiently high should the analysis be pursued. Some modifications of CCA that maximize the redundancy of the two data matrices (DeSarbo 1981, Johannson 1981, van den Wollenberg 1977) are available but they require *a priori* knowledge of the variability and will not likely make much difference in the overall analysis (Wartenberg 1985a). To avoid any bias, they were not used by Wartenberg (1985a) or by me.

The latter CTS limitation of requiring biogeographically representative character variation can really only be assessed through personal knowledge of the species and their possible biogeographic patterns. Any statistical analysis, especially multivariate, should be carefully re-examined if it does not make biological sense (Corruccini 1975, 1978, 1987, Pimentel 1979, Reyment et al. 1984). If the resultant CTS biogeographic pattern is not realistic it should be suspect. Common overall distribution patterns should be kept in mind when analyzing the pattern of any single species. Moreover, it is reasonable to assume that character variation can be used to identify biogeographic patterns (Endler 1977, Morishima 1969, Thorpe 1976, 1985a–c), and in short-history, post-glacial regions it has often been employed for that purpose (Khan and Qadri 1971, Lindsey 1956, 1964, 1975, McAllister and Lindsey 1959, McPhail and Lindsey 1970). While environmental heterogeneity could result in geographically unpatterned data, adjacent or nearby populations of a species are likely to be more closely related and respond more similarly to a common environment than

are those of geographically distant populations of the same species (Endler 1977, Thorpe 1976). This is particularly true in the case of the biogeography of short-history regions such as B.C.. Here the recolonizing populations may have come from several separate glacial refugia and thus already have accumulated substantial morphometric differences as a result of several thousand years (forty thousand in B.C.) of isolation (McPhail and Lindsey 1970, 1986). Geographically adjacent populations of such a species should be more similar to each other and different from populations in other areas. Therefore, the geographic aspect of character variation should identify biogeographic patterns.

Zoogeography of Dolly Varden and Bull Trout in B.C.

Dolly Varden and Bull Trout

The zoogeography of Dolly Varden (*Salvelinus malma*) and bull trout (*Salvelinus confluentus*) in B.C. provides an excellent test and example of CTS. The genus *Salvelinus* is notorious for its variability and until fairly recently Dolly Varden were not even recognized as a distinct species. They were lumped into the Arctic char (*Salvelinus alpinus*) species complex. Upon their recognition as a distinct species, Dolly Varden themselves received attention as a char species complex and eventually the bull trout was suggested (Cavender 1978) and verified as a distinct species (for a complete discussion and references see **chp. 1**).

These difficulties with taxonomy in char stem largely from univariate analyses of their morphometric and meristic character variability being insufficient (Behnke 1980, 1984, Frohne 1973, McPhail 1961, Mednikov et al. 1980, Morrow 1980) and this character variation being more localized than geographic (Chereshnev 1982, Hammar 1984, McPhail 1961, Savvaitova 1980a–b). This makes an analysis of their zoogeography easily confounded. Indeed I first attempted to identify their zoogeographic patterns through PCA and canonical variates analysis, but these analyses resulted in there being no patterns except at a local level. In less variable species, such analyses may have detected general patterns but they still would likely be better elucidated with CTS.

Another feature of char that makes them attractive for a CTS zoogeographic analysis is that they are essentially freshwater fish (Armstrong and Morrow 1980, McPhail and Lindsey 1970, Scott

and Crossman 1973). While anadromy is not uncommon, especially in Dolly Varden, char all spawn in and spend at least their juvenile years in freshwater. Freshwater fish are particularly wedded to geography because of their restricted capacity to disperse and their post-glacial distribution in B.C. must be the result of a limited set of recolonization routes. This freshwater restriction also makes the objective selection of regional watershed localities for the CTS much easier (Crovello 1981, Krumbein 1955, Legendre 1986).

A final but important aspect is that these two char species probably evolved and thus were already distinct before the last glaciation event started (Behnke 1980, 1984, Cavender 1970, 1978, 1980, 1984, 1986, Uyeno and Miller 1963, Smith 1981, Wilson 1977; but see Clemens 1953, Jones 1959, Neave 1958, Norden 1961, Vladykov 1964). Consequently, while their distribution and variation was nonetheless greatly influenced by glaciation, their speciation was probably not the result of it. In a similar way, the last glaciation in Canada resulted in several geographic types of Arctic char but complete speciation is not evident (Behnke 1980, 1984, Johnson 1980, Kircheis 1976, McCart 1980, McPhail 1961, Morrow 1980a, Qadri 1974, Savvaitova 1980b, Scott and Crossman 1973). This is true as well of most other freshwater fish groups in Canada (Cavender 1986).

Glacial History of B.C.

The geological history of B.C. is complex and I will only present those details relevant to char distribution. A complete geological account can be found in McKee (1972) and a more general ichthyological picture in McPhail and Lindsey (1970, 1986) and in Lindsey and McPhail (1986). This geographic complexity will also challenge the ability of the CTS to provide a coherent zoogeographic analysis.

B.C. is characterized by a variety of terrain but is predominanted by mountains. The main mountain building period started in the Miocene and continued through to the Pliocene (McCrossen and Glaister 1964). These periods were also volcanic. Since most of the major rivers in B.C. have maintained a continuous westward flow in deep gorges across these mountains they are believed to predate geological uplifting. Essentially, this implies that they have maintained their present courses since at least the Pliocene (McPhail and Lindsey 1986).

The last glacial period in B.C. occurred in the Pleistocene and was called the Wisconsin (or Fraser). It started about fifty thousand years ago and ended about ten thousand years ago. The Cordilleran ice-sheet of the Wisconsin glaciation extended to just south of the present-day Canada/U.S.A. border (Olympia, Washington) and covered B.C.. It was preceded by three other extensive glacial events all separated by relatively mild ice-free periods (McPhail and Lindsey 1970, 1986).

Three major and possibly one minor glacial refuge provided the fauna that recolonized B.C. (fig. 5). The largest ice-free area was the Pacific refuge which was the lower two-thirds of the Columbia River system (McPhail and Lindsey 1970, 1986). The Pacific refuge was nevertheless affected by glaciation as the upper one-third of the present Columbia River system, which is the part in B.C., was glaciated as were many of its major tributaries.

The second major refuge was the Bering. It existed in the Yukon River basin area and was not glaciated because it lies in a rain shadow behind the high coastal mountains surrounding the Gulf of Alaska (Hoffman 1981, Hopkins 1972, McPhail and Lindsey 1970). This lack of precipitation still exists today and despite lower temperatures in the glacial period this probably prevented the buildup of extensive snow fields. This refuge also received fauna from across the Bering Land Bridge (Lindsey and McPhail 1986). This temporary isthmus contained freshwater connections and linked Alaska and Siberia during the Wisconsin glaciation because the ice-sheets locked up enough water to substantially lower the sea-level (Hopkins 1959, 1973, Nelson et al. 1974).

South of the Bering refuge several minor refugia may have existed, one of which is important to this B.C. scenario. There is some biological argument about whether or not the Queen Charlotte Sound area was glaciated (Calder and Taylor 1968, McPhail and Lindsey 1986, Moodie 1972a–b, Moodie and Reimchen 1976) but there is little geological evidence (Heusser 1960, Howes 1982, Karlstrom 1961, Warner et al. 1982) and several alternative biological explanations (McPhail and Lindsey 1986) for the apparent endemism in this area. Nevertheless, this area was deglaciated early about 15,000–16,000 years ago (Warner et al. 1982) and may have been dry when sea-levels were lower (Klein 1965, McPhail and Lindsey 1970, 1986)

The third major refuge was the Missouri. It lay east of the Rocky Mountains in the present-day northern United States (McPhail and Lindsey 1970). This region was beyond the maximum extent of glaciation and its fauna recolonized many areas on the Great Plains adjacent to the Rocky Mountains. In this analysis it is important because of a possible influence on the zoogeography of char in the Peace River and other nearby drainages. The Peace River is the only major B.C. river to rise in the west and flow east through the Rocky Mountains (Continental Divide). There may also have been temporary water connections between the Missouri refuge and the Columbia River system (Lindsey and McPhail 1986, Malde 1965, McPhail and Lindsey 1986, Minckley et al. 1986, Wheeler and Cook 1954).

Deglaciation began about fifteen thousand years ago but did not proceed uniformly. The coastal areas became ice-free relatively early and the southern regions may have been ice-free as recent as twenty-five thousand years ago (Clague et al. 1980). The coastal areas sunk under the weight of the ice and upon deglaciation were initially flooded by the sea (Matthews et al. 1970). The present coastal watershed system was established about thirteen thousand years ago when these lowlands rebounded. However, there may also have been a glacial re-advance in the Lower Fraser River area about eleven thousand years ago (Armstrong 1981, Tipper 1971). As well, the regions east of the Rocky Mountains rapidly de-glaciated to the northeast (McPhail and Lindsey 1970, 1986).

Deglaciation produced many temporary shifts in river drainage patterns. The glacial ice blockage of major upstream Columbia River tributaries resulted in the ponding of a giant glacial lake called Glacial Lake Missoula (> 300 km long). The Columbia River was similarly constricted in several places further downstream and this formed large ephemeral lakes. When the ice dam broke on Glacial Lake Missoula the entire lake drained apparently in two weeks (Baker 1973, 1988) and a cycle of floods in the downstream constrictions swept the lower Columbia River regions several times (Bunker 1982). This resulted in the massive scouring of the Columbia Plateau evident today in the complex channeled scablands of eastern Washington state (Allen et al. 1986).

The Fraser River was blocked by ice early in deglaciation and initially drained through the Thompson/Okanagan River basin into the Columbia River (Fulton 1969). The upper Fraser

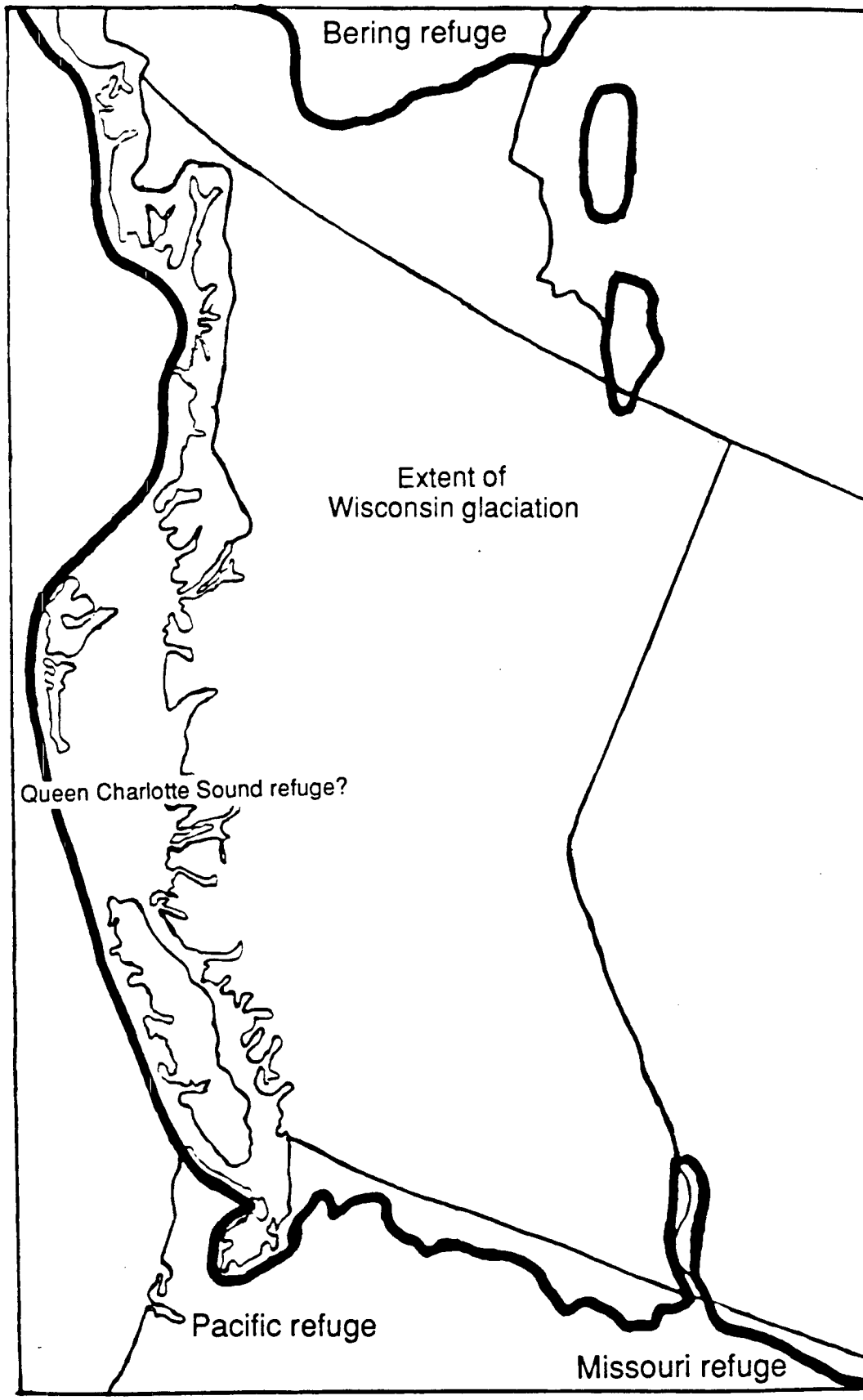


FIGURE 5. Glacial refugia relevant to British Columbia.

River was also a tributary to the Peace River at two separate times and drained eastward across the Rocky Mountains (Holland 1964, Tipper 1971). On removal of the ice barriers, the Fraser River took up its present southwest course. In addition, temporary post-glacial connections were established between the headwaters of the Fraser and Skeena Rivers, Skeena and Peace Rivers, Stikine and Yukon Rivers, and Peace and Yukon Rivers (Bostock 1969, Clague and Rampton 1982, Holland 1964, Lindsey and McPhail 1986, McPhail and Lindsey 1970, 1986, Nelson et al. 1974, Templeman-Kuit 1980, Workman 1978). Similar ephemeral connections through the Snake River system existed between the Columbia River (Pacific refuge) and the Missouri refuge as well (Malde 1965, McPhail and Lindsey 1986, Miller 1965, Wheeler and Cook 1954). Little else is known about deglaciation patterns in B.C.. This deglaciation information suggests that probable reinvasion routes were along the coast, south from the Bering refuge, west from the Peace River drainage, and north from the Columbia River both at its uppermost northern reaches and through its former Fraser River connection.

Materials and Methods

All analyses and graphics are based on computer macros I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Center, University of British Columbia. They are available from me.

Separate CTS's are carried out in the same manner on the zoogeography of Dolly Varden and bull trout. Each species' distribution in B.C. is objectively divided into its appropriate major river drainages (figs. 6a–7a) as given in Carl et al. (1977) and McPhail and Lindsey (1986). Each region is represented by samples collected in different years and at different times, by char of both sexes and where possible by char of all possible habitat types (lakes, rivers, and streams). The less accessible northern drainage regions have smaller sample sizes, but all are still composed of at least two populations and usually more (fig. 4; **chp. 1**). The regional means of the morphometric and meristic characters and of the locality coordinates of the populations sampled in that region are calculated for each species. These character and locality means form the two matrices entered into the species-specific CTS as described in Wartenberg (1985a). Spurious correlations are checked by separately jackknifing (see **chp. 4**) out regions and characters from the analysis and recomputing

the CTS (Chernoff 1982, Neff and Marcus 1980). The jackknife results are essentially identical to the complete CTS.

The resultant first new CTS linear character vector is then entered into unweighted average-linkage cluster analysis based on the Euclidean distance between the regions (Hagmeier 1966, Hodkinson 1980, Hoffman et al. 1979, Johnston 1969, Thorpe 1975b, 1976). Only the first CTS vector is used because it accounts for the most correlated variation and thus presumably the most geographically patterned information. My analysis is robust to the clustering technique and distance matrix used (Boyce 1969). The cluster analysis gives the non-hierarchical relationships between the regions. The degree of these relationships are then used to determine the zoogeographic patterns of each of the char species in B.C. (figs. 6b–7b). Potentially more sophisticated and informative visual representations (Dougenik and Sheehan 1979, Guptil and Starr 1988, Piazza et al. 1981a, Wartenberg 1985a), were not readily available to me.

The most parsimonious recolonization patterns into these regions (figs. 6a–7a) is assessed by a species-specific minimal spanning tree analysis (MST) (Gower and Ross 1969, Prim 1957). The locality coordinate means of the regions for each species are entered into MST. The distance network that fits the minimum connected distance between all these regions is calculated. This shows the shortest possible connections between the regions, but does so without accounting for geographic history. The patterns of deglaciation and the presence of any barriers to dispersal are not part of the MST. Therefore, it presents the most parsimonious statistical explanation for the zoogeographic distribution, but not necessarily the most parsimonious biological one. However, it still represents a simplest pattern and thus provides some background against which to assess the CTS. The other assessment of the CTS comes from descriptive explanations already offered by Lindsey and McPhail (1986) and McPhail and Lindsey (1970, 1986).

The char variables employed in my analyses are the twenty-six truss measurements (Bookstein et al. 1985, Strauss and Bookstein 1982) and the four characters used to obtain the three linear discriminant function parameters as explained in **chapter 1**. The twenty-six truss measurements are displayed in figure 1 (**chp. 1**). The four other characters are branchiostegal number (meristic no. 6), anal fin ray number (meristic no. 2), maxillary length (morphometry nos. 38 + 94) and

standard length (morphometry no. 51). All are explained in **appendix A**. These thirty characters are sufficient to characterize both Dolly Varden and bull trout and also to recognize ecological morphotypes within each species. The PCA scatter plot patterns of these variables are the same as those for a much larger character set (fifty-one characters). They should thus be more than adequate in variability and number for accurately representing and distinguishing zoogeographic patterns.

The locality coordinates used in the analysis are the mean latitudes and longitudes of each region's populations (figs. 4, 6a–7a). The eight major watershed regions for Dolly Varden (fig. 6a) are the Lower Fraser River, Vancouver Island, Central Coast, Queen Charlotte Islands, Skeena River, Nass River, Stikine River and Tatshenshini River. The nine major watershed regions for bull trout (fig. 7a) are the Lower Columbia River, Upper Columbia River, Lower Fraser River, Central Fraser River, Upper Fraser River, Skeena River, Nass River, Stikine River and Peace River.

Results and Discussion

Dolly Varden:

The CTS analysis for the Dolly Varden accounts for 61.7 % of the overall variance in the first vector and its redundancy coefficients are high. This indicates an interpretation of this vector is acceptable. This overall variance level is sufficient, but its relatively low value could suggest the presence of significant localized and ecological variation.

Figure 6b reveals the regional relationships based on geographically patterned character variation. The Lower Fraser River and Vancouver Island regions are distinct, and I believe they were recolonized from the Pacific refuge (McPhail and Lindsey 1986). The other main dendogram branch contains all the northern watersheds and I argue they represent recolonization from the Bering refuge (Lindsey 1975, Lindsey and McPhail 1986).

The Skeena River's dendogramic and geographic positions are closer than the other northern drainages to that of the Lower Fraser River and Vancouver Island. The Skeena River may have received a secondary coastal invasion of Dolly Varden from the Pacific refuge, but since the

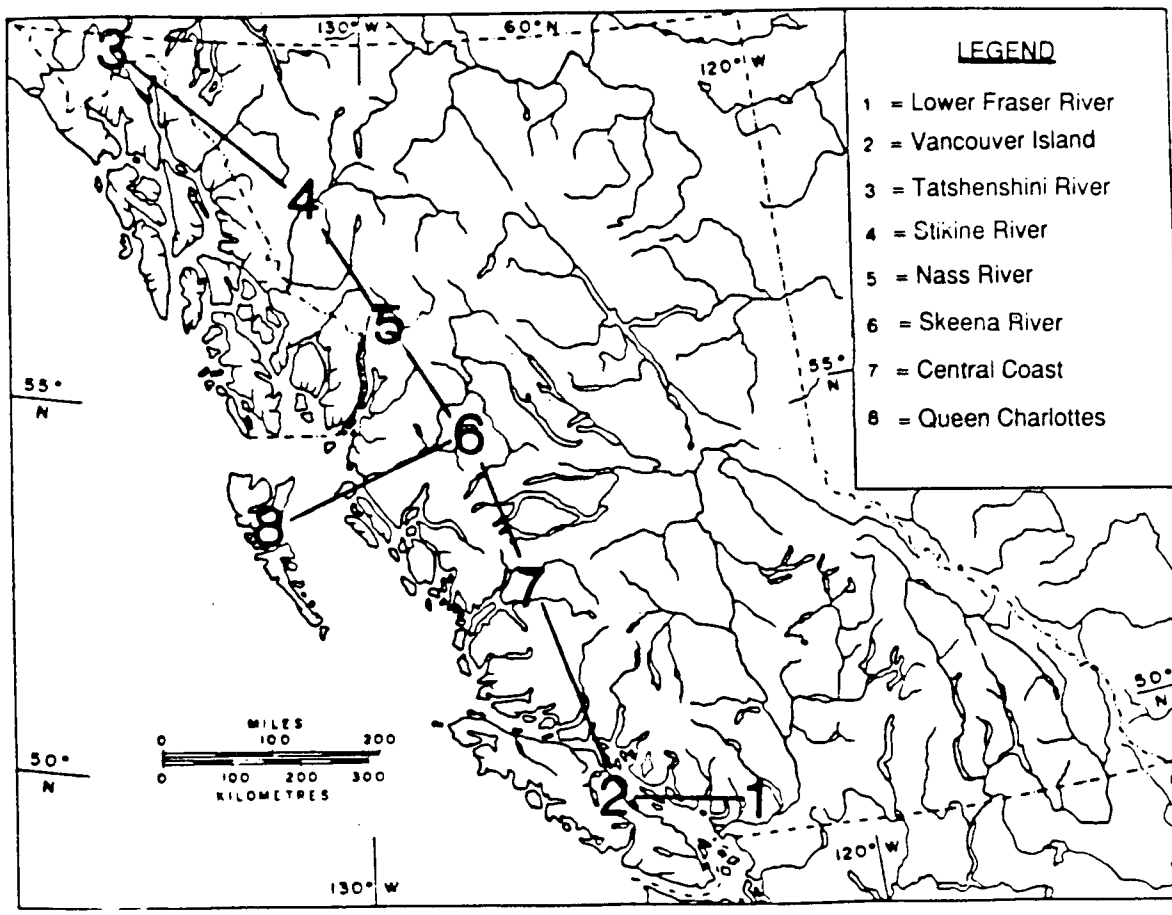
Central Coast and Queen Charlotte Islands watersheds apparently did not this explanation is less satisfactory. It is possible that the Queen Charlotte/Central Coast was already occupied when the Pacific refuge Dolly Varden arrived and that only the larger Skeena River watershed was still open for substantial recolonization.

The main Dolly Varden refugia in the Wisconsin glaciation were the Pacific and Bering. The addition of Columbia River and Yukon River Dolly Varden to this CTS would test this hypothesis but they were unfortunately unavailable.

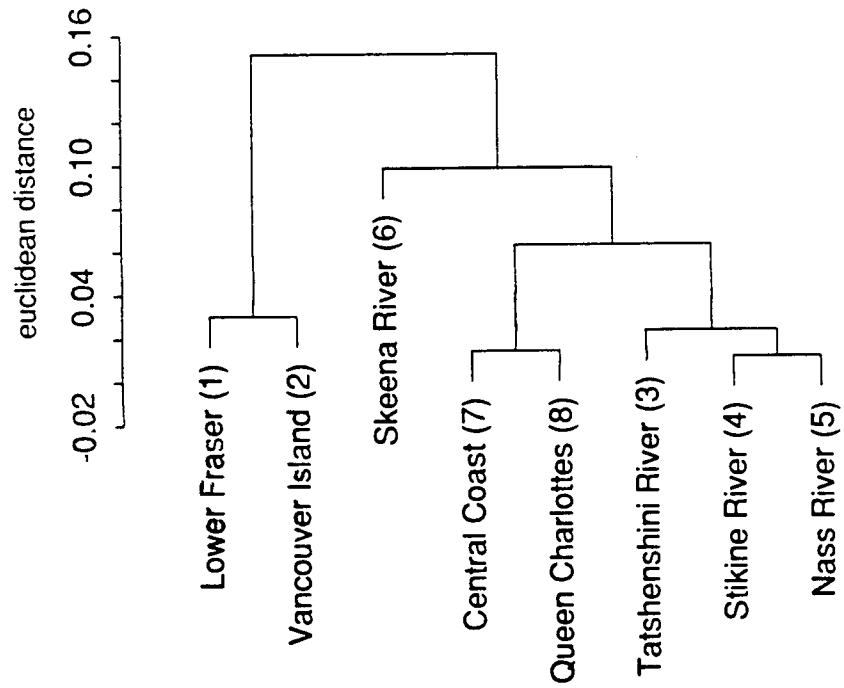
A minor Dolly Varden refuge in the Queen Charlotte Sound area is also possible. The Queen Charlotte Islands and Central Coast watersheds are distinct from those closer to the Bering refuge and yet the Dolly Varden in this area have no close relationship to the Pacific refuge drainages. Therefore, they may have had their own refuge and later received some char input from the Bering refuge. Alternatively, any similarity of the Queen Charlotte/Central Coast system to the other northern Bering refuge drainages could be due to proximity and/or parallel evolution, or the Queen Charlotte Islands and Central Coast were recolonized from the Bering refuge and no Queen Charlotte Sound refuge existed. A Queen Charlotte Sound refuge could also explain why Pacific refuge fish seem to have had no influence in these regions if the Skeena River did indeed secondarily receive Pacific refuge fish.

When the Vancouver Island region is broken up into northern and southern sectors, the northern area comes out similar to the Queen Charlotte/Central Coast region (for geology see Howes 1982). The southern Vancouver Island region, however, remains most similar to the Lower Fraser River and distinct from the rest (for geology see Alley and Chatwin 1979). This analysis is not included in the dendogram because my northern Vancouver Island region is then only composed of a single population. This is nonetheless further evidence for the possibility of a Queen Charlotte Sound refuge.

If the CCA locality vector loadings are directly analyzed, most of the character variation loads most strongly onto latitude. This conforms to the CTS zoogeographic scenario interpreted here and to the MST (fig. 6a). The zoogeographic pattern for Dolly Varden does not exactly follow the statistically shortest MST route but the patterns from each refuge are consistent. This

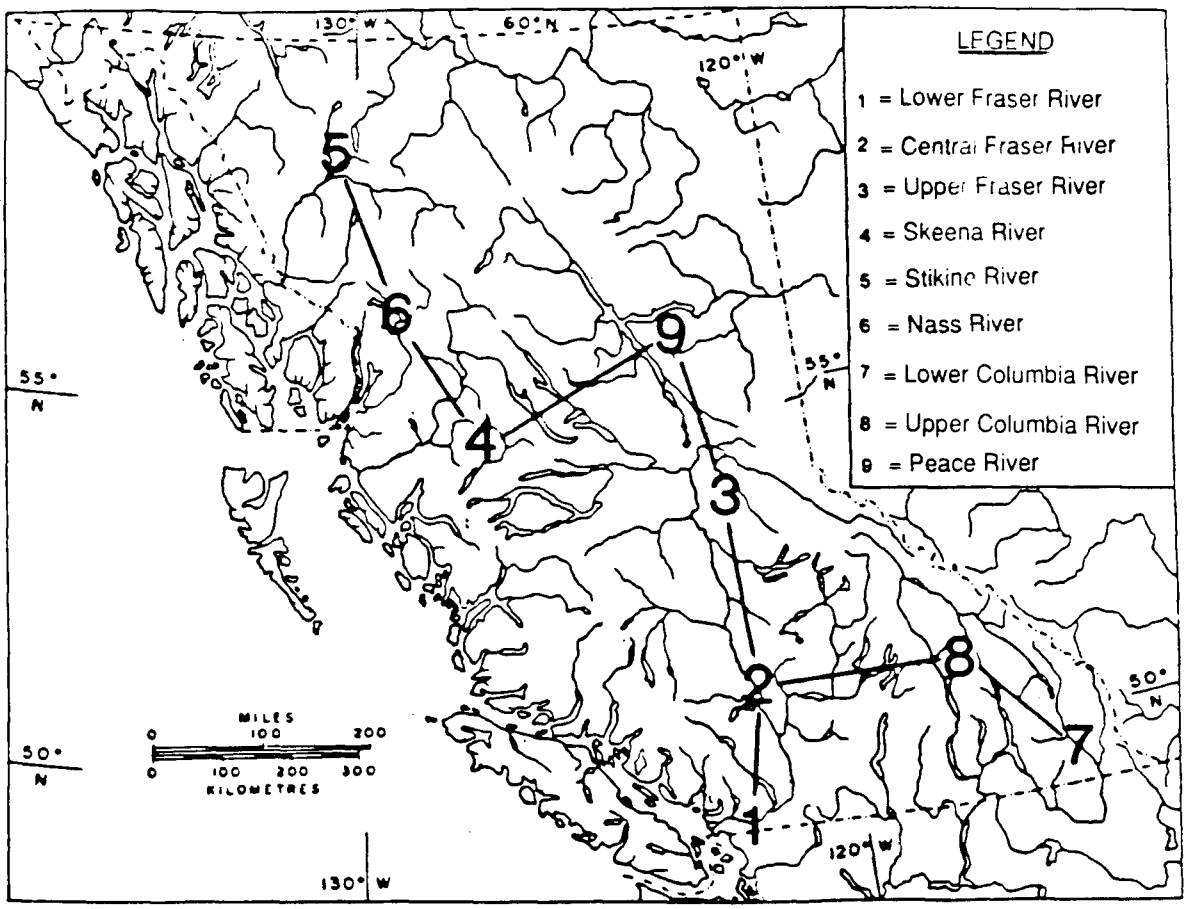


a. Minimal spanning tree (shortest statistical) recolonization routes.

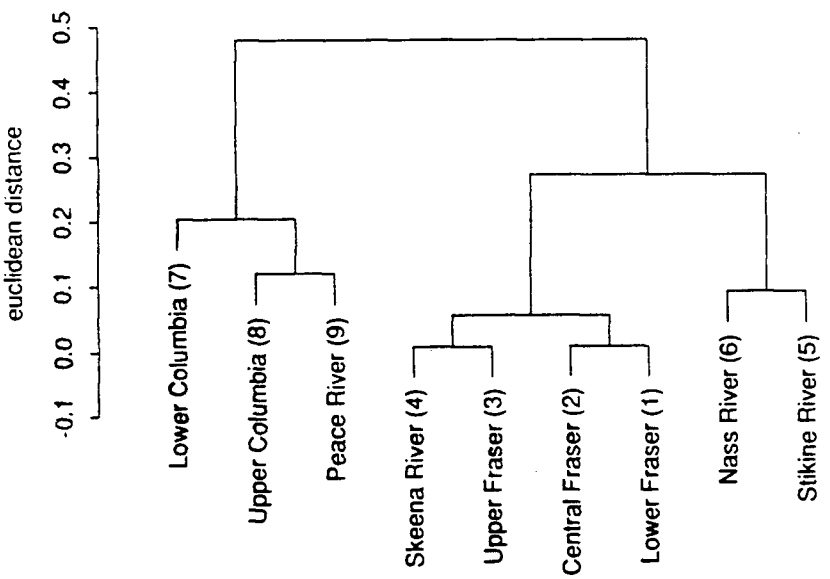


b. Canonical trend surface analysis zoogeographic relationships.

FIGURE 6. Quantitative zoogeography of Dolly Varden in B.C.



a. Minimal spanning tree (shortest statistical) recolonization routes.



b. Canonical trend surface analysis zoogeographic relationships.

FIGURE 7. Quantitative zoogeography of bull trout in B.C.

is probably because their recolonization route was mostly coastal or barely interior and thus the MST did not get misled by geography it cannot account for.

No single character loads much more strongly than others onto the CCA character vector but anal fin ray number and maxillary length are higher. Branchiostegal rays provide the single primary taxonomic distinction between Dolly Varden and bull trout, but their variation is not significant within the Dolly Varden (see **chp. 1**). This further strengthens the separation of the Dolly Varden complex into these two species since the major character separating them has insignificant intra-specific geographic variation (see bull trout results and discussion as well).

Bull Trout:

The CTS analysis for bull trout accounts for 77.8 % of the overall variance in the first vector and its redundancy coefficients are high indicating an interpretation of this vector is acceptable. The overall variance level is sufficient and relatively high. This, and the higher bull trout redundancy coefficients, may indicate that more of their variation is geographically patterned than in the Dolly Varden. This is confirmed by the larger Euclidean distances between bull trout populations compared to Dolly Varden (figs. 6b–7b). The bull trout also appear to be less morphometrically variable than the Dolly Varden as evidenced by the comparative degree of their scatter on the PCA plots in figure 23 (**chp. 6**). This PCA data coupled with the Euclidean distances suggests that the lesser bull trout variability is more between river drainages while the greater Dolly Varden variability is higher within river drainages. Dolly Varden thus seem to have more habitat related variability than bull trout. This information could be utilized in other non-geographical analyses.

Figure 7b shows the regional relationships based on geographically patterned character variation. The Columbia and Peace River bull trout are strikingly distinct from all other drainages. These char are either from the Pacific refuge or the Missouri refuge, but a lack of samples from the actual lower Columbia River (fig. 7a) in present-day Washington/Oregon makes this assessment speculative.

The absence of Lower Columbia River samples may also explain my inability to more definitely account for the recolonization of the Fraser River. While it is not impossible that the Lower

Fraser River was recolonized from the Bering refuge, it is extremely unlikely. I thus speculate that the Fraser River was recolonized from the Pacific refuge and that my Columbia (in B.C.) and Peace River regions received bull trout from the Missouri refuge (Lindsey and McPhail 1986, McPhail and Lindsey 1986, Minckley et al. 1986). This surmise accounts for their disjunct relationship on the dendogram and for the divergence between the Fraser River regions and the northern river drainages. Further evidence is that other distinct and similar differentiation also exists in other species found in the Columbia River (Bisson and Bond 1971, Bond 1973, Hubbs and Miller 1948a–b, Lindsey 1956, Loudenslager and Thorgaard 1979, McAllister and Lindsey 1959, McPhail and Lindsey 1986, Miller 1965, Smith 1966, 1975). This explanation makes sense in light of the more easily interpretable Dolly Varden zoogeographic picture and fits well with the aforementioned geological evidence as well.

In addition, it is doubtful that any bull trout came into the Fraser River through its short-lived Thompson/Okanagan River valley connection with the Columbia River (McPhail and Lindsey 1986) unless they used this connection and have since gone extinct there. Either of these hypotheses is verified by the absence of bull trout in the Okanagan. The overall distribution of the Dolly Varden complex (fig. 4; **chp. 1**) suggests if any char are present in the Okanagan they should be bull trout and not Dolly Varden.

The Stikine and Nass Rivers were probably recolonized from the Bering refuge (Lindsey 1975, Lindsey and McPhail 1986). This Bering dispersal assumes that bull trout survived in that refuge. While the char in Alaska are almost certainly Dolly Varden (see **chp. 1**) it is not clear whether bull trout also exist there. I have a single upper Yukon River watershed sample that are bull trout (also see Lindsey and McPhail 1986) and certainly bull trout exist in most nearby southern watersheds. I have no other Yukon River drainage samples from this area. If this single upper Yukon River population is included in the cluster analysis it groups out with the Stikine and Nass Rivers. I left it out of the dendogram though because a single sample is too small to characterize a region. It is thus possible that bull trout were present in the Bering refuge.

Additional evidence is the apparent relative intolerance of bull trout for sea-water (see **chp. 1**). The bull trout has been collected in saltwater and as anadromous-appearing individuals in freshwater but they are not present on any of B.C.'s coastal islands. Therefore their ability to migrate long distances in the sea may be limited and only have extended to the Fraser River or perhaps not even to there if the bull trout used the temporary Thompson/Okanagan connection. The freshwater influence of the Fraser River in the regions where I have caught "anadromous" bull trout also greatly decreases the salinity of the seawater (Clark and McInerney 1974). .

The Skeena River is most similar to the Fraser River regions on the dendogram and thus apparently was recolonized from the Pacific refuge and not from the Bering (McPhail and Lindsey 1986). Since the bull trout is less saltwater tolerant and not found in the Central Coast region it may have dispersed from the Pacific refuge via the Fraser River drainage. Dispersal through the Fraser drainage is not unreasonable (McPhail and Lindsey 1986) since my hypothesized Missouri refuge fish only reached the upper Columbia River and could have entered the Peace River from waters east of the Rocky Mountains (Christiansen 1979, Crossman and McAllister 1986, Lindsey and McPhail 1986, Minckley et al. 1986, Nelson 1977, Paetz and Nelson 1970, Reeves 1973, Rutter 1980). The Bering and Pacific refuge hypotheses again make sense in light of the more easily interpreted Dolly Varden zoogeographic picture.

The main bull trout refugia in the Wisconsin glaciation were the Pacific, Bering and Missouri. The addition of true lower Columbia River and Alberta bull trout samples to this CTS would more concretely test my hypotheses but they were unavailable.

If the CCA locality vector loadings are directly analyzed, most of the character variation loads onto longitude. This is opposite to that of Dolly Varden but conforms to the CTS zoogeographic scenario interpreted here and to the MST (fig. 7a). The zoogeographic picture for bull trout does not fit the statistically shortest MST route as well as the Dolly Varden probably because bull trout recolonization was largely through the interior of B.C. and the MST cannot account for the complex geography in this area.

No single character loads much more strongly onto the CCA character vector but anal fin ray number and maxillary length are again loading somewhat higher. As for Dolly Varden, branchiostegal ray variation is not significant within bull trout alone. This strengthens the specific separation of the bull trout from the Dolly Varden complex since the major single character separating the species shows insignificant intra-specific geographic variation. Further zoogeographic confirmation for this dual species status is that both species appear to have co-existed in the Pacific and Bering refugia.

Summary and Conclusions

The Dolly Varden seem to have recolonized B.C. from both the Pacific and Bering refugia. The Nass, Stikine and Tatshenshini River watersheds were likely recolonized from the Bering refuge. The Lower Fraser River drainages and Vancouver Island probably received Dolly Varden from the Pacific refuge. The Skeena River watersheds appear to have had Dolly Varden dispersal from the Bering or both of these refugia. The Queen Charlotte Islands and Central Coast evidence argues that they may have had a separate refuge, or that they were recolonized from the Bering refuge as well.

Bull Trout seem to have recolonized B.C. from the Pacific, Missouri and Bering refugia. The Fraser and Skeena River watersheds were probably recolonized from the Pacific refuge. Portions of the Columbia River system in B.C. and the Peace River drainages contain bull trout that appear to have dispersed from the Missouri refuge. The Nass and Stikine River watersheds likely received bull trout from the Bering refuge. The co-existence of Dolly Varden and bull trout in the Pacific and Bering refugia, and the lack of intra-specific geographic character variation in their single major taxonomically distinguishing character further confirms the specific separation of bull trout from Dolly Varden.

Canonical trend surface analysis (CTS) could be an effective method for analyzing geographically unconfounded biogeographic patterns. It provides a realistic and detailed zoogeographic picture for the recolonization of a complex and recently glaciated region by two char species that are notoriously variable and often are at an extremely localized level. This CTS zoogeographic analysis appears to work well in spite of relatively smaller sample sizes for northern regions and

an examination of only part of the overall range of the Dolly Varden species complex. If the whole species complex range was analyzed much of the speculation in this interpretation probably could disappear.

CHAPTER THREE

The Paedomorphic Evolution of Dolly Varden and Bull Trout

Introduction

Ontogeny is the course of growth and development from fertilization to the cessation of growth. Classically, its study has been cellular and sequence descriptive in nature and is commonly referred to as embryology (Balinsky 1981). In the past, such developmental work was central to evolutionary theory (Darwin 1859, de Beer 1958a), but since the demise of the all-encompassing biogenetic laws of Haeckel (1866; also see Garstang 1922, Hertwig 1894, Meyer 1935, Oppenheimer 1959, Shumway 1932, Weismann 1904, Wilkie 1967) it became less significant (Gould 1977, Nelson 1978b). In fact, embryology provided some of the strongest criticisms of the neo-Darwinian evolutionary school that predominates today (Brooks and Wiley 1986, Eldredge and Gould 1972, Goldschmidt 1940, Ho and Saunders 1979, Løvtrup 1974, Rensch 1959, Schindewolf 1950). The role ontogeny does or does not play in this modern evolutionary synthesis is poorly understood (Hamburger 1980, Raff and Kaufman 1983) but it is potentially significant (Alberch 1980, 1985, Alberch et al. 1979, Bonner 1982, de Beer 1958b, Fink 1982, Goodwin 1982, Goodwin et al. 1983, Gould 1977, Maynard Smith et al. 1985, Stanley 1979, Waddington 1962).

Recently, interest in the study of ontogeny in evolution has been revived (Alberch 1985, Alberch et al. 1979, Atchley 1984, Atchley et al. 1984, Blackstone 1987a, Bonner 1982, Creighton and Strauss 1986, Emerson 1986, Fink 1982, Gould 1977, Kluge and Strauss 1985, Maynard Smith et al. 1985, Ricklefs 1979, Wake 1966, Wayne 1986). This new work is on a more gross level than classical embryology, and primarily interprets ontogeny as single developmental events based on allometry. These allometric developmental events are the outcome of processes rather than processes themselves, even though they are usefully viewed and discussed as the latter (Goodwin 1982, Kauffman 1983, Nijhout et al. 1986). The actual roles of intrinsic and extrinsic factors in ontogeny are still to be identified. Nonetheless, the allometric developmental events are supported by empirical cellular allometric studies (Gerhart et al. 1982, Hall 1984, Katz 1980, 1982, Laird 1965, Laird et al. 1965, 1968, Odell et al. 1981, von Bertalanffy 1960).

The new ontogenetic studies examine heterochronic morphometric and osteological data collected on the same individuals throughout their development. Such longitudinal studies reveal the ontogenetic events by which evolution might occur, but such experiments are time-consuming and confined to laboratories. This is not a criticism but an admission that these constraints will continue to be a hindrance to ontogenetic research.

Morphological and osteological data are collected in many biological disciplines, especially systematics, but the data are almost never longitudinal. Usually the data are based only on adults (static data) but sometimes an entire size-range of different individuals of the same species is measured (cross-sectional data). Static data are not ontogenetic and should not be used as such unless interpreted cautiously (Atchley and Rutledge 1980, Bonner 1965, Cheverud 1982b, Gould 1971, Lande 1979, Mosimann and James 1979, Shea 1985, Sweet 1980, White and Gould 1965). Cross-sectional data, however, can provide insights into ontogeny and the role it plays in evolution (Bookstein et al. 1985, Fink 1982, Shea 1983, Strauss and Fuiman 1985, Sweet 1980).

The char, genus *Salvelinus*, show considerable ontogenetic variability and flexibility and this has been suggested as an important component in their evolution (Balon 1980a–e, 1984, Kircheis 1976, Maekawa 1984, Savvaitova 1973, 1980a). For instance, Arctic char (*Salvelinus alpinus*) can rapidly attain several discrete levels of morphometric differentiation, at times within a single individual's ontogeny, that would coincide at least with subspecific designations in taxonomy (Frost 1965, Nordeng 1983, Savvaitova 1980a, Skreslet 1973). Dolly Varden (*Salvelinus malma*) and bull trout (*Salvelinus confluentus*) appear to be less ontogenetically variable as species than Arctic char, and this consistent within-species ontogeny could provide an excellent test of the role of ontogeny in the evolution of these two species. My data set on these two species is cross-sectional and thus also presents an example of how ontogeny can be examined using common morphometric data.

Any analysis of ontogenetic evolution between these two char species requires that their phylogenetic relationship be established (Alberch 1985, Creighton and Strauss 1986, Fink 1982). While no strict phylogenetic systematic analysis (Hennig 1966, Wiley 1981) has been undertaken for *Salvelinus* (but see Balon 1984, Behnke 1980, 1984, Savvaitova 1980a–b; for family Salmonidae see Cavender 1970, Fink and Weitzmann 1982, Holčik 1982, Rosen 1974), the evidence suggests that bull trout are more primitive than Dolly Varden. Most of this evidence is morphological (Behnke

1980, 1984, Cavender 1978, 1980; also see Kolyushev 1971, Medvedeva and Savvaitova 1980, Morrow 1980), cytological (Cavender 1984; also see Abe and Muramoto 1974, Behnke 1984, Chernenko and Viktorovsky 1971, Hartley 1987, Muramoto et al. 1974, Ueda and Ojima 1984, Vasilyev 1975, Viktorovsky 1975a–b, 1978) and embryological (Armstrong and Blackett 1980, Balon 1980e, 1984, Blackett 1968, Gould 1987, Soin 1980). Unfortunately, since no phylogeny exists for the entire genus it is also not known whether Dolly Varden and bull trout are sister species. However, their morphological similarity and widely overlapping geographic ranges (see **chp. 1**) suggest that they are at least closely related. Dolly Varden may be more closely related to Arctic char than to bull trout but both these species probably had a common ancestor like bull trout. In terms of this ontogenetic assessment, their true phylogenetic relationship may influence the interpretation of the data but it does not affect the analytic approach.

Multivariate Morphometric Cross-sectional Ontogenetic Data Analysis

My analysis of ontogeny using cross-sectional data involves multivariate morphometric procedures which partition morphological variability into ontogenetic size and shape parameters. This formal size and shape model was proposed by Gould (1977) and expanded by Alberch et al. (1979; also see O'Grady 1985, Thompson 1942), but its actual implementation has remained theoretical, analytically bivariate, or limited to longitudinal studies. Principal component analysis (PCA) can be used to obtain ontogenetic size and shape factors, even for cross-sectional data (Jolicoeur 1963a, Jolicoeur and Mosimann 1960, Pimentel 1979, Reyment et al. 1984). The first principal component (PC) scores represent size and the second PC scores represent size-adjusted shape for each individual. This shape factor is plotted against the size factor to obtain the cross-sectional data equivalent of a growth curve (fig. 8). These curves will be termed allometric curves since they do not represent true growth (Alberch 1985, Blackstone 1986, 1987a–c, Cheverud et al. 1983, Cock 1966, Strauss and Fuiman 1985).

There are several advantages to a multivariate ontogenetic analysis. Growth, size and shape are all multivariate factors and not directly measured variates (Humphries et al. 1981, Thorpe 1983b, Thorpe and Leamy 1983). Consequently, all the characters have a size measure and overall size is in effect a composite and not a single variable. There is no problem with having to choose

a single representative size variable and the intercorrelations of all the characters are used rather than ignored (Lande and Arnold 1983, Reyment et al. 1984). The allometric hypothesis here is multivariate and therefore probably more realistic than hypotheses based on bivariate comparisons.

True growth curves require longitudinal data because it provides actual chronological ages and overall body morphologies at those ages (eg. Alberch and Alberch 1981, Alberch and Gale 1983, 1985). My allometric curves demonstrate how body shape changes with size, and my assumption therefore is that size is a realistic surrogate for chronological age (Bookstein et al. 1985, Cheverud et al. 1983, Creighton and Strauss 1986, Lohmann 1983, Shea 1983, Strauss and Fuiman 1985, Sweet 1980, Takai 1977). Composite multivariate size is a biological time estimate that may sometimes be more robust than chronological time because it is directly tied to growth and somewhat environmentally adjusted. It also is more consistent and less variable than an individual measure (Alberch 1980, Alberch et al. 1979, Strauss 1987, Strauss and Fuiman 1985; but see Blackstone 1987c, Laird 1965). Therefore, properly identified, ecologically diverse, samples can be better compared in a specific level analysis. Moreover, this is not an unreasonable assumption, especially if the cross-sectional data set analyzed has a sufficient sample size to demonstrate overall trends and limit the influence of outliers. Afterall, growth rates are known to be tightly regulated about their mean (Creighton and Strauss 1986, Eisen 1975, Herbert et al. 1979, Kidwell et al. 1979, Riska et al. 1984, Tanner 1963). In fact, outliers could be removed from the analysis of overall trends and investigated later to determine the reasons for their difference.

Since these are not true longitudinal growth curves they can have such otherwise unusual features as negative slopes (fig. 8). Negative slopes imply negative growth. Clearly this is impossible in true growth curves but in allometric curves it simply means that the largest organisms have shapes similar to the smallest ones (also see Bookstein et al. 1985). While the largest char depicted in figure 8 obviously appear visually different, they are nonetheless similar in shape to the smallest char when their morphometric truss characters are size-adjusted.

To assist with the interpretation of these allometric curves, I fit idealized ontogenetic trajectories (Alberch et al. 1979; also see Waddington 1957, 1962) onto them. These straight lines (fig. 8) are drawn from the plot origin to the means of the largest size-class of each species on their allometric curves. The ontogenetic trajectories are intended to represent the growth of each

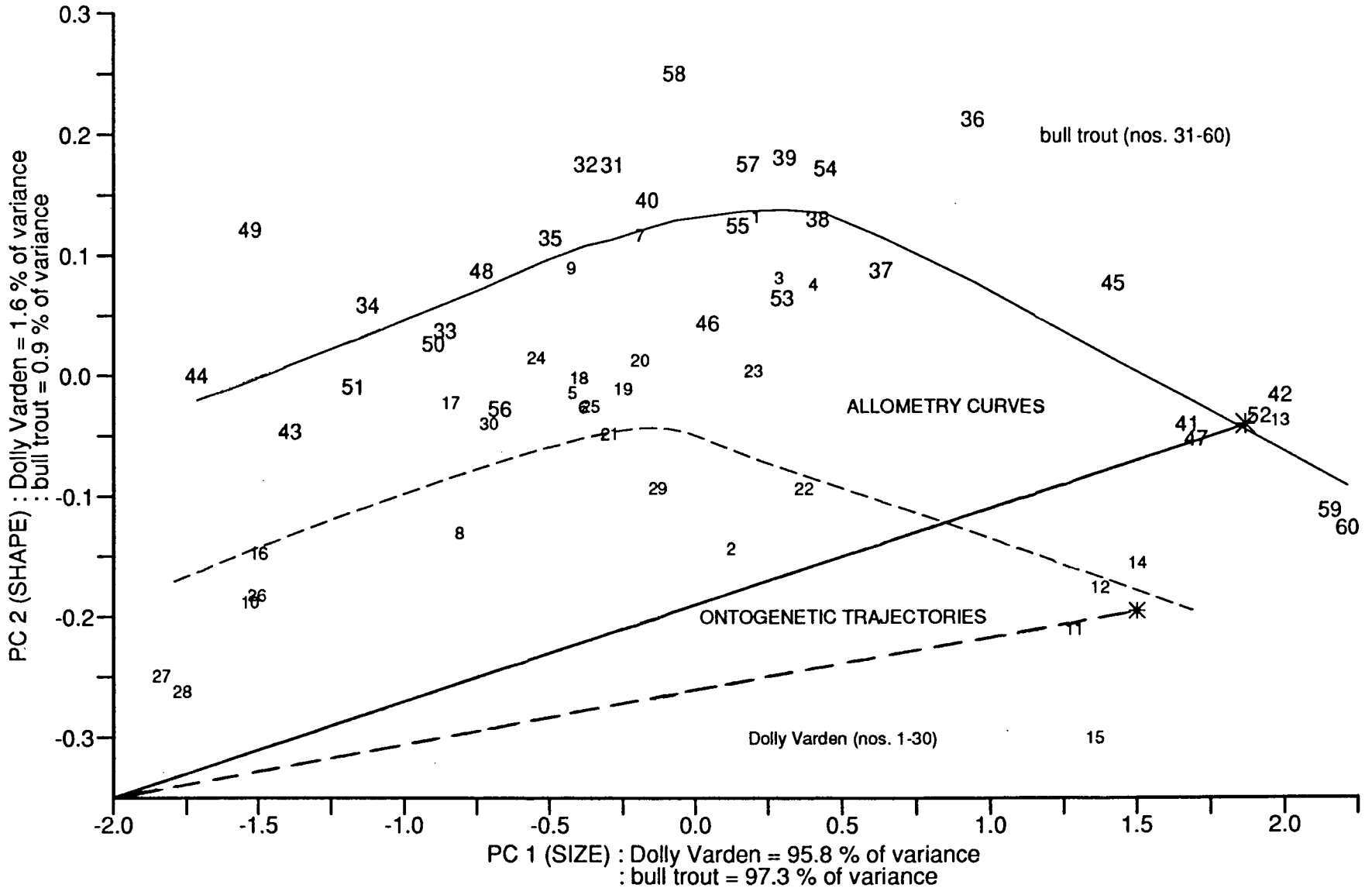


FIGURE 8. PCA allometry curves and idealized ontogenetic trajectories.

species based on all the characters examined in the PCA. It is very difficult to collect char in the smallest size-classes so the interpolation of the ontogenetic trajectory to the plot origin is speculative. However, my preliminary unpublished morphometric data on laboratory-reared crosses of Dolly Varden and bull trout suggest that there are no significant differences in their incubation time, shape at hatching, or overall indeterminate growth rate so this speculation and interpolation is not unreasonable (Alberch et al. 1979, Cheverud 1982b, Katz 1980, Larson 1980). Furthermore, von Baer's law (von Baer 1828) suggests that the early stages of closely related species will be more similar than the adult stages, and this law appears to generally hold true (Cheverud et al. 1983, Gould 1977, Lande 1979).

This PCA representation of ontogeny is supplemented with an analysis of the actual allometric growth rates and the developmental integration of each character in each species. Ontogenetic allometric growth rates are calculated using the entire available size-range of individuals, whereas static allometric growth rates are computed for any size classes that may be of further interest. In my case, static allometric growth rates are calculated for the three size-classes (small, intermediate and large) that are apparent on the allometric curve plot (fig. 8). So allometry can be assessed, the allometric growth rates also are rescaled and centred about an isometric value of one. This isometric value is represented on all plots by a dotted line.

Materials and Methods

Intraspecific PCA is carried out on a covariance matrix of twenty-six $\log_{10}$ transformed truss measurements (Bookstein et al. 1985, Strauss and Bookstein 1982) taken from approximately four hundred char (see **chp. 1**). This large data set all follows the same allometric trends shown in figure 8 and thus strengthens the utility of this procedure. It was not necessary to remove outliers. Only the first two PC's are significant. PCA on correlation or sheared matrices are also alike, as is the PCA on the total (both species combined) matrix. A covariance matrix is used because it provides the most realistic morphometric output, and PCA on separate matrices for each species is employed so that separate character growth rates and developmental integration information could be calculated. The assumptions tested for this data set, the literature and analytic justification of each PC as size or shape, and the PCA techniques and terminology are described in **chapter 6.**

The composite multivariate size measure is assessed as an age indicator, and for its consistency and equivalency between the Dolly Varden and bull trout, through the correlation of their intraspecific normalized size vectors (Cheverud 1982a, Creighton and Strauss 1986). High correlation coefficients indicate strong size parallelism between species. Size differences are also evaluated by cross-checking the actual standard lengths of the char against their PC1 scores. As well, the static allometry coefficients calculated for each intraspecific size category should reveal any allometric differences in the ontogeny of either species that might affect the overall interpretation and their relationship to size.

The twenty-six truss measurements are presented in figure 1 (see **chp. 1**) and **appendix A**. The results of the analysis are similar regardless of which individuals in my total data set are used. Consequently, I opted to employ only my sixty original char (thirty of each species) as my **part II** statistical analyses are based on these individuals and I have a larger character set for them than for any other individuals. Dolly Varden are represented by the numbers 1–30 set in small type and bull trout by the numbers 31–60 set in large type (fig. 8). Only the twenty-six truss measurements are used so that an accurate and interpretable representation of body shape is achieved and because they still meet the necessary statistical assumptions. The curves are fitted to the allometry plots using a locally weighted robust regression technique designed to smooth scatterplots (Cleveland 1979), but they could just have been fitted by eye. Similar, but damped, allometry plot patterns (fig. 23; see **chp. 6**) are seen with the **part II** PCA on fifty-one variables. Therefore, this reduced truss character set appears to give compatible and representative results. Truss characters have previously been used in a study (Winans and Nishioka 1987) of body shape changes in another salmonid, coho salmon (*Oncorhynchus kisutch*), during its developmental transition from freshwater to sea-water (smoltification), and also in a study on sculpin (Family Cottidae) ontogeny (Strauss and Fuiman 1985).

Intraspecific allometric growth rates for each truss character are calculated in the same manner regression and PCA allometry coefficients are respectively estimated in **chps. 5–6**. Only the PCA estimates are used here because in this case the two coefficient types are virtually identical (also see Jolicoeur 1963a–b, Leamy and Bradley 1982, Shea 1985). These allometry coefficients are not simply empirical descriptions of growth patterns. They have been shown to be the solution to

the differential equation relating the growth rates of a character and body size to time (Lande 1985, Reeve and Huxley 1945, Shea 1985, Strauss 1987). Thus, the allometric coefficients are equivalent to the growth rates of morphometric characters relative to body size (Alberch et al. 1979, Creighton and Strauss 1986, Wayne 1986).

For ontogenetic data (cross-sectional or longitudinal), the first eigenvector loadings are estimates of the rates of change of individual characters with size. These loadings become allometry coefficients when they are proportionately rescaled so that overall rate of change is isometric and therefore is centred about one (Hills 1982, Shea 1985, Strauss 1987). If an allometry coefficient is one then that character is isometric, if it is greater than one then positive allometry is present, and if it is less than one the allometry is negative. The size of the allometry coefficients, greater or less than one, indicates how strongly the characters are positively or negatively allometric. It does not indicate their allometric growth rate. Allometric growth rates of each character are represented by the magnitude of their allometry coefficients (or the unscaled eigenvector loadings). Intraspecific mean growth rates are assessed as the mean of the ontogenetic allometric growth rates for all the characters for that species.

Growth rates for static data are not realistic since they are based on individuals of only one size or age class. Consequently, they have no growth within that group (Atchley and Rutledge 1980, Cheverud 1982b, Gould 1971, Lande 1979, Shea 1985, White and Gould 1965). However, they can still be knowingly interpreted and cautiously compared (Bonner 1965, Gould 1971, Mosimann and James 1979), especially if they are only used intraspecifically. In my study, the static allometric growth rates seem to be realistic as they make biological sense (Pimentel 1979, Reyment et al. 1984). This result may be because my static analysis contains three size groups that probably are not completely homogeneous for age but rather represent a limited size range.

Developmental stability for the entire data set is assessed using the formula integration = 1 – (correlation matrix determinant) (Cheverud et al. 1983; also see Olson and Miller 1958, Scagel et al. 1985), using an analysis of the smallest eigenvector (fig. 12) resulting from the PCA (Gower 1967, Holland 1968, Jolicoeur 1963b, Reyment 1979, Reyment et al. 1984), using isometric patterns (fig. 10) of ontogenetic allometry coefficients (Wayne 1986) and using a qualitative assessment of the relative scatter on the first two PC axes (fig. 8) on the PCA scatter plot (Neff and Smith 1979,

Wayne 1986). The integration formula provides an overall estimate of developmental correlation where high integration values indicate strong character correlations and developmental homeostasis. The smallest eigenvector represents that linear combination of variables which is relatively invariant in the sample and thus provides general information on growth invariant or highly canalized developmental patterns. Most small eigenvectors have similar overall patterns. This smallest PC assessment has been mathematically substantiated (Gnanadesikan and Wilk 1969). The ontogenetic allometry coefficients which are near isometry also indicate developmentally canalized characters. The relative amount of scatter on PCA plots gives an indication of whether most of the variation within and between species is based on size or shape, and if the individual relationships to the allometric curves are strong or not.

All analyses and graphics are based on computer macros I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Centre at the University of British Columbia. These programs are all available from me.

Results and Discussion

The mean correlation coefficient between the intraspecific multivariate size vectors (PC1) is 0.97 indicating that the sizes portrayed here are parallel in Dolly Varden and bull trout. This is corroborated by the tight relationship between their size vector scores and their actual standard lengths, and by the similarity of all the static and the ontogenetic allometry coefficients (figs. 10–11). Furthermore, nearly all of the approximately four hundred char analyzed here fit their respective allometry curves (fig. 8) with little deviation. These strong size and shape relationships suggest that this size factor makes a realistic comparative time scale for these two species' ontogenies, and that the ontogenetic differences that are present relate to shape and not to overall body size.

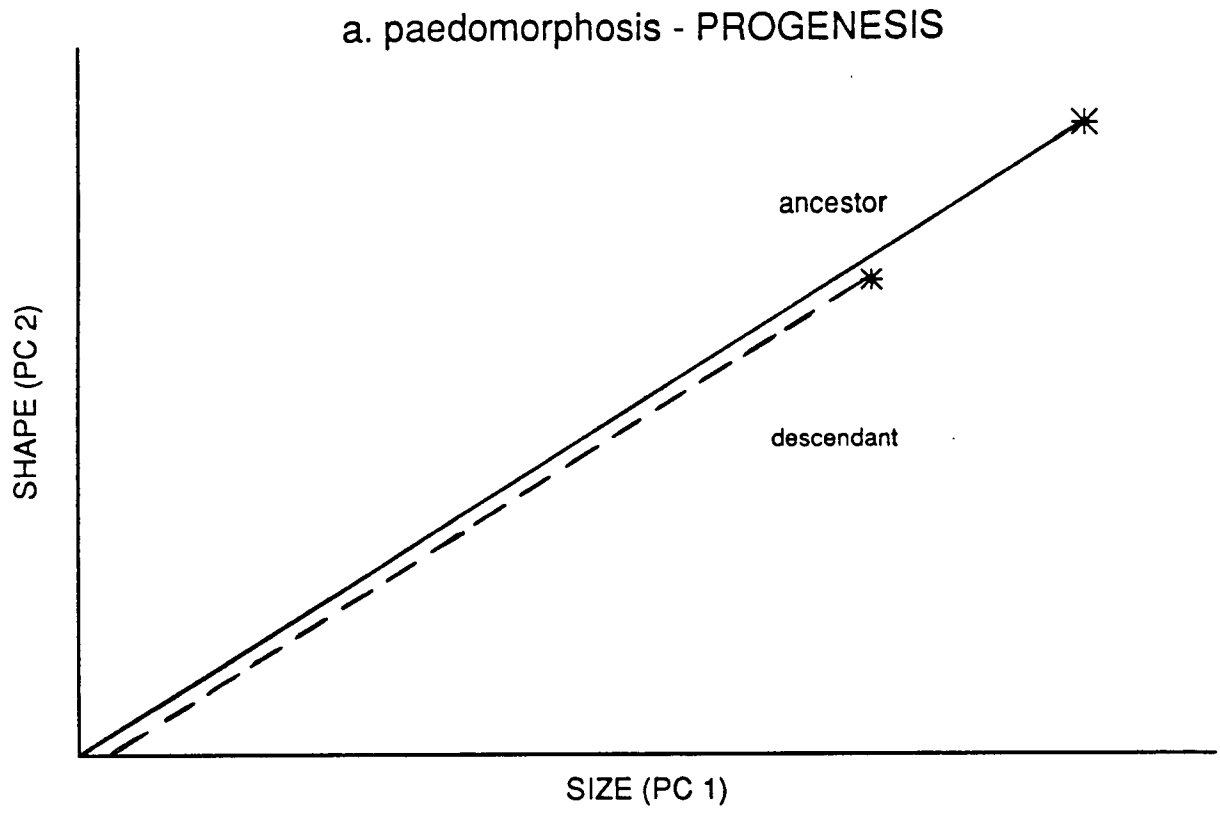
The allometry curves and ontogenetic trajectories for Dolly Varden and bull trout in figure 8 are essentially slightly displaced mirror images of each other. This information coupled with the apparent lack of interspecific variation in size, incubation time, shape at hatching, overall mean growth rate, sexual maturation time and time of growth cessation suggests that the differences

between these species could be the result of ontogenetic changes in their relative growth rates and timing.

Since I have assumed that bull trout are the more primitive species, the only ontogenetic mechanism that can account for this pattern is that Dolly Varden evolved from bull trout through paedomorphosis (juvenilization). Other recent and similar examples of paedomorphosis as an ontogenetic mechanism for speciation are Alberch and Alberch (1981), Bell (1981), Gould (1968), Guerrant (1982), Larson (1980), Shea (1983) and Wake (1966). If my phylogenetic assumption is wrong and bull trout are the more derived species, the ontogenetic explanation would simply be reversed and peramorphosis would be the evolutionary mechanism (Alberch et al. 1979, Fink 1982).

Paedomorphosis is taken to be the retention of ancestral juvenile characteristics by later developmental stages of descendant forms. This can result through two mechanisms, progenesis and neoteny. Progenesis is paedomorphosis produced by the precocious sexual maturation of an organism that is still at a morphometrically juvenile stage. The expected ontogenetic trajectory for progenesis is given in figure 9a (Alberch et al. 1979). There is no evidence in my data or in figure 8 that suggests that Dolly Varden and bull trout have different maturation times. Both species appear to become sexually mature at the intermediate size stages. Therefore, progenesis does not seem to be a probable ontogenetic mechanism for the evolution of Dolly Varden from bull trout. It may, however, play a role within-species in the case of stunted char populations.

Neoteny is paedomorphosis produced by the actual retardation of development in certain characters so that the adult organisms attain sexual maturity at full size while retaining mostly ancestral juvenile characteristics. An idealized ontogentic trajectory for neoteny is depicted in figure 9b (Alberch et al. 1979). This ontogenetic trajectory for neoteny is similar to that seen in Dolly Varden and bull trout. Indeed, the shape of Dolly Varden is like the shape of juvenile bull trout (fig. 8). Other authors have also suggested that the evolution of char, especially those members of the Arctic and Dolly Varden char species complexes (see **chp. 1**), has occurred through paedomorphosis by neoteny (Balon 1980a–e, 1984, Kircheis 1976, Maekawa 1984, Savvaitova 1973, 1980a). This may be true of many other fish species as well (Balon 1979, 1981, 1983, Bell 1981, Hubbs 1926).



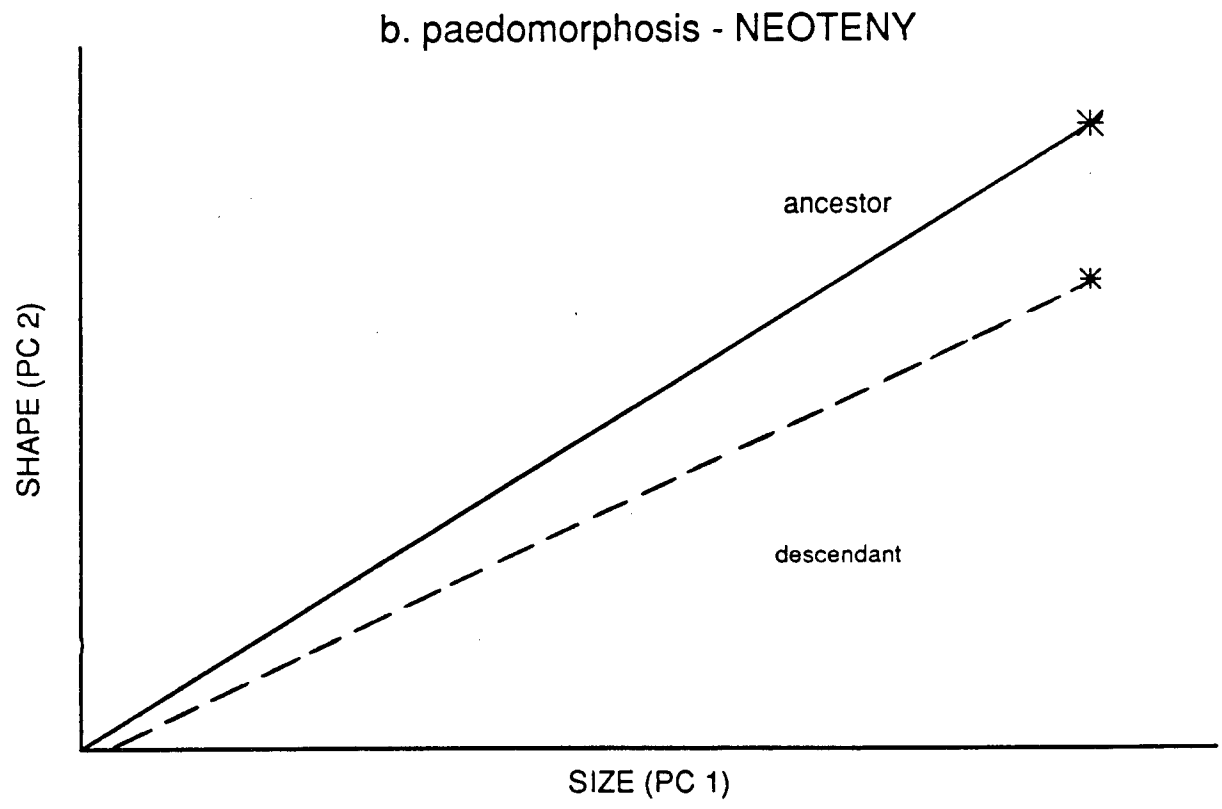


FIGURE 9. Idealized ontogenetic trajectories for paedomorphosis.

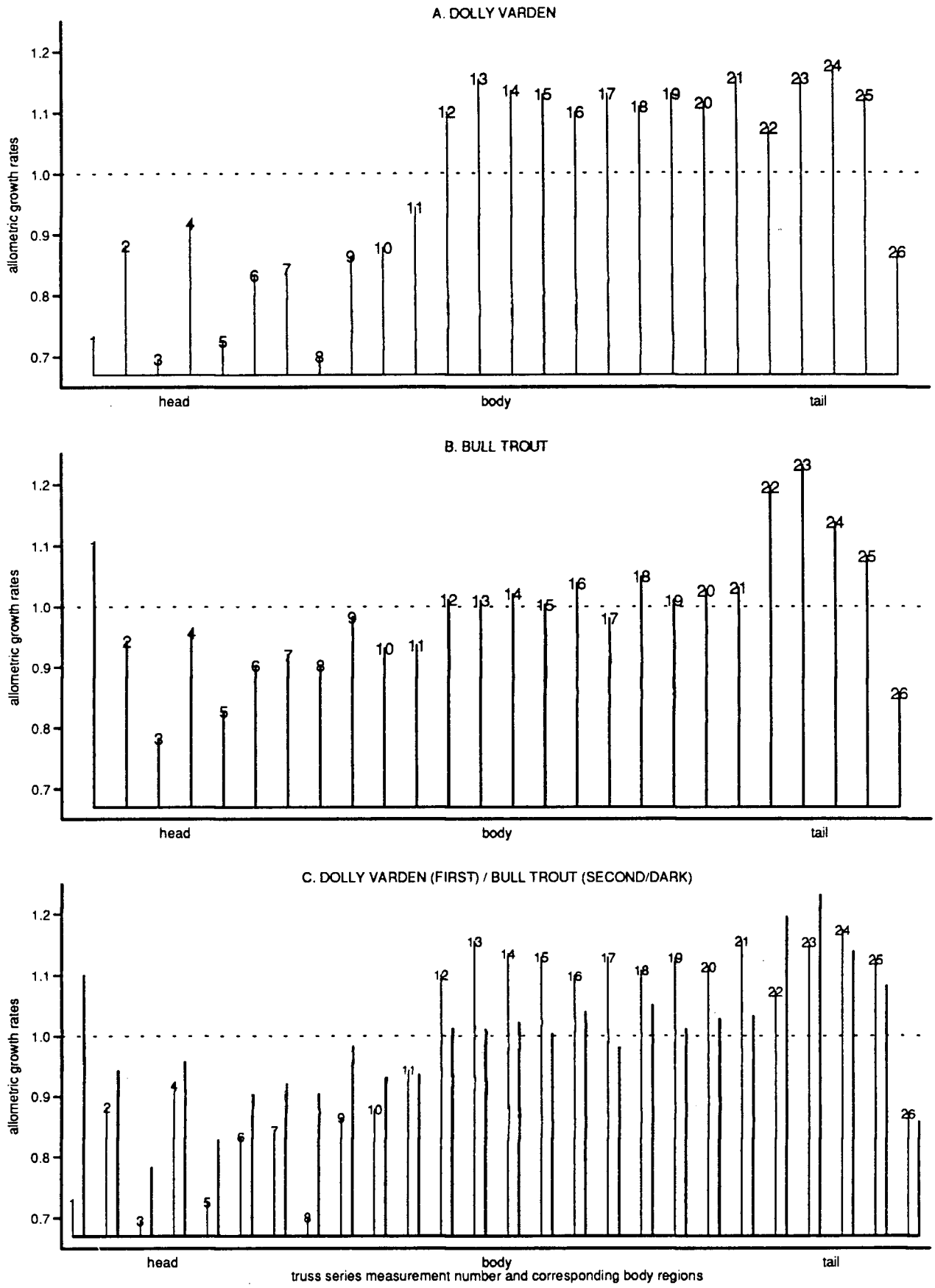


FIGURE 10. Ontogenetic growth rates for Dolly Varden and bull trout.

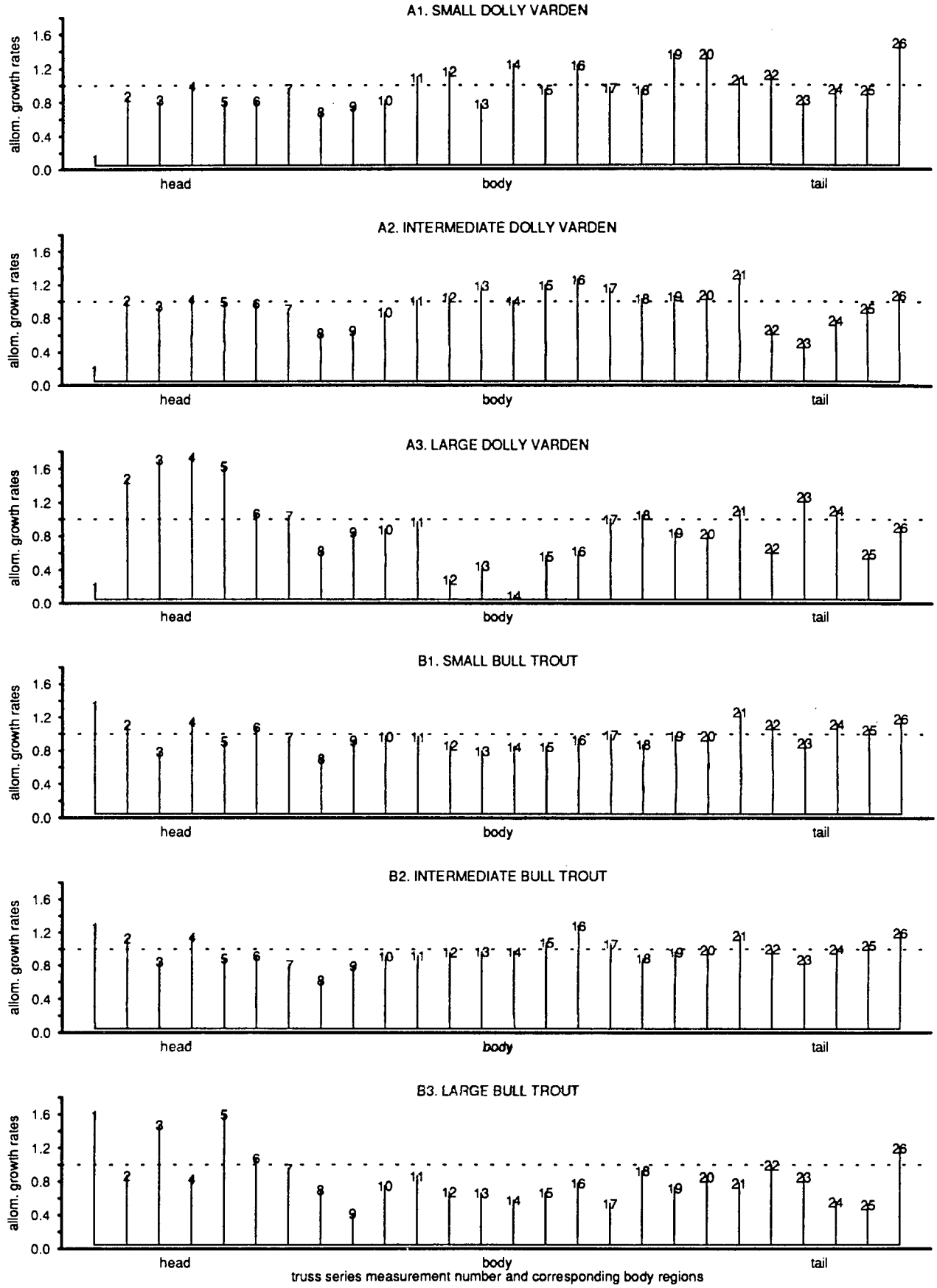


FIGURE 11. Static growth rates for each intraspecific size class.

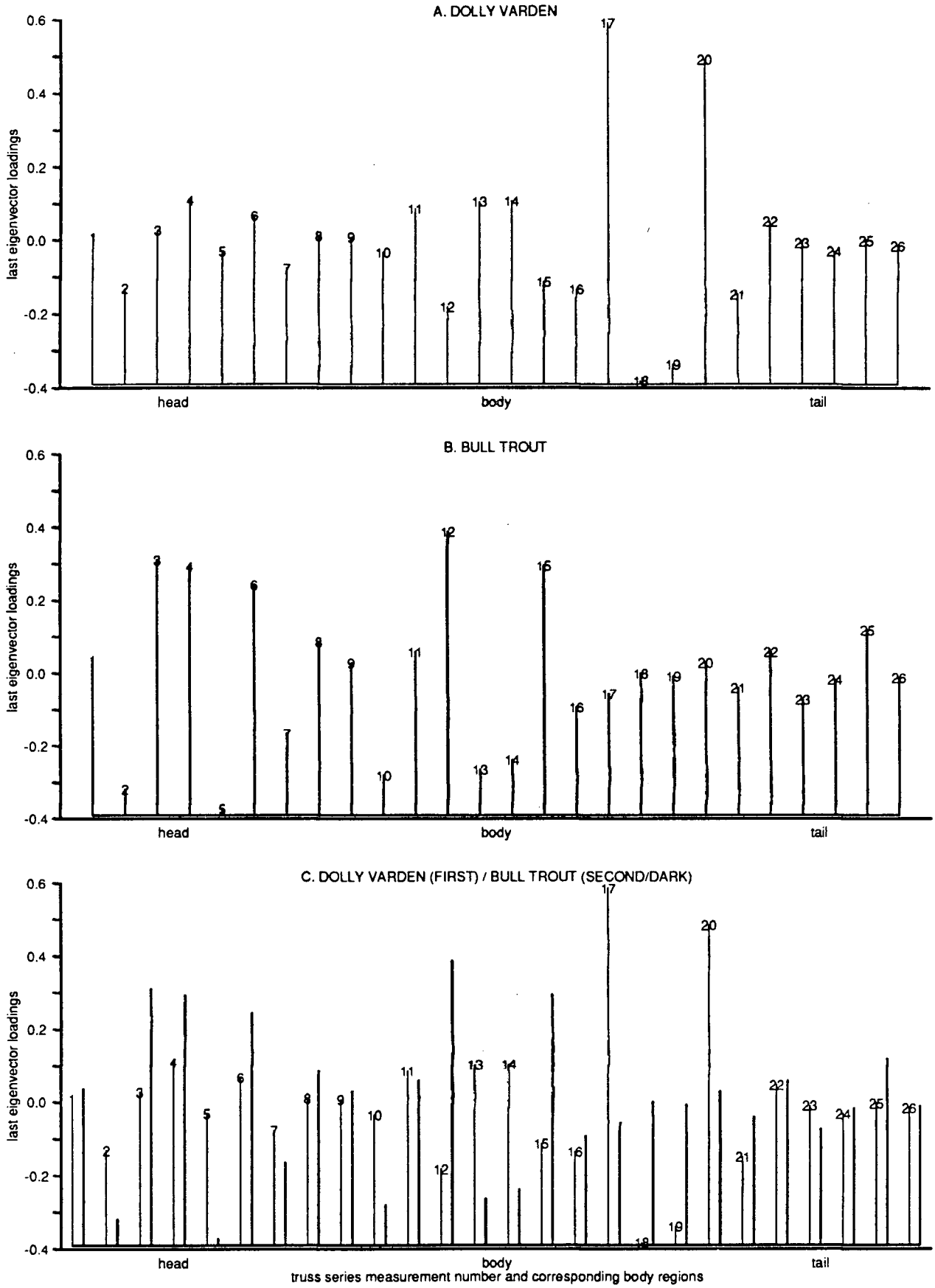


FIGURE 12. Developmental canalization (last eigenvector loadings).

An examination of the ontogenetic allometric growth rates in figure 10 further corroborates this hypothesis (for characters see fig. 1 in **chp. 1** and **appendix A**). The ontogenetic allometric growth rates for head morphology are higher in bull trout (see Hall 1982) whereas in Dolly Varden those for body morphology are higher. This is consistent with the overall morphological differences between these two species. Bull trout are distinguished by having larger, broader and flatter heads with more slender and ventrally flattened bodies than Dolly Varden (see **chp. 1**; figs. 2–3). Dolly Varden have heads which do not dominate their body profile and their bodies are more oval and "snake-like". This Dolly Varden shape morphometry is like that of juvenile bull trout. Also of interest is that Dolly Varden meristic characters are reduced in mean number and trend in comparison to bull trout. This meristic relationship is consistent with neotenic paedomorphosis.

This ontogenetic growth rate analysis is sensitive enough to detect the body morphology differences that suggest bull trout are more slender and flattened in appearance than Dolly Varden. The only body morphology characters (nos. 22–23) in bull trout that have higher allometric growth rates than Dolly Varden account for this difference since they increase the growth rates of the caudal region in bull trout and thus flatten out and decrease the tail and body profile (see fig. 1; **chp. 1**).

The fact that not all of the Dolly Varden characters have comparatively reduced growth rates (fig. 10) is not necessarily evidence against paedomorphosis (Creighton and Strauss 1986, Fink 1982, Kluge and Strauss 1985, Wayne 1986). Dolly Varden certainly appear paedomorphic with respect to bull trout (fig. 8) and the developmental changes which result in this paedomorphosis are a reflection of the growth rates of characters and not the growth rates of whole organisms. Some characters or character complexes might display one form of heterochrony while others will demonstrate some other type of heterochrony or perhaps none at all. This subtlety is often lost in the allometric framework used here and elsewhere where ontogeny is necessarily perceived as various phenomena. Paedomorphosis is really a term for the outcome of certain processes rather than a process itself.

The overall similarity of the allometry curves (fig. 8) and the static and ontogenetic allometric growth rates (figs. 10–11) of each species suggests that their development may be highly canalized (Alberch et al. 1979, Alberch 1982, Lerner 1954, Maynard Smith et al. 1985, Waddington 1962, Wayne 1986). This suggestion is strengthened by the high overall integration values calculated for

each species correlation matrix (both > 0.8), by an analysis of the smallest (growth invariant) PCA eigenvectors (fig. 12), and by the tight correspondence of the data to isometry (fig. 10) and to the PCA scatter plot curves (fig. 8). The bull trout characters are nearly all closer to isometry and load more heavily onto the last eigenvector than those for the Dolly Varden. This indicates that the bull trout are developmentally more strongly canalized. This indication is supported by the observation that bull trout apparently have reduced morphometric variability (fig. 23, see **chp. 6**; figs. 6b–7b, see **chp. 2**).

The shape differences that distinguish Dolly Varden and bull trout are not the same in all size classes. There appears to be "paedomorphosis" within each species as well (fig. 8; also see fig. 23 in chp. 6). The static allometric growth rates (fig. 11) for the largest size class of both species also are somewhat different and suggestive of paedomorphosis. Such intraspecific ontogenetic changes are probably not related to phylogeny but more likely are related to life-history (Creighton and Strauss 1986, Fink 1982, Strauss and Fuiman 1985). A good example of this is the Arctic char which can attain several discrete levels of morphometric variation within a single individual's ontogeny in response to ecological factors (Nordeng 1983). What the intraspecific differentiation means in my case is unknown, but Gould (1977) and others (Alberch et al. 1979, Balon 1979, 1980e, 1981, 1983, 1984) have suggested that paedomorphosis may be related to selection for competitive ability (ie. K-selection; Pianka 1970, Stearns 1976, 1977).

Summary and Conclusions

A multivariate morphometric analysis of ontogeny using cross-sectional data seems to be effective, realistic and simple. Multivariate size appears to be a practical surrogate for chronological age and permits the calculation of multivariate allometric shape curves and idealized ontogenetic trajectories that can be fitted as ontogenetic growth indicators. The ontogenetic mechanisms thus established are also supported by the allometric growth rates and developmental canalization of individual characters.

Dolly Varden appear to have evolved from bull trout through neotenic paedomorphosis. Their strongly canalized morphometric differences are likely the result of simple changes in relative

growth rates and timing. Allometric shape changes in ontogeny translate into the morphometric differences between Dolly Varden and bull trout.

PART II — Morphometric Statistics

CHAPTER FOUR

Data Attributes and Statistical Requirements for Morphometric Studies

Introduction

Morphometrics is the quantitative study of phenotypic variation, and attempts to describe the phenotype in terms of size and shape features. Size deals with absolute magnitude and growth, while shape describes general form. These two parameters are usually confounded due to allometry (Huxley and Tessier 1936). This means that shape change is size-related and thus the comparison of characters from individuals of different sizes and the separation of size and shape information from them is difficult. Since shape information is usually a much more reliable and significant indicator of relationships (Corruccini 1973, Jolicoeur and Mosimann 1960, May 1969, Steyskal 1968, Werner 1971, Wiley 1981), most morphometric procedures attempt to obtain unconfounded measurements representing shape. Size is still of interest though (Bonner 1965, Calder 1963, McMahon and Bonner 1983, Peters 1983, Platt and Silvert 1981, Schmidt-Nielsen 1984) and thus morphometric techniques should provide good size estimates and not simply remove it from the data.

Morphometric studies all share certain data and statistical requirements. These requirements generally are poorly understood and rarely assessed. This appears to be due to a lack of appreciation for the analytical problems that can result from not approximating these underlying assumptions. The availability of simple tests for these requirements is also apparently not realized or widely-known.

The comparisons of morphometric procedures in **part II** are all based on the same data set. Consequently, it is introduced and justified here to demonstrate its characteristics and utility for these analyses. This data set is therefore also tested for all the general data and statistical requirements and thus serves as an example of how to treat data prior to morphometric analyses. General recommendations and warnings concerning the type of data that meet the statistical and study requirements, and how to simply test for them, are presented. This should establish the background for the remaining chapters in **part I** and for morphometric studies in general.

This first chapter in **part II** is sequentially arranged under the following major headings: background; statistical assumptions; character selection; data transformation; data pooling; and summary.

Background

Methods

Most of the analyses and all the graphics for **part II** are based on computer macros written within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Centre, University of British Columbia (U.B.C.). These programs are available from me. The two exceptions are the 2-way ANOVA and the 2-way MANOVA which were respectively run on the pc-SAS (SAS 1985) and mainframe SAS (SAS 1982) statistical packages. The mainframe SAS was run on the MTS operating system (MTS 1976) at the Computing Centre at U.B.C..

Data Set

The data set for **part II** consists of fifty-one morphometric (continuous) and ten meristic (discontinuous) characters taken from sixty fish. The sixty individuals represent two closely related hypothetical groups (see **chp. 1**). Each group consists of thirty fish. The morphometric and meristic variables are analyzed independently to avoid problems with mixed character data sets (Pimentel 1979, Thorpe 1983a) and so that any statistical effects on the two classes can be clearly seen. Since all the results in this chapter are the same for both morphometric and meristic characters, they are jointly discussed here.

Group one is composed of fifteen males and fifteen females, and group two of eleven males and nineteen females. Two-way ANOVA (Corruccini 1987, Thorpe 1976, 1980) and 2-way MANOVA (Lande and Arnold 1983, Neff and Marcus 1980, Thorpe 1976, Willig et al. 1986, Willig and Owen 1987) suggest that there is no significant univariate or multivariate sexual dimorphism in either group (p at least > 0.1). This lack of sexual dimorphism is also demonstrated in figure 13(i). This figure represents a principal component analysis (PCA) (Hotelling 1933) scatter plot of the sixty

individuals with their sexes depicted by different symbols (for further PCA explanation see **chp. 6**). It reveals no clear differences between males and females.

Utility of the Data Set

There are three types of allometric data (Cock 1966, Gould 1966, Leamy and Bradley 1982). Cross-sectional data involves different individuals at different sizes or ages. Longitudinal data is from the same individuals but at different sizes or ages. Static data comes from specimens that are all in one size or age class. Only the first two types give real ontogenetic information and the use of static data can be misleading (Atchley and Rutledge 1980, Cheverud 1982b, Gould 1971, Lande 1979, Shea 1985, White and Gould 1965). Morphometric studies based on one size or age group, or on adults in organisms with determinate growth, should be cautiously interpreted (Bonner 1965, Gould 1971, Mosimann and James 1979).

The data in this study are cross-sectional. Furthermore, these data are realistically allometric since they represent fish that have indeterminate growth, and in this case range greatly in size (8.5–49.3 cm) (Humphries et al. 1981). The size range present in the two putative groups is completely overlapping (Claytor and MacCrimmon 1987) and their mean sizes are almost the same. Consequently, the differences between these two groups are not likely based on size alone.

Morphometric statistical procedures work best and are simplest with two closely-related groups (Blackith and Reyment 1971, Corruccini 1973, 1975, McKay and Campbell 1982a, Pimentel 1979, Siegel and Benson 1982, Thorpe 1980, 1983). In particular, the statistical assumptions underlying the tests are more easily met with such two group data. This data further meet these criteria because the two groups are equally represented and contain almost equal numbers of both sexes.

A complete set of counts and measurements was made on all sixty specimens and thus there are no missing data values. While the specimens were collected and then randomly selected from these collections, unusual fish were removed so outliers would not affect the analyses. Outliers are also easily detected with multivariate statistical techniques such as PCA (Everitt 1978, Gnanadesikan and Kettenring 1972, Rohwer and Kilgore 1972) and no outliers are found in this data.

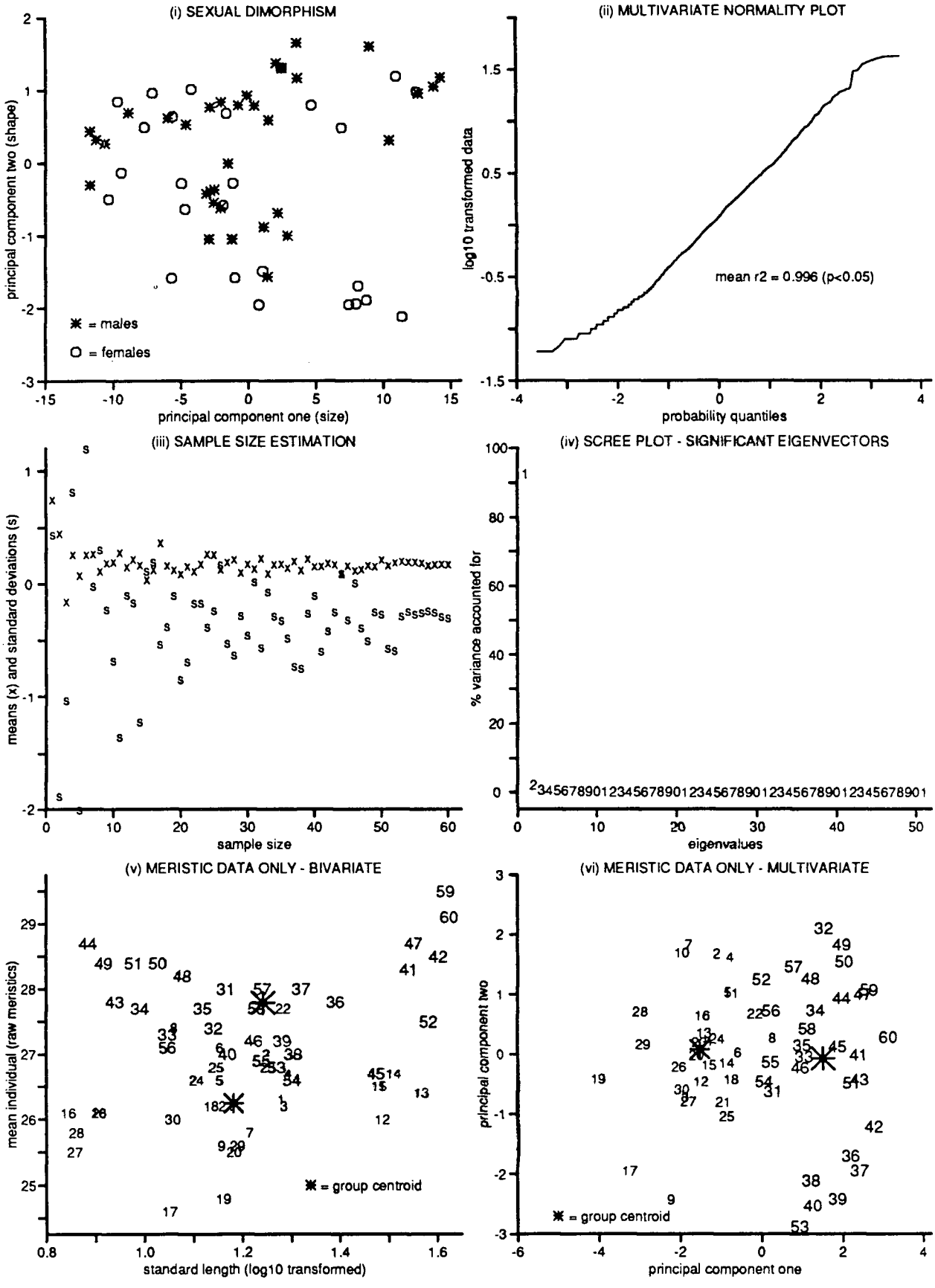


FIGURE 13. Representative plots of various data features.

The results of morphometric procedures are very data specific (Brown and Davies 1972). The data and its variances and intercorrelations have a dramatic effect on the outcome of these analyses. Therefore, this large number of characters was deliberately selected to provide balance. They help stabilize the analyses and strengthen the comparisons.

Jackknife analysis (Bissell and Ferguson 1975, Miller 1974, Mosteller and Tukey 1977, Quenouille 1949) is designed to test whether a subset of the data will produce different results (Claytor and MacCrimmon 1987, Gibson et al. 1984, McKay and Campbell 1982a, Neff and Marcus 1980, Pimentel 1981, Schaafsma and van Vark 1979, Srivastava and Carter 1983). My jackknife results are almost identical to the total data set for all the characters and data subsets analyzed. This suggests that this data set is robust and suitable for evaluating and comparing morphometric techniques. Other studies have demonstrated that if the total character number is sufficient then the ommission or inclusion of specific variables has little effect on the outcome of morphometric analyses (Bigelow and Reimer 1954, Boratynski and Davies 1971, Joliffe 1972, 1973, Thorpe 1976, 1985a–c).

Statistical Assumptions

Morphometric procedures all share certain statistical assumptions. Multivariate methods also have some additional assumptions implicit in their total matrix approach. All these requirements must be at least approximated (Lande and Arnold 1983) if the techniques are to be reliable and interpretable. Erring a small amount, however, is usually better than doing nothing (Pimentel 1979, Tukey 1962), especially when that small amount and its possible effects are known and understood. Ignoring all the assumptions or making large violations of them is unacceptable though.

Multivariate procedures are sometimes deemed free of assumptions if they are used only for descriptive purposes (see **chp. 6**). Such an application is rare in morphometrics. Meeting statistical criteria improves confidence in the results, regardless of the intended purpose of the analysis. In addition, the tests of the assumptions presented here are simple and provide useful insights into the data and all the analyses.

Normality

The major statistical assumption in morphometric analyses is that the data are normally distributed. Bivariate procedures work best with univariate normality, and multivariate techniques

require that each measurement and all its linear combinations be normally distributed. Univariate normality does not necessarily imply multivariate normality (Andrews et al. 1973, Neff and Marcus 1980, Pimentel 1979), and, indeed, multivariate non-normality has been demonstrated where univariate normality existed (Reyment 1971). However, univariate normality tests should at least be used for multivariate statistics if multivariate normality tests are unavailable.

The best univariate normality test (Chen 1971, D'Agostino and Pearson 1973, Mardia 1975, Shapiro et al. 1968, Zar 1984) is the Shapiro-Wilk statistic (Shapiro and Wilk 1965; also see Shapiro and Francia 1972). Unfortunately, it is not readily available on its own because it is hard to program and is difficult to calculate for large samples. It is, however, an incidental part of the output of some multivariate statistical package routines.

The most common univariate normality test is the Kolmogorov-Smirnov. This test is insensitive to departures from normality and this problem is magnified in the multivariate situation (Andrews et al. 1973). This insensitivity was noticed in this study, and the rarely used normality test based on probability (quantile) plot correlation coefficients (Filliben 1975, Ryan et al. 1976, Ryan and Joiner 1971) was therefore employed. This correlation coefficient test is simple and provides a close approximation of the Shapiro-Wilk statistic (Filliben 1975, Ryan and Joiner 1971). A table for its use is available (Filliben 1975, Ryan and Joiner 1971), and it is part of the Minitab statistical program (Ryan et al. 1976). Figure 14 (**chp. 5**) gives the results of this univariate normality analysis and is further explained in **chapter 5.**

Multivariate normality is almost never directly assessed because the tests are apparently difficult. The literature, however, suggests the use of probability plots of the data matrix as a test for multivariate normality as well (Andrews et al. 1973, Campbell 1980, Cox 1968, Everitt 1978, Gabriel 1985, Gnanadesikan 1977, Healey 1968, Hills 1969, Kimball 1960, Srivastava and Carter 1983, Wilk and Gnanadesikan 1968, Wilk and Shapiro 1968, Wilk et al. 1962). Like its univariate counterpart, this test is rarely used but simple and effective. It is suitable at least for a subjective visual assessment, and this could be supplemented by employing the univariate correlation coefficient test on the multivariate probability plot correlation coefficient. It is perhaps less likely that statistically significant multivariate normality will be obtained through this univariate correlation coefficient test though as it is probably too sensitive for the multivariate case. Nevertheless, a probability

plot resembling a straight line certainly approximates normality. At the very least, these plots will identify multivariate data that are very non-normal. Figure 13(ii) shows that my data plot as a straight line and so probably are multivariately normal (even significantly so based on the univariate correlation coefficient test).

Multivariate normality can also be subjectively assessed through scatter plots of the principal components (PC's) resulting from the multivariate analyses. PC plots tend to have an ellipsoidal distribution if the data they are based on is multivariately normally distributed (Reyment et al. 1984, Thorpe 1976). These plots again suggest that my data are multivariately normal (fig. 23 in **chp. 6**). There are other tests for multivariate normality (Anderson 1971, Andrews et al. 1973, Campbell 1980, Cox and Small 1978, Day 1969, Mardia 1970, Smith and Spiegelhalter 1981, Wagle 1968), but these are indeed complex and unnecessary in the light of these other simpler and equally effective tests.

Linearity

Linearity is another statistical assumption influencing morphometric statistics (Neff and Marcus 1980, Pimentel 1979). It is strongly related to normality and if a data set is linear it is probably normal. Therefore, linearity can also be used as a test for normality. Linearity can be best assessed through pair-wise variable plots (Chambers et al. 1983, Jöreskog et al. 1976, Leamy and Bradley 1982, Neff and Marcus 1980, Reist 1985), but histograms (Eickwort 1969) can be used as well. A statistical test for linearity is a runs test (Leamy and Bradley 1982), but the linearity plotting procedures permit better data appreciation. Pair-wise variable plots and the normality statistics indicate that these data are linear.

Homoscedasticity

Another important statistical assumption for morphometric studies based on multiple groups or both sexes is homoscedasticity (homogeneity of group variances or dispersions). The assessment of distinct entities in a procedure that does not account for within-group relationships (eg. regression, PCA) requires that the groups be homoscedastic. Box's (1954) modification of Bartlett's test (Pimentel 1979) is used to assess homoscedasticity. It is extremely conservative, however, and often gives negative results even when homoscedasticity is present (Phillips et al. 1973, Somers 1986).

Homoscedasticity will not be found if the data are not normally distributed either (Neff and Marcus 1980, Pimentel 1979, Reyment et al. 1984, Van Valen 1978).

At any rate, some heteroscedasticity does not seem to have much of an effect on single-group procedures (Ito 1969, Ito and Schull 1964, Phillips et al. 1973), and many of the tests are quite robust to it (Pimentel 1979). In short, single group procedures are permissible if a data set is homoscedastic. If a data set is not homoscedastic, consider other aspects of the data before using different possibly inappropriate analyses or pooling the data for homoscedastic groups. These data are homoscedastic ($p > 0.5$; Box's test) for both groups and sexes. See the data pooling section in this chapter for some further discussion.

Matrix Singularity

Matrix singularity (determinant $\neq 0$) is the final statistical assumption and it pertains only to multivariate analyses done on certain data sets. Singularity is rarely a problem except with large numbers of characters because these variables could then be completely linearly dependent and redundant. If a data set contains far fewer variables than individuals, and if multivariate analyses within statistical packages run smoothly, then singularity is not a problem. Most standard multivariate statistical programs will not analyze singular matrices.

This data set could be singular because it is deliberately set up with many variables to provide clear analytic comparisons. It is not singular though, as the matrix is at least positive semi-definite if not positive definite, matrix inversion is possible and no negative eigenvalues were encountered in the multivariate analyses which ran smoothly (Pimentel 1979, Somers 1986). I looked for negative instead of zero eigenvalues because fifty-one characters result in fifty-one eigenvalues the latter of which consequently come close to zero and thus are hard to discern (for multivariate terminology see **chp. 6**). The number of negative eigenvalues obtained in an analysis can be useful as their number reveals the number of ipsative measures present in the data set (Pimentel 1979). Ipsative measures are those which sum to a constant or to another measure. Unfortunately, it does not show which ones they are, but this can often be deduced through logical screening of the characters.

Character Selection

While the number of characters used in an analysis is often determined by study economics certain specific features of them are important. Foremost, the selected characters must adequately describe the organism (Bookstein et al. 1985, Strauss and Bookstein 1982, Thorpe 1976, 1980) and meet the statistical requirements. At the same time, too many characters can be redundant (Corruccini 1975, Crovello 1970, Power 1971, Reist 1985, Rohlf 1967) and ipsative measures (Neff and Marcus 1980, Pimentel 1979, Sacher 1970) must be avoided. Ipsative measures should be removed by logical screening, through matrix singularity checks and by jackknifing. These checks revealed no ipsative measures in this data set.

Characters with high, relatively uniform intercorrelations are best for effective size/shape separation (Campbell 1976, Corruccini 1983, Somers 1986) and are recommended for statistical technique comparisons (Corruccini 1983). The variables should especially be correlated with size (Reist 1985, Somers 1986) because otherwise the removal of nonexistent size can result in false shape variables. This latter effect is evident in some of my meristic analyses (see **chps. 5–6**). As is usually the case, my morphometric variables are strongly positively correlated both with each other and with size, but my meristic characters are not.

High character intercorrelation can result in a problem for multivariate statistics known as Rao's paradox (Corruccini 1987, Healy 1969, Kowalski 1972, Rao 1966a, Willig and Owen 1987). In this situation, the multivariate analyses overcorrect for size differences between groups if their shapes are nearly identical. This brings their sample centroids closer together than they would be in a univariate state. My data has more statistically significant group differences based on MANOVA than on ANOVA so this situation does not exist. I also get much better group separation in the multivariate case than in the univariate case (figs. 16 and 23; **chps. 5–6**).

Some authors advocate using only characters that are significantly different between groups (Newman and Jancey 1983, Thorpe 1975a, 1975b), but this practise does not allow for effective phenotypic description and is biased towards group separation (Thorpe 1976). Such an approach should only be used in discrimination and not in morphometric description.

Characters also should be easy to measure, well-defined and repeatable (Corruccini 1978, Croy and Dix 1984, Sjøvold 1975, Thomas 1968). In addition, common variables are good (Croy and Dix 1984) since they allow for comparison, especially to earlier work. Most traditional measures are linear and may be insufficient for shape analyses. Therefore, measurements done in all three dimensions are excellent adjuncts to the customary ones. They add more shape information (Reyment et al. 1984) and do not preclude technological advances (Bookstein et al. 1985, McGlade and Boulding 1986, Strauss and Bookstein 1982; also see **chps. 1–3**). Characters which permit biological and theoretical interpretation are particularly encouraged (Pimentel 1979).

Measurement Error

Measurement error is rarely assessed yet it could greatly affect analyses (Winans 1984). Ten randomly chosen specimens (five from each group) were remeasrued here and measurement error was negligible. This was evaluated by looking at the mean amount of measurement error (Winans 1984) and also through a one-way ANOVA (Baumgartner et al. 1988). With the ANOVA, repeatability is calculated as the ratio of among individual variation to total (among plus within) variation (Falconer 1981). If this ratio is large (close to one) then the repeatability is high and measurement error is insignificant. In this data set, each group has similar and insignificant patterns of error so these factors did not affect the analyses either.

Sample Size

While sample size is partly determined by study economics, an estimate of the sample size required is easily obtained and greatly increases confidence in the results. The simplest univariate technique to estimate sample size is to randomly select and plot the cumulative individual specimen means and associated standard deviations or standard errors against the sample size (fig. 13(iii)). A sample size is chosen where the plotted curve stabilizes and asymptotes. When tested this way, these data asymptote before thirty and thus their univariate sample size is adequate in each group and for the total data matrix. A suggested univariate rule-of-thumb minimum sample size for morphometrics is twenty-five individuals (Neff and Marcus 1980, Mardia 1971, Reist 1985) and is supported by this study.

The simplest multivariate technique to estimate sample size is to randomly choose and plot the cumulative determinants of a correlation matrix and their standard errors against sample size (Scagel et al. 1985; also see Cheverud et al. 1983). Again, an adequate sample size is indicated by where the curve stabilizes and asymptotes. My multivariate estimates are almost identical to the univariate estimates and further suggest this data set is appropriate and sufficient. This robustness of sample size is also demonstrated by my jackknife analyses which produce similar results based on subsets of the data.

Because of this similarity, only the univariate plot is shown. It also is more readily understood and does not require logarithmic axis transformation for plotting (Scagel et al. 1985). If different multivariate sample estimates are obtained and multivariate statistics are to be employed determinant plots should be used. A morphometric multivariate rule-of-thumb for sample size is that it should be greater than the number of characters used in the analysis (Orloci 1967, Pimentel 1979). Many standard computer programs will not run if there are more characters than individuals in the sample.

Other sample size estimators exist (Cochran 1977, Croy and Dix 1984, Falconer 1981, Green 1979, Newman and Jancey 1981, Odeh and Fox 1975, Sokal and Rohlf 1969), but they are not as intuitive or simple as the techniques described above.

Data Transformation

In this study, logarithmic (base 10) data transformation is used for all the multivariate statistics and for specified bivariate statistics. This results in my data conforming better to the statistical assumptions. In fact, $\log_{10}$ transformation is necessary in this case to obtain linearity and normality.

Logarithmic transformation usually improves both univariate (Kermack and Haldane 1950) and multivariate normality (Gower 1972, Jöreskog et al. 1976, Pimentel 1979, Smith 1980), helps the data approximate homoscedasticity (Gower 1972, Smith 1980, Thorpe 1976), stabilizes variances (Jolicoeur 1963a, Reyment and Banfield 1976, Ricker 1973, Thorpe 1976), helps make results independent of scale and magnitude (Jolicoeur 1963a, 1963b, Humphries et al. 1981, Reyment and Banfield 1976, Smith 1980), reduces outlier problems (Jöreskog et al. 1976, Smith 1980), preserves

allometries (Humphries et al. 1981) and is necessary for the calculation of allometry coefficients (see **chps. 5–6**). Logarithmic transformation also improves linearity in most data sets (Kuhry and Marcus 1977, Pimentel 1979, Ricker 1973, Smith 1980, Thompson 1942, Thorpe and Leamy 1983).

Logarithmic transformation of morphometric data is almost always recommended (Bryant 1986, Burnaby 1966, Bookstein et al. 1985, Harvey 1982, Kermack and Haldane 1950, Marriott 1974, Sacher 1970, Sokal 1965, Shea 1985, Thorpe 1983). There is no real alternative transformation available for morphometric data, but a square-root transform is sometimes suggested for meristic data (Jöreskog et al. 1976, Pimentel 1979). This is especially true if the meristics are in low numbers (Sokal 1965) or if they follow a Poisson (random) distribution (Marriott 1974). Since square-root transformation of this meristic data produces results similar to the logarithmic transformation, logarithmic transformation is used to maintain consistency with the morphometric measurements. If weight is to be used as a variable it will probably require cube-root tranformation in order to render it more dimensionally equivalent (Gould 1971, Leamy and Bradley 1982).

Data Pooling

Pooling data within-groups (Pimentel 1979, Reist 1986, Shea 1985, Somers 1986, Thorpe 1975a, 1976, Thorpe and Leamy 1983) is often used when heteroscedastic groups occur in the data set. The term "groups" could represent statistically distinct entities in a multiple group analysis or statistical differences between sexes within a single group. In this study, groups represent distinct entities in a multiple group analysis as there is no statistically detectable sexual dimorphism or heteroscedasticity in this data set.

Other authors suggest pooled within-group data analyses so that the differences in group data structure, particularly variances, are accounted for. This, however, necessitates *a priori* group assignment which is subjective and assumes that the groups are both real and completely distinguishable. This is a problem if unknown individuals, groups, hybrids or introgressed individuals exist in the sample. The most variable group may also dominate the analysis, especially in multivariate procedures such as PCA (Pimentel 1979, Somers 1986). Such a multigroup PCA (Pimentel 1979) can result in correlated size and shape vectors (non-orthogonal eigenvectors) caused by the influence of grouping on the standardization of the original data matrix (Bookstein et al. 1985,

Burnaby 1966, Humphries et al. 1981, Rohlf and Bookstein 1987, Somers pers. comm.) Pooling within-groups may also increase the importance of discriminating characters (Rohwer and Kilgore 1972), which is desirable in discrimination but not in morphometric description or size/shape adjustment. These could all potentially be bigger problems than the use of a total data matrix, especially in multivariate statistics.

If specimens from different populations are combined in the analyses, the rationale for pooling within-groups is further weakened (Baker et al. 1972, Hiernaux 1972, Sokal 1965, Thorpe 1976, 1980). Such compound localities are often necessary, but the within-group pooling should then be based on the populations which form the compound localities and not just on the groups themselves. This data set is composed of compound localities.

Since the groups in this data set are homoscedastic, the use of pooled within-group data is unnecessary. However, the following tests can act as further checks on pooling within-groups for analyses based on data that the conservative Box's test deems slightly heteroscedastic. Pooled within-group data, or single group data, made no differences in my regression or PCA. Group regression slopes and intercepts are not significantly different (Clarke 1980, Claytor and MacCrimmon 1987, Reist 1986, Zar 1984) in either case and the PCA plots are much alike (Shea 1985). The sheared PCA procedure (Humphries et al. 1981; see **chp. 6**) does take the group sizes into account and the results are almost identical to PCA's based on the total data matrix (fig. 23). Furthermore, the use of another multivariate procedure which accounts for groups (linear discriminant function analysis (Fisher 1936)) produces virtually identical results to the PCA. The jackknife tests removed certain specimens from each group and then reanalyzed the results (Gibson et al. 1984). This test also revealed no differences. The effects on the actual data values are also minimal and consistent in all these alternative within-group pooling procedures. A final reason to not pool data within-groups here is analytic consistency, since it cannot be done for ratios and is not yet part of the available size-constrained PCA procedure (Somers 1986, pers. comm.).

These reasons, and the similar results obtained from all these analyses, further suggest that pooling within-groups is not necessary in this study. In addition, some heteroscedasticity does not have much of an effect on single-group procedures (Ito 1969, Ito and Schull 1964, Phillips et al. 1973) and many of the methods are quite robust to it (Pimentel 1979). Somers (1986) mentions pooling in

regards to PCA on heteroscedastic sexes and then ignores it for the aformentioned reasons as well. Campbell (1976) and Shea (1985) say the overall relationships and interpretation of total and pooled within-group data are usually the same. This is especially true with closely related groups that overlap in size (Claytor and MacCrimmon 1987; also see Mosimann and James 1979). Reist (1986) discusses various poolings in regression analyses and finds significantly different results. However, his data set is composed of compound localities and his delineation of groupings is by cluster analysis in which the groups are designated *a priori*. Therefore, their natural discrimination is not assessed. Furthermore, the overall patterns and relationships for his characters are identical to those in the total data analyses and only individual characters display subtle, yet similar, changes. The final interpretation would remain the same. Reist (1986) still recommends that pooling within-groups only be done *a posteriori* and where necessary.

The best way to deal with pooling data within-groups is to be aware of the problem and then test for it if necessary. Initially, within-group pooling should not be done in compound samples or in difficult and unknown closely-related taxonomic groups with overlapping size ranges, unless it is absolutely necessary. Pooling within-groups should probably be undertaken if the samples are based on well-known populations, on good previous studies, on desired groups, on groups with non-overlapping size ranges, or on static data. However, even here check to ensure that is necessary.

Within-group samples are easier to pool for bivariate techniques such as regression and it may be more valuable to do so here as well (Kuhry and Marcus 1977, Reist 1986, Thorpe 1975a–b). It is more complicated, and often less beneficial, to pool in single group multivariate procedures such PCA. When pooling within-groups, always keep its effects in mind. Often neither the total nor pooled within-group matrices are ideal, so as a rule-of-thumb do what is simplest and most appropriate.

Summary

This is a summary of general morphometric data characteristics and assumptions and not of my study data.

1. There are three types of allometric data of which only the first two provide real ontogenetic and allometric information:

a) Cross-sectional data;

b) Longitudinal data;

c) Static data.

2. Characters should be adequate to describe the organism and to meet the study objectives. They should have relatively high and uniform intercorrelations, and be easy to measure, well-defined, repeatable, common and practical.

3. Characters should not be ipsative, redundant or have statistically significant measurement error.

   test: Logical screening, matrix singularity checks and jackknifing for ipsative or redundant characters.

   test: Mean error and ANOVA for significant measurement errors.

4. Is statistically significant sexual dimorphism present?

   test: MANOVA (or ANOVA).

      absent: Analyze the sexes together.

      present: Analyze the sexes separately in a multiple group analysis. In a single group analysis, test for homoscedasticity and see summmary discussion of it below to decide what to do. Think of the sexes as groups in a single group analysis.

5. Sample size is usually determined by study economics but the actual appropriate sample sizes can be estimated.

   test: For univariate or bivariate analyses, plot the cumulative individual specimen means and associated standard errors against sample size. Rule-of-thumb is twenty-five individuals.

   test: For multivariate analyses, plot the cumulative determinants of a correlation matrix and their standard errors against sample size. Rule-of-thumb is that the sample size be greater than the variable number.

The following statistical assumptions should be at least approximated.

note: Logarithmic transformation of morphometric data is almost always recommended as it improves the data and helps meet the statistical assumptions. Weight is cube-root transformed, and a square-root transformation is sometimes suggested for meristic characters.

1. Univariate normality.

   test: Probability (quantile) plot correlation coefficient test.

2. Multivariate normality.

   test: Probability (quantile) plots — obtain an approximate straight line. Another indicator of multivariate normality is ellipsoidal PCA scatter plots.

3. Linearity.

   test: Pair-wise variable plots, and good normality statistics.

4. Matrix singularity. It is only a potential problem in multivariate studies where the character number approaches the sample size or the variables are ipsative. Otherwise, it can likely be ignored.

   test: Computer statistical programs run smoothly. Matrix determinant is greater than zero, matrix inversion is possible, and programs result in no zero or negative eigenvalues.

5. A final statistical assumption is homoscedasticity. Its consequences warrant a separate summation.

test: Box's test.

note: Homoscedasticity may be difficult to prove due to the very conservative nature of Box's test. However, some heteroscedasticity is often not a problem. Therefore, use this protocol:

i. Test for homoscedasticity.

   a) If present use the total data matrix.

   b) If absent consider the following points before pooling the data within-groups or using a procedure that accounts for groups.

ii. What are the research objectives? Would total or pooled within-group analyses be better?

iii. Other tests (since Box's test is conservative).

   a) Are the regression slopes and intercepts of each group statistically different? If so pool, but if not use the total data matrix.

   b) In PCA, does the shear matrix procedure, pooling within-groups or doing separate group analyses have any effect? If so pool or use the sheared matrix procedure, but if not use the total data matrix.

iv. Other considerations for not pooling within-groups:

   a) Compound samples.

   b) Difficult or unknown closely related individuals and groups.

   c) Groups have completely overlapping size ranges.

v. Deliberate considerations or reasons to pool within-groups.

   a) Specifically want to analyze particular *a priori* groups.

   b) Groups represent very well-known populations or samples.

   c) Groups have non-overlapping size ranges.

   d) Data used is static.

vi. Pool within-groups only if necessary or desired. If you do not specifically want to pool within-groups then initially do not do so. If you must pool within-groups keep the pooling effects in mind. Since neither approach is always ideal, do what is simplest and most appropriate.

CHAPTER FIVE

Assessment of Bivariate Morphometric Procedures

Introduction

Bivariate morphometric procedures independently adjust each character with a single, arbitrary measure of overall size. These adjusted characters provide the shape information and the single size measure gives the size information. The procedures do not account for correlations between characters and cannot directly statistically test their complete, overall affiliations.

The choice of size variable is critical (Hills 1982, Jungers and German 1981, Leamy and Bradley 1982, Mosimann 1970, Pimentel 1979) because it determines the shape aspect of each character and is the only size measure. Shape need not be related in the same way to different size variables, so the size measure chosen must be representative (Mosimann and James 1979) and correlate strongly to the other variables (see **chp. 4**). Standard length (Hubbs and Lagler 1958) is used here, but two other ichthyological size measures (total and caudal length) produced identical results (Baumgartner et al. 1988, Rohlf and Bookstein 1987). Weight is often used as an effective size measure.

A problem with bivariate procedures is that size has relevance and should not be computed only to separate it from shape (Bonner 1965, Bookstein et al. 1985, Pimentel 1979, Smith 1980). Also, since the size measurement in bivariate techniques is a single measure it does not provide separate size information for each of the other variables. Therefore, bivariate procedures assume that the underlying allometry is univariate and this may be incorrect. In addition, the size measurement is often linear and not of functional interest (Smith 1980), and this can affect any biological interpretations (Leamy and Bradley 1982).

While not as sophisticated as multivariate techniques, bivariate procedures are more easily understood and still frequently used (Corruccini 1975, Reist 1985). Older morphometric studies are based exclusively on these techniques and thus bivariate techniques should also be investigated for the sake of comparability. For these reasons bivariate techniques cannot be ignored and should not be abandoned (Corrucini 1983, Hatheway 1962, Holloway and Jardine 1968). Their relationships

and data effects must be understood, and their procedures are simpler and adequate for many objectives. Their understanding also helps interpret the more complex multivariate procedures and establishes the compatibility of the bivariate and multivariate morphometric techniques.

The only two types of bivariate morphometric procedures are ratios and regressions. There are several modifications of each of these techniques but only two ratio variants and two regression methods are common enough to be discussed here. The other modifications usually result in very similar output anyway (Atchley 1978, Corruccini 1975, 1977, Pimentel 1979, Reist 1985).

Ratios

Ratios are the oldest morphometric methods. Their statistical problems have also long been demonstrated (Pearson 1897, Simpson and Roe 1939) and these difficulties have since been expounded on. Ratios are, however, still used (Baltz and Moyle 1981, Mosimann and James 1979, Shaklee and Tamaru 1981, Wilk et al. 1980) both because of their simplicity and through ignorance concerning their problems (Atchley and Anderson 1978, Barraclough and Blackith 1962, Blackwelder 1964 , Burnaby 1966, Christensen 1954, Jeffers 1967, Middleton 1962, Pimentel 1979, Reyment et al. 1984).

The two major problems with ratios involve spurious correlations (Anderson and Lydic 1977a, 1977b, Atchley et al. 1976, Chayes 1949, Pearson 1897, Reist 1985, Schuessler 1974), and leptokurtic, skewed or Cauchy distributions (Albrecht 1978, Anderson and Lydic 1977a, 1977b, Atchley et al. 1976, Reist 1985, Thorington 1972). As well, ratios assume a linear relationship between the variables involved (Albrecht 1978, Hills 1978) and also assume that the axis describing this relationship passes through the origin (Thorpe 1983a). The relationship may be linear, especially with logarithmic transformed data, but it is rare that their axis intersects the origin (Thorpe 1983a). Indeed, for this data set, pair-wise plots of each of the characters on the size variable indicate linearity but the intercept was almost never near the origin (see **chp. 4**).

Other problems involving ratios are that they do not remove scaling effects (Anderson and Lydic 1977b, Atchley et al. 1976), they compound error terms (Reist 1985, Simpson et al. 1960), and they result in information that may be unpredictably due to either the numerator, denominator or both (Atchley et al. 1976, Croy and Dix 1984, Sokal 1965). Also, the use of ratios often obscures

data relationships, especially those of size and shape. (Anderson and Lydic 1977b, Dodson 1978, Humphries et al. 1981, Phillips 1983, Pimentel 1979, Reist 1985). Furthermore, it is claimed that ratios do not address an allometric hypothesis (Burnaby 1966, Dodson 1978) and thus should not be employed for allometric adjustments. Generally, ratios are only recommended for problems where the hypotheses tested deal directly with ratios (Blackith and Reyment 1971, Corruccini 1977, Kowalski 1972, Reyment et al. 1984).

Two ratio methods are examined. The first is the division of each character for each individual by that individual's size measure, and the second is this same quotient but it is $\log_{10}$ transformed. This latter $\log_{10}$ transformation is supposed to help with the linearity and scaling problems of untransformed ratios (Hills 1978, Reist 1985).

Untransformed Ratio Formula

The formula (Reist 1985, Shea 1985) used for calculating the untransformed ratios is:

$$\hat{y}_{ip} = y_{ip}/x_i$$

where:

$\hat{y}_{ip}$ = adjusted p^{th} character for the i^{th} individual;

y_{ip} = original unadjusted p^{th} character for the i^{th} individual;

x_i = size measure for the i^{th} individual.

Logarithmic Transformed Ratio Formula

The formula (Hills 1978, Reist 1985) used for calculating the $\log_{10}$ transformed ratios is:

$$\hat{y}_{ip} = log_{10}(y_{ip}/x_i)$$

where:

$\hat{y}_{ip}$ = adjusted p^{th} character for the i^{th} individual;

y_{ip} = original unadjusted p^{th} character for the i^{th} individual;

x_i = size measure for the i^{th} individual.

Regressions

Regression morphometric techniques were developed as an alternative to ratios (Huxley 1932, Thompson 1942) and are still considered the best bivariate procedures (Corruccini 1978, Gould 1966, Reist 1985, Schuessler 1974). They are related to the power function $y = ax^b$ (Huxley 1932, Snell 1891) which describes the exponential growth of each part of an organism (see **chp. 4**) and has both a cellular (Gerhart et al. 1982, Katz 1980, Laird 1965, Laird et al. 1965, 1968) and morphometric (Blackstone 1987a, Creighton and Strauss 1986, Strauss and Fuiman 1985) basis. Regression techniques thus better approximate real allometric hypotheses.

There are several types of regression, but the least-squares method (Draper and Smith 1981) is employed throughout this study because it is in general use, and is simple and readily available. Furthermore, if the character correlations are high and if the groups are closely related the other regression techniques provide almost identical results (Brown and Davies 1972, Cock 1966, Gould 1966, Leamy and Bradley 1982, Misra and Reeve 1964, Röhrs 1961, Siegel and Benson 1982). These data criteria are met here and usually are.

While keeping this in mind, reduced major axis regression often is nonetheless deemed preferable to least-squares regression on theoretical grounds (Clarke 1980, Gould 1966, Hayami and Matsukuma 1970, Imbrie 1956, Kermack and Haldane 1950, Ricker 1973, Sacher 1970, Tessier 1948), and other similar recommended alternatives are major axis regression (Claytor and MacCrimmon 1987, Kuhry and Marcus 1977), principal axis regression (Jolicoeur 1965, Kermack and Haldane 1950, Sacher 1970), Bartlett's method of regression (Bartlett 1949, Brown and Davies 1972, Kidwell and Chase 1967, Simpson et al. 1960, Sokal and Rohlf 1969; for contrary see Kuhry and Marcus 1977, Madansky 1959, Neyman and Scott 1951, Kuhry and Marcus 1977) and robust regression (Siegel and Benson 1982).

Most of these other regression techniques are trying to solve the problem that the size measure (x-variable) in least-squares regression is theoretically not independent because in morphometrics size is subject to measurement error (Claytor and MacCrimmon 1987, Kuhry and Marcus 1977, Sacher 1970). This dependence results in regression estimates being unpredictably biased downward (Cock 1966, Leamy and Bradley 1982; also see Manaster and Manaster 1975, Zar 1968).

However, least-squares regression on most morphometric data actually gives similar results to these alternatives. Furthermore, any measurement error is usually minimal and in this study it is statistically insignificant (see **chp. 4**). These alternative regression methods also are often not readily available and are less well understood. In addition, they have their own problems, and besides their increased complexity defeats the advantage of simplicity that bivariate morphometric procedures have over multivariate techniques (Sacher 1970). My least-squares regression analysis was verified (Gould 1966) with reduced major axis regression. As expected, least-squares and reduced major axis regressions gave virtually identical results.

Two regression techniques are looked at in this study. The first is based on $\log_{10}$ transformed data and uses the slopes derived from the regression to adjust the variables for each individual to an overall grand mean body size. The second uses raw data and the adjusted characters are taken as the residuals of this same but untransformed regression. Each residual is the measure of deviation of each character of each individual from the regression line.

Regression Formula

The actual formula for the first regression is presented in the next formula section on regression residuals. The formula used for calculating the mean regression data is (Claytor and MacCrimmon 1987, Reist 1985, Shea 1985, Thorpe 1975):

$$\hat{y}_{ip} = log_{10} y_{ip} - k_p(log_{10} x_i - log_{10} \bar{x})$$

where:

$10^{\hat{y}_{ip}}$ = adjusted p^{th} character for the i^{th} individual;

y_{ip} = unadjusted original p^{th} character for the i^{th} individual;

x_i = size measure of the i^{th} individual (standard length here);

$\bar{x}$ = grand mean of size measures (or an arbitrary comparative standard size);

k_p = allometry coefficient for p^{th} character (slope (b) of $\log_{10}$ regression).

Regression Residuals Formula

The formulas used for calculating the regression residual data and the regressions are (Claytor and MacCrimmon 1987, Reist 1985; also see Smith 1981):

$$\text{regression} : y_p = a + bx + e$$

$$\text{regression residual adjustment} : \hat{y}_{ip} = e_{ip}$$

where:

y_p = unadjusted p^{th} character for all individuals;

x = size measure for all individuals (standard length here);

a = regression intercept;

b = regression slope;

e = regression residuals;

$\hat{y}_{ip}$ = adjusted p^{th} character for the i^{th} specimen;

e_{ip} = residuals for the p^{th} characters of the i^{th} individuals.

Bivariate Morphometric Procedures — Assessment Methods

The bivariate results are presented in figures 14–17 and table 1. All the morphological and meristic characters are respectively numbered 1–51 and 1–10, and are independently represented on figures 14–15 by separate layouts and captions. The sixty individual fish are portrayed on figures 16–17 which are completely separate from the character representations. Group one is depicted by numbers 1–30 set in small type and group two by numbers 31–60 set in large type. Centroids (group means) for each of these groups are also plotted in small and large sizes. Each graph and table is for the raw data (labelled **a**), $\log_{10}$ transformed data (**b**), ratio data (**c**), $\log_{10}$ transformed ratio data (**d**), regression data (**e**) and regression residual data (**f**).

Figure 14 presents the central tendency statistics, and allows for a complete assessment of the effects of transformation and of each bivariate procedure on the data. The numbers plotted represent the character means. If the numbers are circled they are normally distributed at $p < 0.05$ (probability plot correlation coefficient test; see **chp. 4**). These numbers on figure 14 are always single digits to allow the normality circles to be drawn neatly around them. Since the x-axis is variable number it still allows for their exact interpretation. The size variable (no. 51) is the first character on the meristic portion of figure 14, and only this figure, because its range is in that region and this permitted better use of the limited space. Since it is presented there it is labelled in two digits as 51. The vertical lines for each number in figure 14 are the standard deviations of the characters. Since the regression residuals are such small numbers, these lines are not presented for **f** because their distributional pattern would be lost if their larger standard deviations are plotted. There is nothing unusual about the standard deviations for **f**.

Figure 15 represents a graph devised to portray and compare the bivariate and multivariate results. Since its style is new and used throughout this thesis, I refer to them as punk plots. This is due to their spiked appearance which is fascinating to look at and easy to pass judgment on (correct in this case). Both figures 16 and 17 are referred to as scatter plots.

Figure 15 presents the allometry coefficients which provide a relative measure of allometry and its direction for each character. The numbers are plotted on equivalent axes about an isometric value of one so that the individual and overall character allometric patterns from each technique can be seen. If the allometry coefficient is one then that character is isometric (shape change is exactly proportional to size change). If it is less than one then positive allometry is present, and if greater than one negative allometry is. The size of the allometry coefficients, greater than or less than one, indicate how strongly the characters are positively or negatively allometric. Since size often is not part of meristic characters (see fig. 17 discussion below), the meristic allometry coefficients may be somewhat unrealistic. Regardless, they still indicate how bivariate procedures can change data relationships from those already represented by the raw meristic allometry coefficients. The allometry coefficients also are the only directly comparable link to the multivariate procedures (Jolicoeur 1963a–b, Leamy and Bradley 1982, Shea 1985).

The allometry coefficients for the regression analyses are the slopes of the regression lines based on $\log_{10}$ transformed data. Since $\log_{10}$ transformation is necessary and since only least-squares regression is employed here, the allometry coefficients are the same for both regresssion techniques. Ratio techniques do not directly provide allometry coefficients, so a calculation from reduced major axis regression (Leamy and Bradley 1982) was modified to approximate them. From this the ratio allometry coeffients are the $\log_{10}$ transformation of the standard deviations (S) of each character (y) divided by the standard deviations of the $\log_{10}$ transformed size measure (x). These ratio allometry coefficients are again the same for both ratio adjustments since one of the ratio procedures is simply the $\log_{10}$ transformation of the other. Allometry coefficients for the raw and $\log_{10}$ transformed data are also calculated using this modified ratio formula, and thus they also are the same for both of these data sets. The formula is $log_{10}(S_y/s_x)$.

Figure 16 presents the mean individuals for all fifty-one morphological measurements plotted against the $\log_{10}$ transformed size variable. It demonstrates how effectively size and shape are separated by the bivariate morphometric procedures, and how well and on which axis the two hypothetical groups are separated. The mean individuals were calculated by taking the mean of all the measurements for each individual, and are done separately for the morphological and meristic character sets. The mean individuals represent shape for each fish. The standard length axis is size and it is $\log_{10}$ transformed to again make effective use of space. The plots using untransformed standard length were essentially identical except that the numbers were more broadly scattered and clumped in size-groups along the x-axis. The use of mean individuals is novel, yet it can be justified by their effectiveness and consistency on the figures, and by their very uniform, if high, standard deviations.

These mean individual plots are used to assess bivariate size/shape separation and group portrayal because they do so effectively for the appropriate bivariate allometric hypothesis. They are intended to be analogous to the traditional multivariate scatter plots (figs. 23–24; see **chp. 6**). The use of multivariate scatter plots is often recommended for such an assessment of bivariate techniques (Reist 1985, Shea 1985, Thorpe 1976) but this is testing bivariate effectiveness with the wrong hypothesis (multivariate). Moreover, if size is even partially removed by the bivariate procedure it is difficult to characterize any of the resulting multivariate eigenvectors or principal

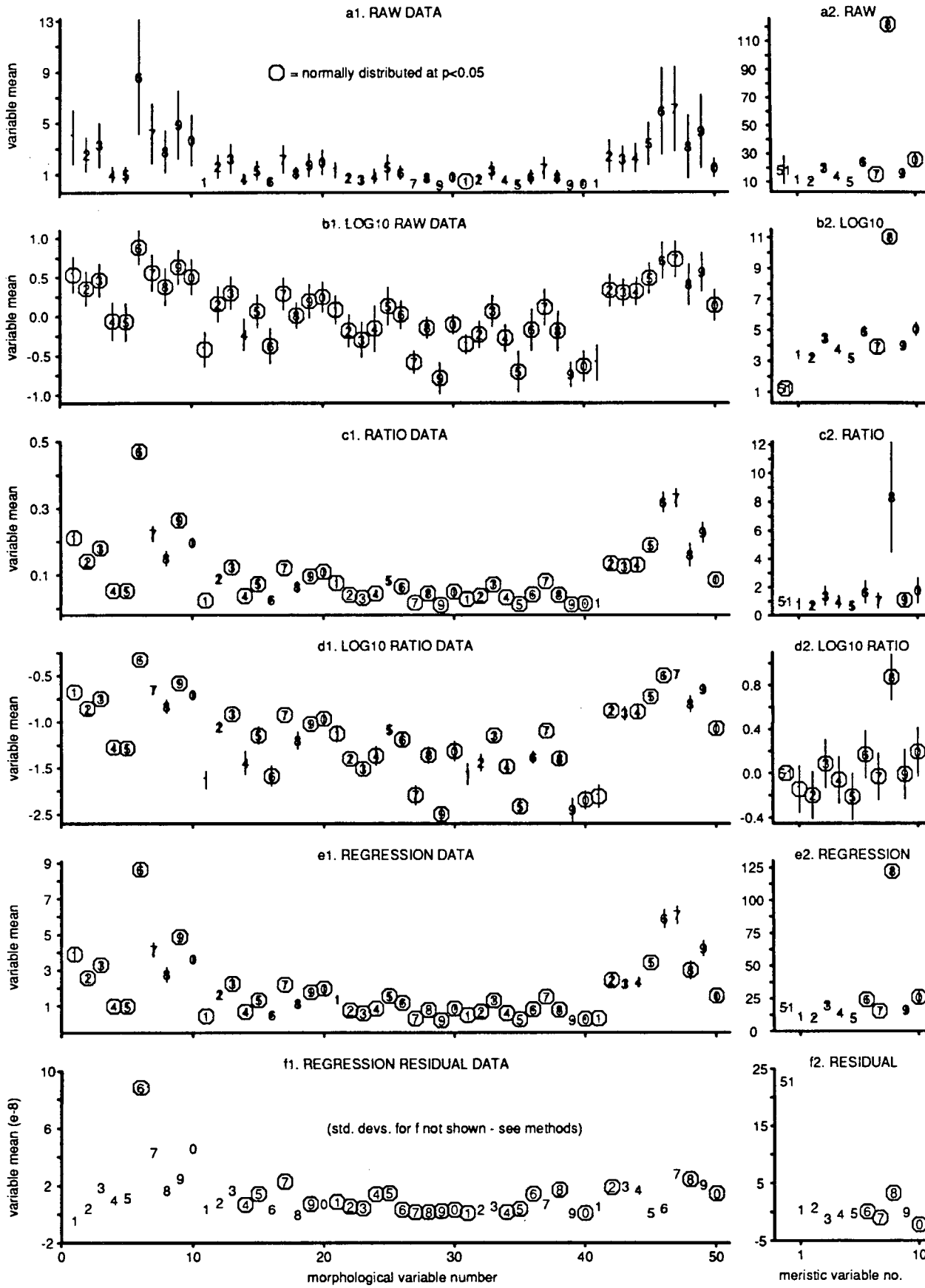


FIGURE 14. Mean, standard deviation and normality for the variables.

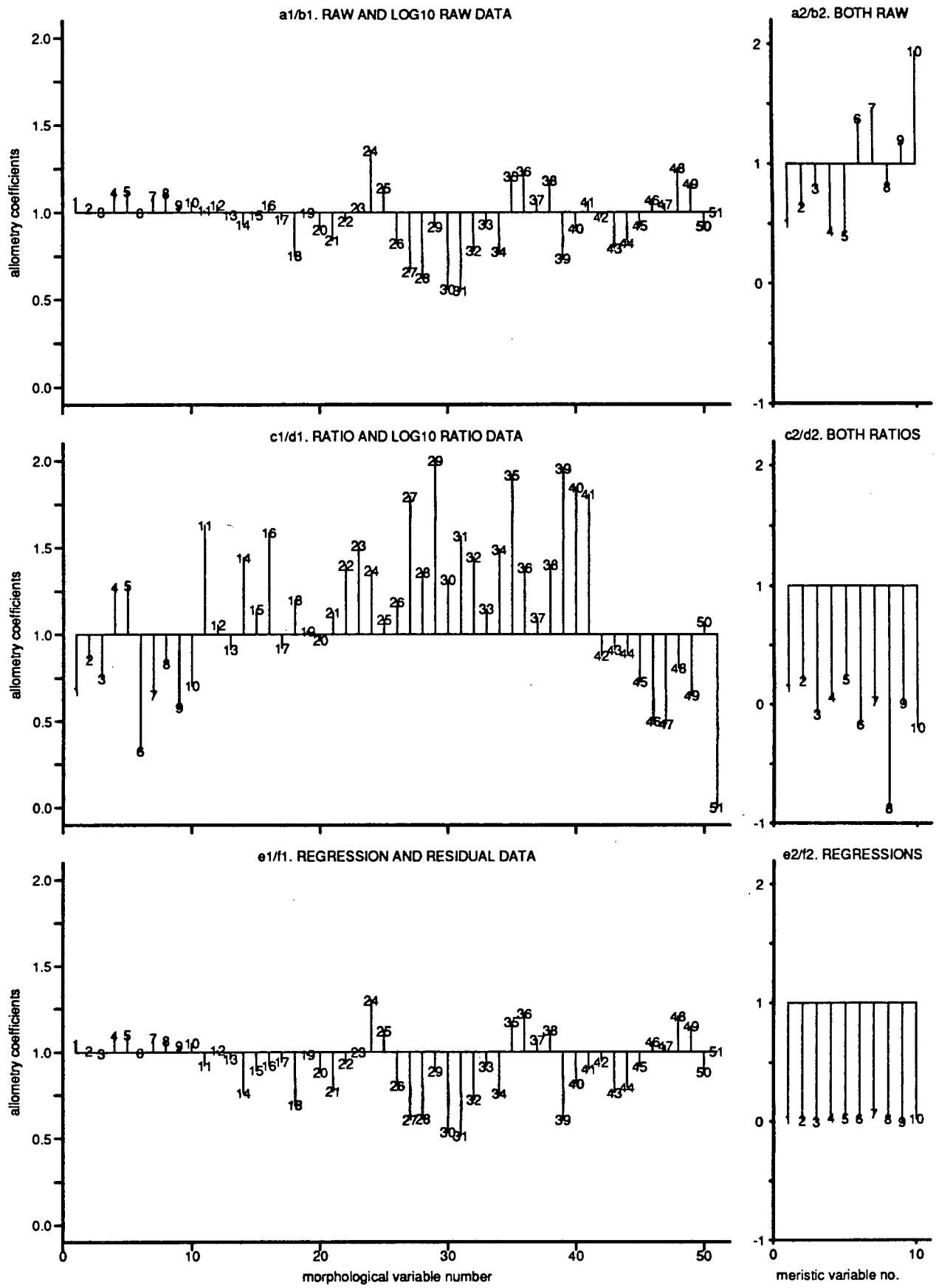


FIGURE 15. Allometry coefficients for the bivariate analyses.

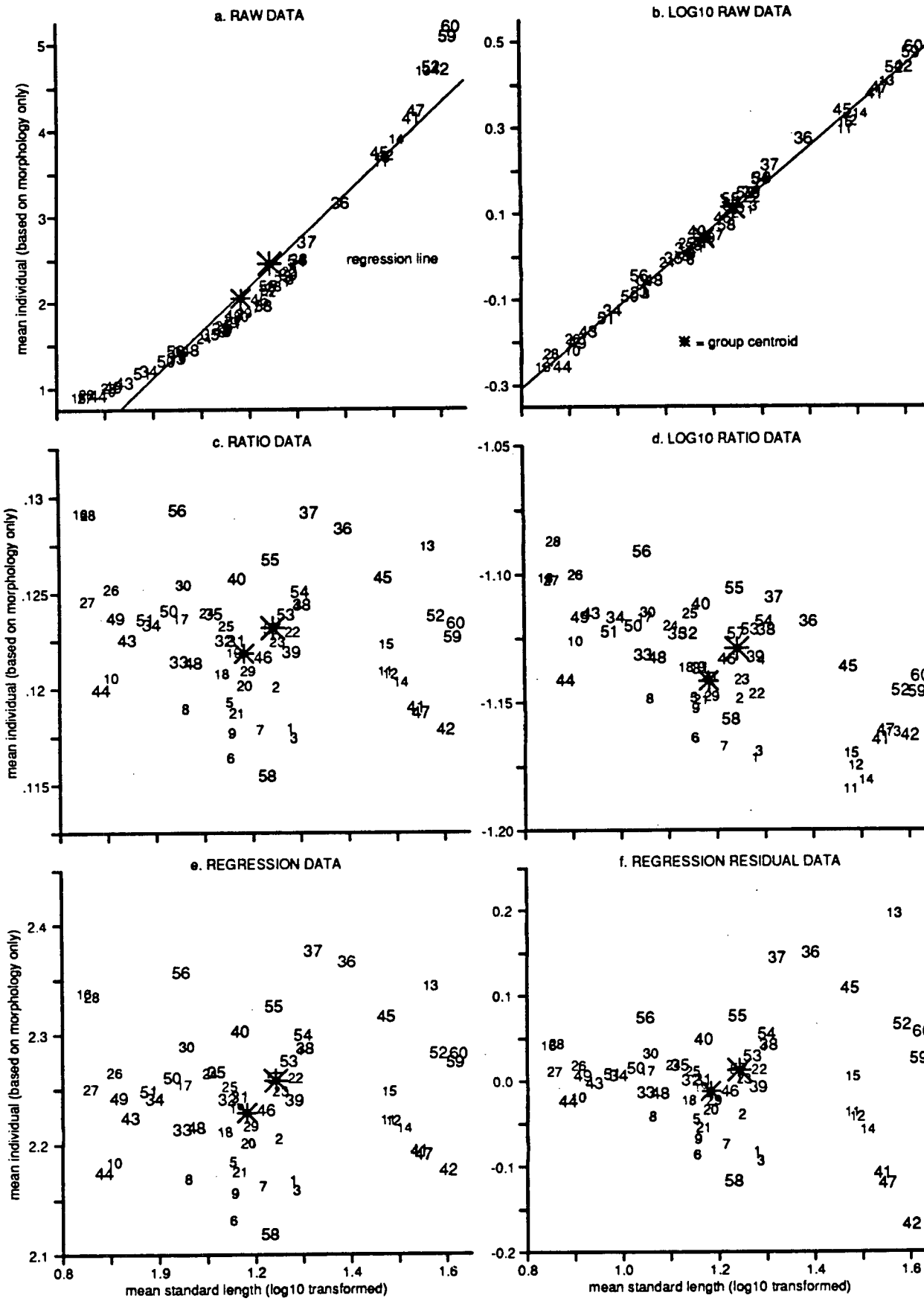


FIGURE 16. Patterns for individuals based on bivariate morphological data.

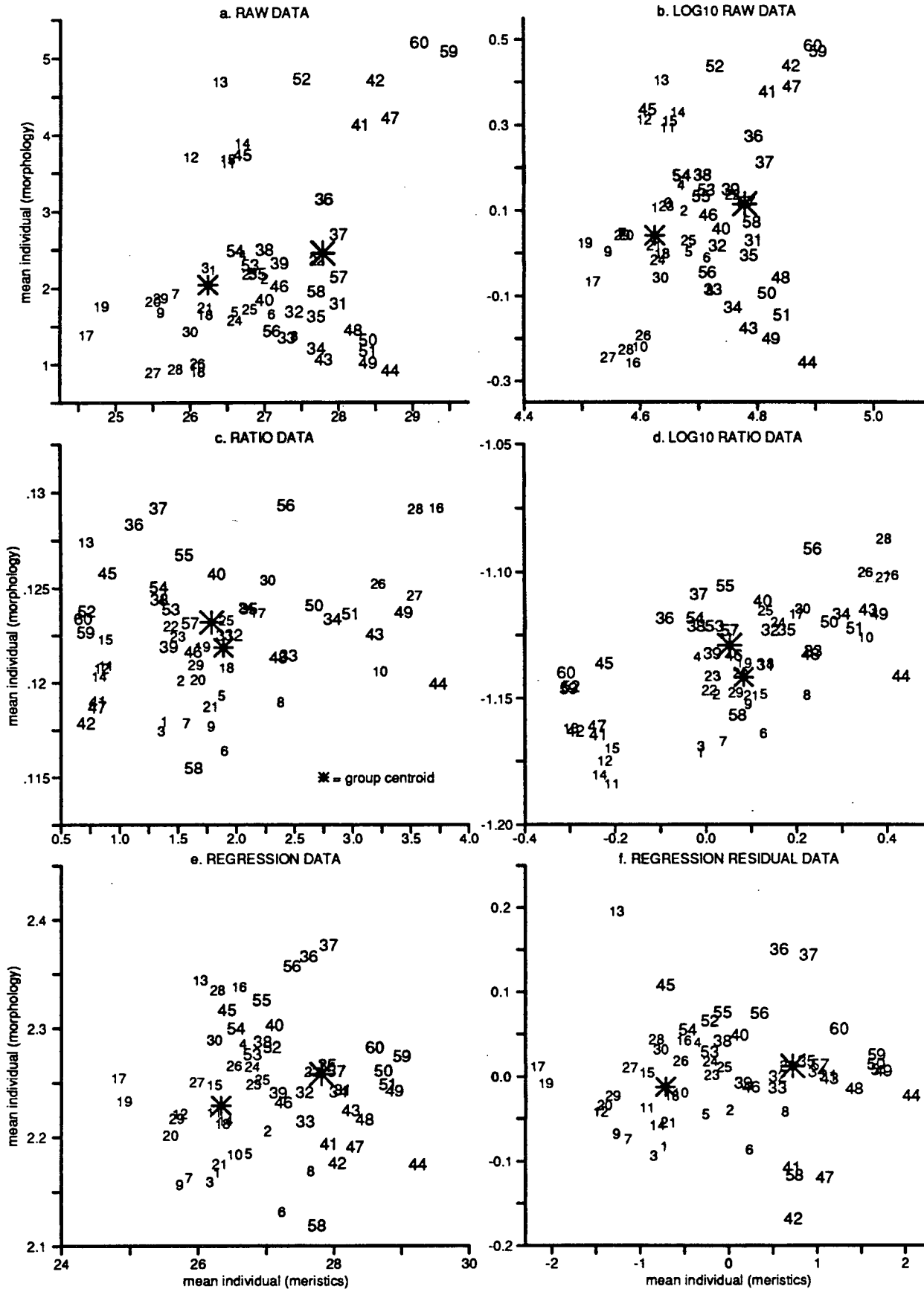


FIGURE 17. Patterns for individuals based on total bivariate data.

Table 1. Regression statistics for data and bivariate procedures.

		mean r^2	N	R^2	R^2 significance
morphology	raw (a)	0.854	51	0.999	$p < 0.0005$
morphology	$\log_{10}$ raw (b)	0.918	51	1.0	$p < 0.0005$
morphology	ratio (c)	0.248	35	0.994	$p < 0.0005$
morphology	$\log_{10}$ ratio (d)	0.254	34	0.994	$p < 0.0005$
morphology	regression (e)	0.007	1	0.443	$p >> 0.5$
morphology	regression residual (f)	0.004	0	0.062	$p >> 0.5$
meristics	raw (a)	0.034	2	0.309	$p > 0.05$
meristics	$\log_{10}$ raw (b)	0.034	2	0.307	$p > 0.05$
meristics	ratio (c)	0.884	10	0.938	$p < 0.0005$
meristics	$\log_{10}$ ratio (d)	0.955	10	0.998	$p < 0.0005$
meristics	regression (e)	0.001	0	0.007	$p >> 0.5$
meristics	regression residual (f)	0.002	0	0.010	$p >> 0.5$

components as size or shape vectors and thus the procedural effectiveness is not easily assessed. The bivariate data also have different distributions and central tendency statistics (fig. 14) and these could affect and confound the multivariate procedures. A multivariate assessment of a bivariate allometric technique is examined in **chp. 7.**

Figure 17 presents the mean morphological individuals plotted against the mean individuals for all ten meristic measurements. Since meristics are not size-dependent in this study (fig. 13(v) (**chp. 4**); table 1) their mean individual plots against size are essentially meaningless. However, the mean meristic individuals thus make a good standard against which shape variables (mean morphological individuals) can be plotted (Bookstein et al. 1985). Figure 17 shows how well and where the conceived groups are delineated.

Table 1 presents the least-squares regression statistics (Mosimann 1970, Mosimann and James 1979, Neff and Marcus 1980, Reist 1985) for each bivariate technique **a–f.** For each method, every character is regressed first against $\log_{10}$ transformed size, and then a multiple regression of $\log_{10}$ transformed size is performed separately against the morphological and meristic data sets. In the first regression, a univariate F is calculated to determine if each character's slope is significantly different ($p < 0.05$) from zero. Zero represents complete size/shape separation. The number of characters that are not significantly separated is given as N. The mean univariate correlation coefficients are presented as r^2. The latter regression is used to obtain the multiple regression F-value to determine if this slope is similarly significantly different from zero and also to get the multiple correlation coefficient (R^2). Size and shape are best separated where the correlation coefficients are low and N is small.

The analyses and graphics are based on computer macros I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Centre, University of British Columbia. These programs are available from me. The macros were validated using literature examples (Reyment et al. 1984).

Bivariate Morphometric Procedures — Assessment Results

The central tendency statistics in figure 14 show that $\log_{10}$ transformation of the raw data (labelled **b**) stabilizes the variances, helps make the data independent of scale, and greatly improves

the univariate normality of the raw data (**a**) itself. The ratio data (**c**) and $\log_{10}$ transformation of ratios (**d**) result in similar distribution patterns to **a** and **b**. Univariate normality, however, is better in **c** (and in **d**) than in **a**. The data are not corrected for scale by **c**, but are by **d**, but univariate normality is slightly decreased in **d**. Neither **c** or **d** stabilize the variances or even affect them in any consistent way. These effects are the same for both morphological and meristic characters, but outwardly **d** seems to result in the most desirable data for meristics.

The regression (**e**) and regression residual (**f**) data are similar to each other and to **a**. The regressions, however, affect some characters differently (eg. size variable no. 51), especially in the meristic data set. Otherwise, these two techniques have the same effects on both the morphological and meristic data. The variances in **e** and **f** are also not stable, and are less so for **f** than **e**. The univariate normality is good for **e**, and while still adequate for **f** it is reduced. Neither regression is independent of scale and this effect of character magnitude is inconsistent.

The allometry coefficients in figure 15 demonstrate that the both **a**–**b** and **e**–**f** have similar patterns for morphological variables. The ratio procedures produce very different results. The allometry coefficients for **c** and **d** are either very high or low and most show positive allometry. This is the opposite to what is seen in **a**–**b** and **e**–**f**. Furthermore, the ratio allometry coefficient for size (variable no. 51) is zero when it should approximate an isometric value of one. The only morphological character pattern noted in the allometry coefficients for **e** and **f** is that those characters with the highest or lowest loadings are those that are not strongly size-related. Measures such as eye size (no. 30) load heavily whereas body proportions such as body depth (no. 1) or body width (no. 2) are closer to isometry.

The raw meristic allometry coefficients presented in figure 15 appear to be realistic and indicative of changes in data relationships in spite of the fact that size is not part of the meristic data. They reveal bad effects for all four bivariate morphometric procedures. The values do not resemble those of the original data at all. The ratio allometry coefficients sometimes reach negative values or are otherwise near zero. The regression allometry coefficients are all near zero.

Table 1 demonstrates that both **e** and **f** effectively remove size from all the morphological variables. In **e**, only the size variable itself (no. 51) is not size compensated. The R^2 value for **f**

is much better than for **e** but this does not affect the success of **e** in separating size. The ratio procedures significantly remove size from about half of the variables. This lack of success is further reflected in their high R^2 and mean r^2 values.

For meristic characters, size is quite effectively accounted for in the original raw data or its $\log_{10}$ transform (table 1). The regressions **e–f** again appear to be the most effective size adjustment techniques. They have accounted for size in all 10 variables and have the lowest R^2 and mean r^2 values. The ratios cause the meristic characters to become completely size-confounded. They cause size to become a problem in meristic data where it is not otherwise.

Figure 16 shows that neither size, nor any of the resultant shape variates, completely delineate the groups or their centroids effectively. However, large and small individuals can no longer be completely distinguished on the shape axis (y-axis; mean individuals). No size group within the total range is better separated than any other group either. All of the bivariate procedure patterns are similar. Procedures **c** and **e** give the most consistent distribution across the entire shape axis but this distinction is not strong. For large individuals, technique **d** is skewed somewhat downward and **f** spreads outward.

Figure 17 demonstrates that such mean individual plots of morphology against meristics can provide reasonable but slight group separation. Separation is best for procedures **a–b** and **e–f** where even the centroids are fairly well separated. This separation, however, is along the meristic axis alone. The patterns for individuals in **a–b** and **e–f** are also quite similar with only the overall distributions changing somewhat. The groups are not separated at all in **c**, while in **d** there is some group separation along the axis of the data cloud, but not along either graph axis. Neither **c** nor **d** separate the centroids well either.

Bivariate Morphometric Procedures — Assessment Discussion

Based on Morphometric Variables

The four bivariate procedures result in different character effects, but these do not produce obvious differences in the analysis of individuals or groups. If individual ordination and group separation are the only objectives then any of these techniques would provide useable, but also

inadequate, results. While the groups are resolved in the scatter plots of mean individuals against standard length (fig. 16), the separation is weak and the group centroids are barely different. In this study, the centroids will never be strongly separated on the size axis because the mean group sizes are very similar. Separation should occur on the shape axis if the morphometric techniques are properly differentiating size and shape. These scatter plots reveal that the univariate size measure is effectively removed in individuals, but their ordination suggests that this removal is insufficient in terms of overall size. There likely still is size information left in the univariate shape characters that comprise each individual. In other words, there is size information in the data that is not represented by the standard length measure and thus is not compensated for by the bivariate morphometric procedures.

In morphometric studies, character information is almost always desired in addition to effective ordination and group separation. This information varies with the procedures employed in this study. Therefore, character interpretations will be different depending on which bivariate technique is used. The two regression techniques are somewhat compatible in terms of characters, but the two ratio procedures are not comparable either amongst themselves or to the regression methods (Reist 1985).

The regression procedures provide the most realistic character information (Reist 1985). Their central tendency statistics (fig. 12) are consistent with the raw data and have a good level of univariate normality. They do not, however, remove scale or variance effects from the data. This latter effect may be undesirable if the data are to be used in further statistical analyses. If the data is not be further analyzed, these regression techniques portray the original data quite well and the deficiencies of scaling and variance are of little consequence. The problems in using such data for further analyses are discussed in **chp. 4.**

The regression procedures also give the most realistic portrayal of the allometric relationships present in the original data. Their allometry coefficients (fig. 15) are virtually the same. The only pattern noticed for the allometry coefficients of the raw and regression data is that characters that intuitively are not very strongly size-related have the highest values, both positive and negative. Shape measures load heavily while body size proportions are closer to isometry. This pattern

makes sense and supports the assessment that the regression techniques correctly are portray the allometric relationships.

The regression procedures are the most effective methods for removing univariate size from the data (table 1). They completely remove univariate size (Reist 1985). While the R^2 values are higher for the regression residuals (labelled **f**) this did not affect the success of the straight regression technique (**e**) in separating size. This assessment is supported by the p- and mean r^2 values which are similar for both regressions. In **e**, only the size variable itself (no. 51) is not compensated and this is not surprising since this is a $\log_{10}/\log_{10}$ regression. The size variable is adjusted in **f** because this regression is untransformed and thus less tight.

Neither ratio procedure removes size effectively (Reist 1985). More than half the variables still contain a statistically significant amount of the univariate size measure. The effect of the ratio procedures on the characters is of more concern. While their central tendency statistics (fig. 14) are somewhat similar to the raw, $\log_{10}$ transformed and regression data, their allometry coefficients are very different. The ratios do not provide data that are size-free and do not result in a realistic portrayal of the allometric relationships (fig. 15) that exist among the characters. It is perhaps not unexpected that the ratios have discrepant allometry coefficients, since these procedures do not allow for size to become a direct part of the allometric calculation. This is another drawback of the ratio methods.

My choice for best bivariate technique is straight regression (**e**). While regression residuals (**f**) are at least equally as good as the regression itself (Reist 1985), the residuals are quite different numbers than the original data. The regression technique (**e**) not only provides data in the same dimensions as that of the raw data but it also adjusts them all to a mean body size. This allows for a natural and intuitive understanding of the procedural output and permits overall mean individuals for a group to be estimated. These group mean individuals can then be readily compared and the differences between groups become apparent. Furthermore, the regression technique calculates allometry coefficients directly. Finally, the central tendency statistics, the distribution of shape represented by the mean individuals and the overall character variances are slightly better for straight regression than for residual regression.

Based on Meristic Characters

All of the effects of the bivariate procedures on the morphometric variables are the same for the meristic characters. What is more important here, however, is that it is clear that these meristic characters do not need allometric compensation. This is also true of all non-size related variables (Reist 1985, Somers 1986).

The central tendency statistics of the meristic characters (fig. 14) are often different and the resultant allometry coefficients (fig. 15) do not resemble those of the original data. The raw and $\log_{10}$ transformed mean morphological and mean meristic individual scatter plots (fig. 17) provided group separation which, while still minimal, are as effective as that of the regression procedures. The mean individual scatter plots based on meristic characters adjusted by the ratio techniques did not reveal either hypothetical group. In short, allometric compensation of meristic variables results in characters which do not portray the original data. These compensated variables also do not help in ordination of individuals or groups.

The regression procedures appear to completely remove size from the meristic characters (table 1), but this may be a consequence of loose regression and not the result of good size removal. The raw (**a**) and $\log_{10}$transformed (**b**) meristic variables do not regress strongly on size and thus are not related to size. While **a** and **b** do not remove size for two characters at $p < 0.05$, they do so at $p < 0.1$. Meristic characters are therefore probably more than adequate unadjusted by any bivariate technique. Generally, bivariate statistical procedures should not be used to compensate meristic or other characters for allometry unless size effects actually are present (Reist 1985, Somers 1986).

If, however, size effects are present in the meristic characters, straight regression (**e**) is probably the best technique to deal with it. The reasons are the same as mentioned for the morphological variables, and in addition this study has shown that regression has the least negative effects on the meristic characters. The central tendency statistics and mean individual scatter plots based on straight regression data are the most similar to those of the raw data. However, allometry coefficients from **e** are still different.

Bivariate Morphometric Procedures — Summary

1. Mean individual ordination and group separation achieved by all four bivariate procedures is similar but inadequate.
    a) Univariate size is removed effectively from the mean individuals but other confounding size information remains.
2. The underlying effects on characters and their relationships differ greatly between these procedures.
3. Regression (technique **e**) is the best bivariate morphometric procedure.
    a) Univariate size is completely removed from the characters.
    b) It protrays the data and its allometric relationships the most realistically.
    c) It results in the simplest, most informative and useful output.
4. Do not adjust meristic characters or any other variables if they have no size information.
    a) If size information is present in the meristic characters use the same regression procedure (**e**) for allometric adjustment.
5. In this study, logarithmic transformation of the raw data stabilizes variances, helps make the data independent of scale, and greatly improves univariate (and multivariate) normality.

CHAPTER SIX

Assessment of Multivariate Morphometric Procedures

Introduction

Multivariate morphometric procedures simultaneously analyze all characters without reference to any size measure, and produce new independent size and shape variables for each of the characters. The major benefits of multivariate techniques are data synthesis and pattern recognition. Their main initial drawback is the seemingly difficult mathematics. The complex interactions of all the characters are simultaneously analyzed and reduced to a smaller number of relationships (Pauken and Metter 1971, Reyment et al. 1984) that often may not have been discernible in the original data (Corruccini 1978, Jolicoeur 1959, Shea 1985). Some of the individual coefficients resulting from multivariate analyses may have no immediate biological interpretation, but the combined overall patterns they define might have significance (Holland 1968, Reyment et al. 1984).

The multivariate statistics used in this study all involve the rigid rotation of the axes describing the original variables and result in new axes that each subsequently explain the maximum variation possible while remaining orthogonal (uncorrelated) to each other. Each new character axis is called an eigenvector (a.k.a. eigenroot, characteristic vector/root or latent vector/root) and the analysis produces as many eigenvectors as the number of original variables. Equivalent new axes for individuals are derived from these eigenvectors and from the original variables. Principal component analysis (PCA) is used predominantly in this multivariate study and in PCA these vectors for individuals are referred to as principal components (PC's).

Each number in an eigenvector represents and corresponds to an original character, and together the numbers define that eigenvector. The same is true of the PC's but here the numbers represent individuals and define that PC. The eigenvector numbers are termed loadings or coefficients, and those subsequently calculated for individuals are called scores. They can all be read just like any other data values (Davis and Baker 1974, Neff and Marcus 1980, Pimentel 1979, Shea 1985). The magnitude and sign of the numbers, and each eigenvector's or PC's collective signs, are important. The larger each number, either positive or negative, the more significance its corresponding original character has in that eigenvector or PC (Marriott 1974). The signs of the

numbers reveal how they are related. The collective signs of each eigenvector or PC indicate the overall relationships in it and represented by it.

Each eigenvector, and by default each PC, is represented by another number called the eigenvalue. Eigenvalues indicate how much variation their corresponding eigenvectors and PC's account for. To reiterate somewhat, multivariate analyses explain the most variation possible in the first eigenvector and PC, and then in the second eigenvector and PC they attempt to explain as much as possible of the variation remaining that is orthogonal to the first, and so on. The eigenvalues reveal exactly how much variation is actually accounted for in each of these resultant vectors.

In multivariate morphometrics, the first eigenvector usually accounts for not only the most but the vast majority of the variation and is also general (all the loadings have the same sign). This first eigenvector is the size vector, and size vectors are characterized by large general morphometric eigenvectors. Also, if the first eigenvector is large and general, size is usually effectively removed (Campbell 1976, Reyment et al. 1984). This size removal, however, is not necessarily total or isometric (Shea 1985, Somers 1986). The second eigenvector usually accounts for most of the remaining significant variation and is bipolar (the loadings have mixed signs). This second eigenvector represents the shape vector, and in fact shape vectors are characterized by any significant bipolar morphometric eigenvectors subsequent to the first eigenvector (Somers 1986). Therefore, the scores of individuals on the first PC are measures of their overall body size, while the loadings on the first eigenvector are estimates of the rates of change of individual characters with size (Lande 1985, Leamy and Bradley 1982, Strauss 1987). The scores and loadings on the second PC and eigenvector are respectively measures and estimates of shape with non-contributing size and its rate of change removed.

This overall interpretation has been justified mathematically (Rao 1964), and it makes intuitive sense since large size-related variation normally predominates in morphometric data (Sacher 1970). Furthermore, larger individuals have larger scores on the first PC (Brower and Veinus 1978), whereas the second PC usually does not distinguish between small and large individuals but rather between their groups (see figs. 23-24). The first eigenvector's general sign is also indicative of a single type of variation (Pimentel 1979). As well, organism weight is strongly size-related and

tends to be highly correlated with the first eigenvector and uncorrelated with the second (Phillips et al. 1973).

Usually only two, but sometimes three, morphometric eigenvectors and PC's are interpreted as they account for most of the variation (Thorpe 1976). The number of eigenvectors (and hence PC's) which summarize statistically significant variation can be conservatively estimated using Bartlett's chi-square test of sphericity (Cooley and Lohnes 1971, Phillips et al. 1973, Pimentel 1979) or the Scree test (Cattell 1966, Somers 1986). While Bartlett's and Scree tests often are comparable (Horn and Engstrom 1979), Bartlett's test generally is better for smaller sample sizes and the Scree test for larger ones (Reyment et al. 1984). Bartlett's test is actually a modification of a more effective test (Anderson 1963) which unfortunately requires very large sample sizes and is thus rarely possible (Pimentel 1979). Besides, Bartlett's modification usually is just as good. Jackknifing (see **chp. 4**) can also help determine which eigenvectors are stable and significant (Gibson et al. 1984). Furthermore, Kaiser's rule (Kaiser 1960) is a rule-of-thumb which says that no eigenvectors with eigenvalues less than one should be interpreted. This advice is usually borne out by statistical tests and should be kept in mind.

All these tests reveal that only the first two eigenvectors and PC's are significant (Bartlett's test: $p < 0.05$) in my study. The simple graphical Scree test for this is representatively shown for morphological variables in figure 13(iv) (**chp. 4**). The Scree test shows which eigenvectors are important by where the curve stabilizes completely and tapers off. Conservatively, only those eigenvectors that are significant should be interpreted (Thorpe 1983a), but this is often justifiably ignored, especially when only the first eigenvector is significant (Pimentel 1979). Interpreting more than three eigenvectors, however, can be confusing and is not recommended except under defensable circumstances. The objectives of the analysis should help decide how many and which eigenvectors to interpret (Reyment et al. 1984).

Since the first three eigenvectors sequentially account for greatly decreasing variation yet receive the same or sometimes more visual emphasis on graphs they are also difficult to portray (Baker et al. 1972, Ball and Hall 1970, Boratynski and Davies 1971, Everitt 1978, Reyment et al. 1984). In such situations, three-dimensional graphs are not the solution and symbolics on two-dimensional graphs appear to be much better and more appropriate (Atchley et al. 1982, Marriott

1974). Furthermore, if the significant variation is expressed in the first two PC's their scatter plots represent real overall individual distances (Everitt 1978, Marriott 1974, Reyment et al. 1984). Tests on the effectiveness of multivariate graphical presentation have also shown that PCA scatter plots are one of the best techniques (Corruccini 1978, Friedman and Rafsky 1981, Marriott 1974, Page 1978; for partially opposite see Jamison and Zegura 1974).

The rationale for multivariate size and shape vectors is different from the bivariate morphometric procedures (see **chp. 5**), because here growth, size and shape are multivariate factors and not directly measured variates (Humphries et al. 1981, Thorpe 1983b, Thorpe and Leamy 1983). Size is also not simply negated or reduced to one measure as with the bivariate techniques. In multivariate procedures, size becomes a distinct part of the analysis and of each variable in the character set (Bonner 1965, Clutton-Brock and Harvey 1977, 1979, Humphries et al. 1981, Thorpe 1976). Furthermore, there is no problem choosing a representative size variable and no independent/dependent variable semantics (Thorpe 1983b; also see **chp. 5**). Variable correlations are used rather than ignored (Lande and Arnold 1983, Reyment et al. 1984), and both overall and univariate significances can be directly assessed. The underlying allometric hypothesis here is multivariate and realistic.

Multivariate Procedures — Discriminant Function and Canonical Variates Analyses

Principal component analysis (PCA), linear discriminant function analysis (LDFA), and canonical variates analysis (CVA) are the three common multivariate approaches to morphometrics. In this study, however, all four multivariate procedures employed are based on PCA, and this section reveals why.

Dealing with hypothesized *a priori* group structure in multivariate data implies using LDFA or CVA because both these multivariate techniques require *a priori* group designation and account for group relationships (Reyment et al. 1984, Somers 1986, Thorpe 1980). There are many reasons, however, why PCA is usually a better alternative. Foremost, *a priori* group designation is subjective and assumes that there is only one taxon per group and that it can be completely distinguished (Humphries et al. 1981, Thorpe 1976, 1980). If this subjective approach is desired or the necesary background information supporting it is available it may be appropriate to use LDFA or CVA.

However, *a priori* designation should usually be initially avoided in morphometrics. LDFA and CVA are discriminating procedures, not descriptive ones, and should only be used as such. For a further detailed discussion on *a priori* group assignment see **chapter 4** (data pooling).

The LDFA and CVA procedures also have many related undesirable features. They are mainly involved in ordinating groups and not individuals (Thorpe 1976) because between-group variation is maximized in relation to within-group variation (Humphries et al. 1981, Lachenbruch 1975, Pimentel 1979, Thorpe 1983a). This maximizing of group separation and minimizing of group overlap emphasizes best separated populations when interest is often more in the least separated populations (Habbema and Hermans 1977). Also, information on individuals and characters can be spurious or lost through this maximization, and size/shape relationships can be confounded (Humphries et al. 1981, Somers 1986). The highest discriminating variables are loaded most heavily even though the lower discriminating variables may contain just as much size or shape information.

LDFA and CVA also have more exacting statistical requirements and are somewhat less robust than PCA (Gilbert 1968, Harris 1975, Holloway and Dunn 1967, Krzanowski 1977, Lachenbruch et al. 1973, Lachenbruch and Goldstein 1979, Moore 1973, Pimentel 1979, Thorpe 1980), especially in regards to homoscedasticity (Gilbert 1969, Marks and Dunn 1974, Sneath and Sokal 1973, Thorpe 1976). The procedures also work best with large sample sizes and a smaller number of characters (Dunn and Vardy 1966) since better group separation is often achieved with fewer but more significant variables (Dunn 1971, Farver and Dunn 1979, Jain and Walker 1979, McKay and Campbell 1982b, McLachlan 1976, Srivastava and Carter 1983).

Unknown hybrid or introgressed individuals or groups are also a difficulty, even though LDFA and CVA can be effective discriminators for known ones (Bloom 1976, Eyles and Blackith 1965, Hatheway 1962, Neff and Smith 1979, Reist and Crossman 1987, Schueler and Rising 1976, Szij 1962, Yang and Selander 1968). In addition, the effects of unequal group sample sizes on LDFA and CVA are potentiallya more serious problem (Neff and Smith 1979) than in PCA.

PCA presents none of these problems since it is a descriptive, and not a discriminating, procedure. It can be used to find groups and then if desired or necessary these groups can be appropriately dealt with in further analyses (eg. Humphries et al. 1981). Subsequently using LDFA

or CVA, however, does not remove their difficulties, and thus regression or another PCA should be performed on any pooled groups deemed necessary. If group discriminations and no character information is desired then LDFA or CVA should be used (see **chp. 1**).

LDFA with equal assignment probabilities is used to verify my PCA analyses. This practise is recommended by many authors (Claytor and MacCrimmon 1987, Crovello 1970, Mosimann and Malley 1979, Thorpe 1976, 1983a, Thorpe and Leamy 1983) and here it gave virtually identical results for the group separations and similar results even for the individuals and characters. This similarity increases as variable number goes up because a high number of characters helps balance out the weightings of discriminant variables in LDFA and CVA. This is probably one of the reasons why the similarity is so high in this study. Nonetheless, this is a confirmation of these PCA results (on data not pooled for groups) because LDFA and CVA operate differently from PCA and they also directly account for within-group character correlations through this difference (Atchley 1980, Campbell 1976, Reyment et al. 1984, Thorpe 1976, 1980 ,1983).

Cluster and fourier analyses are rarely used as multivariate techniques and the reason for this deserves explanation. Cluster analysis is relatively unrobust for such work (Boratynski and Davies 1971) and can impose misleading categorical structure, fail to separate groups and cannot assess character contributions (Hiernaux 1972, Thorpe 1983a). Fourier analysis does not deal with homology (Bookstein et al. 1982) and is thus not a comparative technique. It may, however, describe single shapes quite well (Read and Lestrel 1986). Neither is recommended for morphometrics, and definitely not for allometric compensation.

Multivariate Procedures — Principal Component Analysis

The four multivariate allometric procedures looked at here are all based on Q-mode (done on individuals) principal component analysis (PCA). It is the original (Jolicoeur 1963a, Jolicoeur and Mosimann 1960) and usually considered the best (Bookstein et al. 1985, Holmes 1975, Humphries et al. 1981, Pimentel 1979, Somers 1986, Timm and Price 1980) multivariate way to deal with morphometrics and allometry. It has none of the LDFA or CVA problems and is unsubjective, repeatable and robust (Corruccini 1983, Dudzinski et al. 1975, Harris 1975, Reyment et al. 1984,

Thorpe 1976, 1980). It also effectively deals with unkown hybrids and introgressants (Clifford and Binet 1954, Lawrence and Bossert 1969, Neff and Smith 1979, Pimentel 1981, Sokal 1965).

Some authors suggest that PCA requires no statistical or data assumptions if it is used descriptively and not statistically (Boratynski and Davies 1971, Campbell 1976, Crovello 1970, Dudzinski et al. 1975, Marriott 1974, Pimentel 1981, Rao 1952, Reist 1985). Commonly, however, most morphometric applications have some statistical tests involved, and these in fact are recommended provided that they not be strictly and solely relied on (Gower and Ross 1969, Reyment et al. 1984, Tukey 1962). Moreover, meeting the assumptions and test requirements increases confidence in the results, simplifies explanation and prevents overinterpretation.

PCA on a covariance or correlation matrix (also called a z-score matrix (Pimentel 1979)) are the two standard analyses. Two variants looked at here are shear analysis (Humphries et al. 1981) and the size-constrained method (Somers 1986). The former basically involves supplementing PCA and is based on a covariance matrix, and the latter manipulates PCA directly and is based on a correlation matrix. All four will be discussed in turn.

One additional PCA variant that has only seen rare use (Baumgartner et al. 1988, Delany and Healy 1964, Reyment and Banfield 1976, Rohlf and Bookstein 1987) is that of Burnaby (1966). It works through a series of matrix manipulations prior to the PCA, and has been theoretically justified (Rao 1966b). It was attempted here but the necessary matrix algebra was found to be unwieldy with the large number of characters in this data set. As well, the estimation of an appropriate *a priori* size vector is also a necessary part of Burnaby's technique and this is the classic drawback of this method (Burnaby 1966, Gower 1976, Humphries et al. 1981, Reyment and Banfield 1976, Rohlf and Bookstein 1987). The technique is also more like LDFA or CVA in its procedures and subjectivity. Furthermore, Bookstein et al. (1985) and Humphries et al. (1981) say that this procedure is only a partial discriminator and that the resultant coefficients are not loadings and cannot be compared among themselves. However, Rohlf and Bookstein (1987) later state that this is not a problem and is the result of Burnaby's method performing complete size correction and not just allometric adjustment. It was thus left out of this study since keeping this large character set was important in helping stabilize the overall procedure comparisons (see **chp. 4**) and because Burnaby's technique has seen so little use and is partly subjective.

Burnaby's technique may deserve another look with an appropriate variable set but is unlikely to yield better results than the PCA techniques analyzed here (Rohlf and Bookstein 1987). This limited use of Burnaby's procedure suggests that it produces virtually identical results to the other PCA techniques, particularly to the shear method, and that any differences between them are consistent and do not affect the ultimate interpretations (Rohlf and Bookstein 1987). Burnaby's technique is computationally simple on reasonably sized variable sets (Burnaby 1966, Rohlf and Bookstein 1987) and if complete, orthogonal and subjective *a priori* size removal is desired then it should be investigated further. It will probably result in better taxa ordination and discrimination in these cases, but its character loadings and description of forms in these taxa likley will not be as realistic. Its use with isometric size vectors (*sensu* Somers 1986; see size-constrained PCA in this chapter) may also be of interest (Rohlf and Bookstein 1987).

Standard Principal Component Analyses Formulas

The formulas (see any multivariate statistics text; eg. Pimentel 1979) for calculating the standard principal component analyses based on the covariance and correlation matrices are:

characteristic equation: $|S^2 - \lambda I| = 0$

simplification step: $\mathrm{L} = \lambda_i I$

eigenvector calculation: $(S^2 - \lambda_i I)a_i = 0$

principal component calculation: $y_i = a_i'(x - \bar{x})$

eigenvalue calculation: $\lambda_i = a_i' S^2 a_i$

where:

S^2 = covariance or correlation matrix;

x = original variables;

$\bar{x}$ = mean of original variables;

y_i = i^{th} principal component scores;

a_i = i^{th} eigenvector;

a_i' = transposed i^{th} eigenvector;

λ_i = i^{th} eigenvalue;

I = identity matrix;

L = diagonal matrix.

$$\text{ie. } I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{ie. L} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_i \end{pmatrix}$$

Sheared Principal Component Analysis and Formula

The shear method (Bookstein et al. 1985, Humphries et al. 1981, Rohlf and Bookstein 1987) uses traditional PCA on a $\log_{10}$ transformed covariance matrix to identify the group structure in the data. A pooled within-group covariance matrix is then constructed with the groups based on the ordination results of the previous traditional PCA. A second PCA is then done on this pooled within-group covariance matrix to extract a within-group size vector ($PC1$). The shape components ($PC2$) of the original total PCA are then regressed on this within-group size vector so that they are independent of it. A new final shape vector is then calculated from this regression.

This shear procedure can sometimes provide somewhat better separation of size and shape than traditional PCA (Claytor and MacCrimmon 1987, Shea 1985, Thorpe 1983), and partially

deals with any potential multiple group problems (see **chp. 4** — data pooling). The final shape vector is also mostly uncorrelated to size within groups (Rohlf and Bookstein 1987), and thus holds all the size-free discriminatory (between groups) information. This procedure only initially allows for the examination of the first two eigenvectors or PC's (more can be looked at later if desired), and Reyment et al. (1984) warn that this method may become onerous if a large number of groups are analyzed.

The calculations for the shear method are the same as given for standard principal component analyses except they are supplemented as discussed. The mathematical formulation of this discussion reveals no new information and is all available in the references cited. The only aspect of the formulae that needs explanation is the calculation of the final shape vector after the regression of the original total shape vector on the pooled within-group size vector.

$$H = PC1(-\alpha\beta) + PC2(1 - \alpha\alpha)$$

where:

H = final shape vector;

$PC1$ = pooled within group size vector (pooled within group PC1);

$PC2$ = original shape vector (total PC2);

α = regression intercept (total $PC2$ regressed on pooled within group $PC1$);

β = regression slope (total $PC2$ regressed on pooled within group $PC1$).

Size – Constrained Principal Component Analysis and Formulas

The size-constrained method (Somers 1986) manipulates the first eigenvector so that it represents isometric size alone. The remaining information is then partitioned into the subsequent eigenvectors and represents shape. This method attempts to completely isolate isometric size from shape, but assumes that an isometric size vector exists and that all characters are correlated with it (Somers 1986).

Somers (1986) warns against using this method if negative eigenvalues result from the calculations. He suggests that PCA on a logarithmically transformed correlation matrix be employed in this event (also see Chatfield and Collins 1980, Rohlf and Bookstein 1987). Negative eigenvalues were not encountered in this study (see **chp. 4**).

The isometric size vector for a PCA of logarithmically transformed data (character number = p) is a first eigenvector with all values of $p^{-0.5}$ (Jolicoeur 1963a, Mosimann 1970, Pimentel 1979, Somers 1986). In the size-constrained method, the isometric size vector is extracted from a logarithmically transformed correlation matrix because this correlation matrix standardizes the data to zero mean and unit standard deviation and more closely approximates an isometric size vector itself than does a logarithmically transformed covariance matrix (Reyment et al. 1984, Somers 1986; figs. 18 and 21). The residual matrix after this extraction is factored into eigenvectors which represent shape and random variation alone (for factoring method see Cooley and Lohnes 1971, Holland 1968, Somers 1986). All the principal components are calculated as described in the standard PCA formula section.

The calculations for the size-constrained method are the same as given for standard PCA except that they are manipulated as follows:

isometric size vector: $a_1' = (p^{-0.5}, p^{-0.5}, \ldots, p^{-0.5})$,

corresponding eigenvalue: $\lambda_1 = a_1' R a_1$,

residual matrix: $R_1 = R - (a_1)^2 \lambda_1$.

where:

a_1 = isometric eigenvector;

a_1' = transposed isometric eigenvector;

λ_1 = eigenvalue associated with a_1;

R = correlation matrix ($\log_{10}$ transformed);

R_1 = residual matrix.

Multivariate Procedures — Assessment Methods

The multivariate results are presented in figures 18–25 and table 2. All the morphological and meristic characters are respectively numbered 1–51 and 1–10, and are independently represented on figures 18–22 by separate layouts and captions. The sixty individual fish are portrayed on figures 23–25 which are completely separate from the character representations. Group one is represented by numbers 1–30 set in small type and group two by numbers 31–60 set in large type. Centroids (group means) for each of these groups are also plotted in small and large sizes. Each graph and table is for the covariance matrix data (labelled **A**), shear matrix data (**B**), correlation matrix data (**C**) and size-constrained matrix data (**D**).

Figures 18–20 respectively present the first two eigenvector loadings for the morphological characters, how much of the variance of the original morphological characters the first two eigenvectors account for, and the correlations between the first two eigenvectors and the original morphological characters. These figures permit an assessment of the effects on the individual and overall character patterns by each of the four multivariate morphometric analyses. These punk plots (for term see **chp. 5**) are extrapolated from or centred about zero, and within each figure the data are plotted against equivalent axes. Column one in each figure corresponds to eigenvector one, column two to eigenvector two, and each row represents one of the multivariate techniques **A–D**.

Figure 25 presents the punk plots for the variance of the original individuals accounted for by the PC scores of each multivariate morphometric procedure. It is arranged like figures 18–20 discussed above. This figure reveals how well and where the techniques are accounting for the majority of the individual and overall variances. This plot of individuals is based only on morphological variables for reasons discussed below in regards to figures 23–24.

Figure 21 presents the allometry coefficients for both the morphological and meristic variables. Each row again corresponds to one of the multivariate procedures examined. This figure permits an assessment of the individual and overall patterns of allometry for the character data resulting from each multivariate technique. The allometry coefficients are plotted about an isometric value of one and against equivalent axes. If the allometry coefficient for a character is one then that character is isometric. If it is greater than one then positive allometry is present, and if less

than one negative allometry is. The size of the allometry coefficients, greater than or less than one, indicate how strongly the characters are positively or negatively allometric. Since size is not part of the meristic characters here (see discussion below for figs. 23–24), their allometry coefficients may be unrealistic. Regardless, these meristic allometry coefficients still indicate which multivariate procedures alter data relationships, and they are the only direct link between multivariate and bivariate morphometric procedures (Jolicoeur 1963a–b, Leamy and Bradley 1982, Shea 1985).

Since all the multivariate procedures employed here are based on PCA, their allometry coefficients are all calculated in the same way (Jolicoeur 1963a–b,Lande 1985, Leamy and Bradley 1982, Sacher 1970, Shea 1985). The first eigenvector is made isometric for all the characters (number = p) by tranforming each of their eigenvector loadings to values of $p^{-0.5}$ (Jolicoeur 1963a, Mosimann 1970, Somers 1986). Allometry coefficients are then obtained by dividing the actual first eigenvector by this first isometric eigenvector. While the allometry coefficient calculations for all four PCA's are the same, their allometry coefficients are different since their eigenvectors are different. These allometry coefficient calculations are based on $\log_{10}$ transformed data.

Figure 22 presents the same graphics as in figures 18–20, but this time only for meristic variables. It permits an assessment of the effects of each morphometric procedure on the individual and overall meristic character patterns. Its three columns each respectively correspond to eigenvector loadings, percentage variances and correlations. Each row corresponds to one of the multivariate techniques (**A–D**) employed.

Figure 23 presents scatter plots for individuals based only on morphological variables. PC2 is plotted against PC1. It demonstrates how well size and shape are separated by each technique, and how well and on what axis the groups are delineated. A punk plot for individual PC scores (ie. as for characters in fig. 18) was not constructed as this information is on figure 23. Look at each PC axis of figure 23 from the perspective of a punk plot. How the individuals score onto it and their overall relationships should become apparent.

Since my meristic characters are not size-dependent a scatter plot of their PC1 and PC2 is meaningless in terms of size/shape analyses. Such a plot reveals complete group separation but only along the PC1 axis (fig. 13(vi); **chp. 4**). The meristic PC1 axis also accounts for a greatly reduced

percentage variance (fig. 24) and is no longer general (fig. 22). It therefore does not correspond to a size vector. Meristic PC1 does, however, make an excellent vector against which morphological shape vectors can be plotted and groups discriminated (Bookstein et al. 1985). Figure 24 presents the scatter plot for individuals based on PC2 for morphological variables plotted against PC1 for meristic characters. It demonstrates how well and on what axis the groups are delineated.

The respective eigenvalue percentage variances of PC1 and PC2 in are presented in figures 18 and 23–24. These eigenvalues demonstrate how much overall variation is being accounted for by the first two eigenvectors and PC's in the analyses. Cumulative eigenvalue percentage variances (PC1 + PC2) are only given for figure 23 since addition of the two separate PCA's in figure 24 would be incorrect. Cumulative eigenvalues are not presented in figure 18 because of lack of space, but they can still be calculated by summing those from eigenvectors one and two. The cumulative eigenvalues in figure 18 are also the same as in figure 23.

The axes of the plots based on covariance matrices (**A** and **B**) in figures 23–24 are mean-centred. This is standard practise and occurs automatically in most statistical packages. Confidence ellipses (Jolicoeur 1959, 1963a, Owen and Chmielewski 1985, Phillips et al. 1973) are not drawn for any of the scatter plots because they are virtually non-overlapping (at 95 %) for **A**–**C** and are useless in **D**. Their contribution to the plots is clutter only.

Table 2 presents the statistical isometry tests (Pimentel 1979; also see Leamy and Bradley 1982, Somers 1986, Thorington 1972). The "degrees" from isometry show how many degrees each of the first two eigenvectors are from isometry. The number of degrees each eigenvector is from isometry gives some impression as to how well size and shape are separated (assuming isometry exists) and to how orthogonal the two eigenvectors are. Perfect orthogonality is ninety degrees. The "χ-value" part gives the nearest p-value derived from Anderson's (1963) chi-square test. If this value is significant ($p < 0.05$), eigenvector isometry is not achieved and some size information is assumed to be present in it.

All these analyses and graphics are based on computer macros I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Centre, University of British Columbia (U.B.C.). These programs are available

from me. The standard multivariate statistics were verified (Rhoads and Trinkaus 1977) using the MIDAS statistical package (Fox and McGuire 1976) run on the MTS operating system (MTS 1976) at the Computing Centre at U.B.C.. The macros were all validated using literature examples (Reyment et al. 1984).

Some of the multivariate manipulative methods used here are generally not available in canned programs. Several references (Bookstein et al. 1985, Cooley and Lohnes 1971, Harris 1975, Manly 1987, Neff and Marcus 1980, Pimentel 1979, Srivastava and Carter 1983, Tabachnik and Fidell 1983) contain computer programs for some of these interpretive procedures and many also list the availability of stock programs. The size-constrained PCA (Somers 1986) with my verified corrections (Somers pers. comm.), and the sheared PCA (Humphries et al. 1981), are available from their respective authors. Copies of their original programs were used to verify mine.

Multivariate Procedures — Assessment Results

Based on Morphological Data

The EV1 loadings in figure 18 demonstrate a strong similarity between the covariance (labelled **A**) and sheared (**B**) matrices, and between the correlation (**C**) and size-constrained (**D**) matrices. All the EV1's are large general vectors. This suggests that they represent size. **A** and **B** account for marginally more variance in EV1 than **C** and **D**, with **B** accounting for the highest amount. **D**, of course, is isometric for this size vector, and **C** is already almost isometric itself.

The EV2 loadings in figure 18 are all smaller and bipolar (except maybe **D**) and thus overall they represent shape. The EV2 loading patterns and the overall variances they account for are very similar for **A**–**C**. There also is correspondence between EV1 and EV2 in these three procedures. When one variable is not strongly size-related (EV1) it usually loads strongly onto the shape axis (EV2) and vice-versa. This correspondence cannot be assessed for **D**. **D** is very different in all these respects, and there is no consistent pattern between **A**–**C** and **D**. The EV2 variables in **D** almost all load positively and it accounts for more overall variance than the other three techniques.

Furthermore, those characters which load strongly positive in **A–C** load strongly negative in **D**, and those which loaded strongly negative in **A–C** are not strongly loaded in **D**.

Figure 19 shows that the percentage variances of each of the original variables accounted for in EV1 and EV2 are virtually identical for **A** and **C**. Overall, **B** is similar but some specific differences exist. There also is correspondence between EV1 and EV2 for **A–C**. Again, **D** is very different. **D** accounts for a slightly lower mean level of variance in EV1 and of course is isometric here as well. The pattern in EV2 for **D** is somewhat similar to **A–C** but a much higher percentage of the variance of each character is accounted for. Once more, some variables most strongly accounted for in **A–C** are not strongly represented in **D**. Most of these are almost zero in **D**, and this pattern is even more inconsistent than it was for the EV loadings in figure 18.

The correlation patterns between the original variables and their eigenvector loadings in figure 20 are similar to the variance patterns in figure 19. **B** is even more like **A** and **C** in this case, and these three matrices result in virtually identical correlations. Again, there is correspondence between EV1 and EV2 for **A–C**. **D** is still very different, and once more those variables which correlate strongly in **A–C** do not necessarily do so in **D**. The EV2 correlation pattern in **D** is inconsistent and mostly positive. It is also not possible to determine if there is any correspondence between EV1 and EV2 in **D**.

The allometry coefficients in figure 21 reveal a somewhat similar pattern. **A** and **B** have allometry coefficients that are consistent with those of the original variables in figure 15 (see **chp. 5**). The overall picture in **C** is somewhat the same but there are subtle differences. All the **C** allometry coefficients are closer to isometry. This is true even if the punk plot for **C** is spread out on a smaller y-axis. Since EV1 for **D** is already isometric, its allometry coefficients are all one.

Table 2 reveals that EV1 in **A–C** still possesses allometric size, whereas in **D** EV1 is isometric. **C** is again the next closest to isometry. In all four procedures, the EV2 values of course are not isometric. These EV2 calculations are presented to demonstrate that EV2 is orthogonal ($\approx$ 90 degrees) to an isometric vector in **A–C**, whereas in **D** it is not. These EV2 calculations also are indicative of whether EV1 and EV2 are orthogonal in each procedure (**A–D**). They appear to be

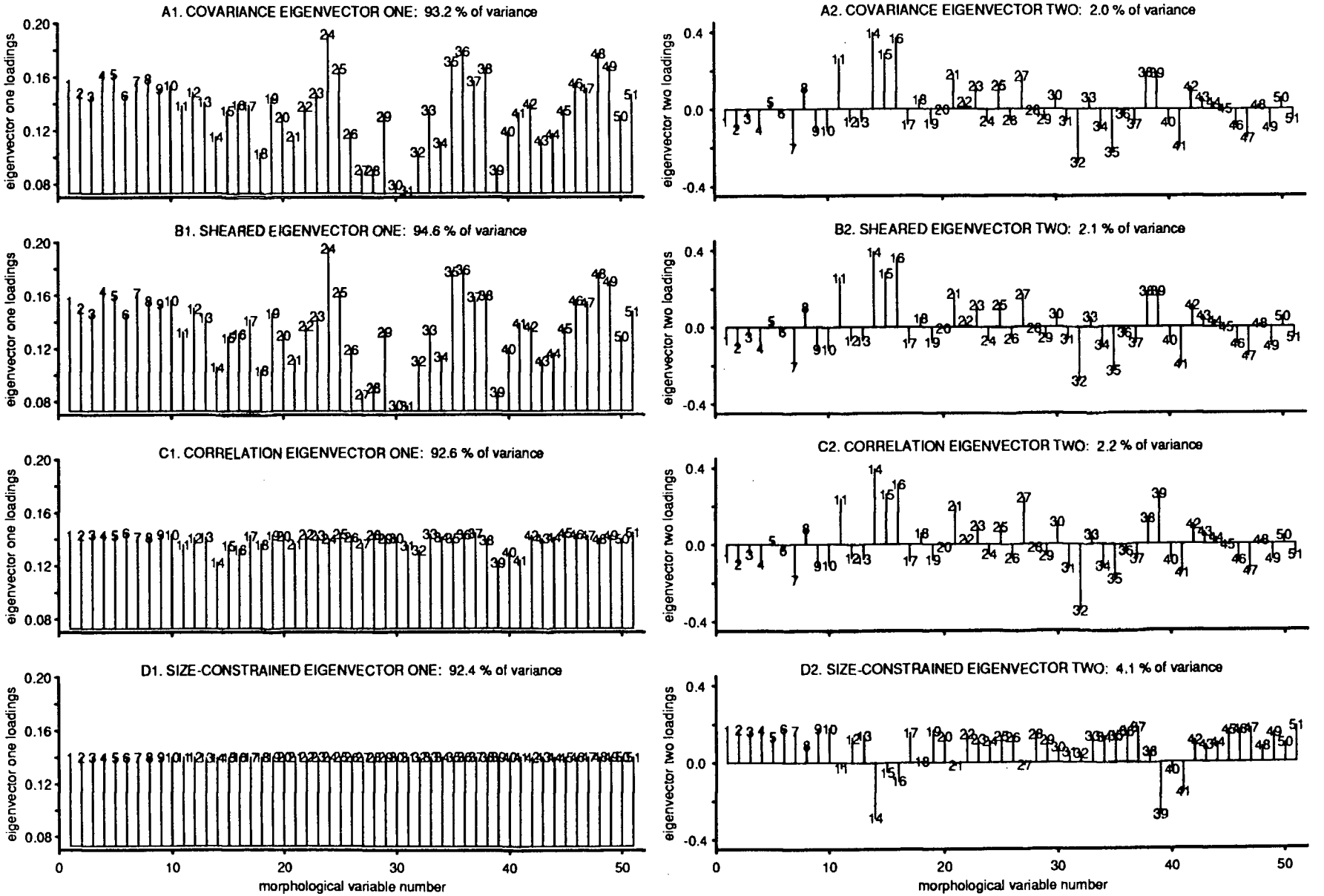


FIGURE 18. PCA patterns for morphological variables (extrapolated from or plotted about zero).

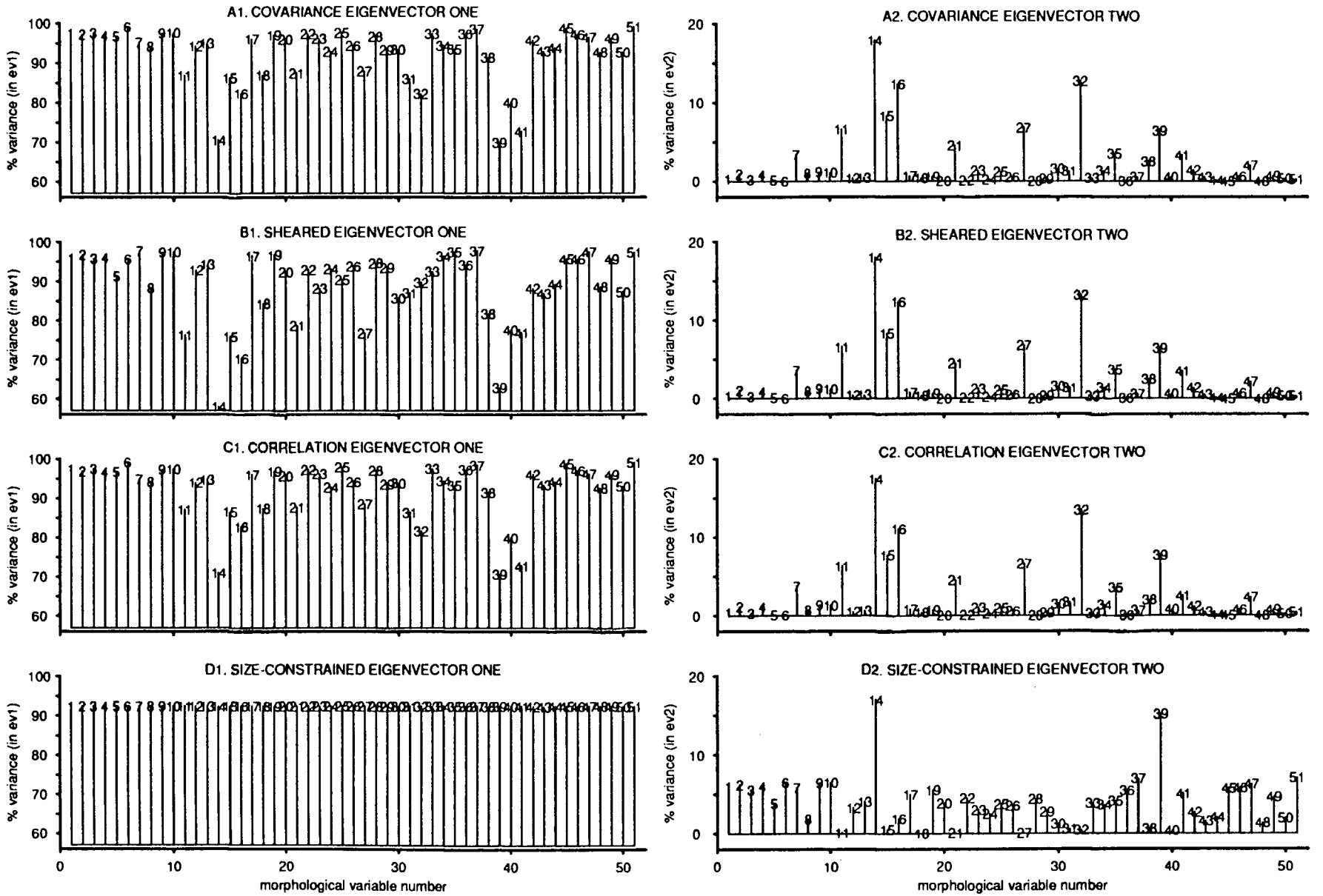


FIGURE 19. Patterns of % variance of each morphological variable accounted for in ev1 and ev2.

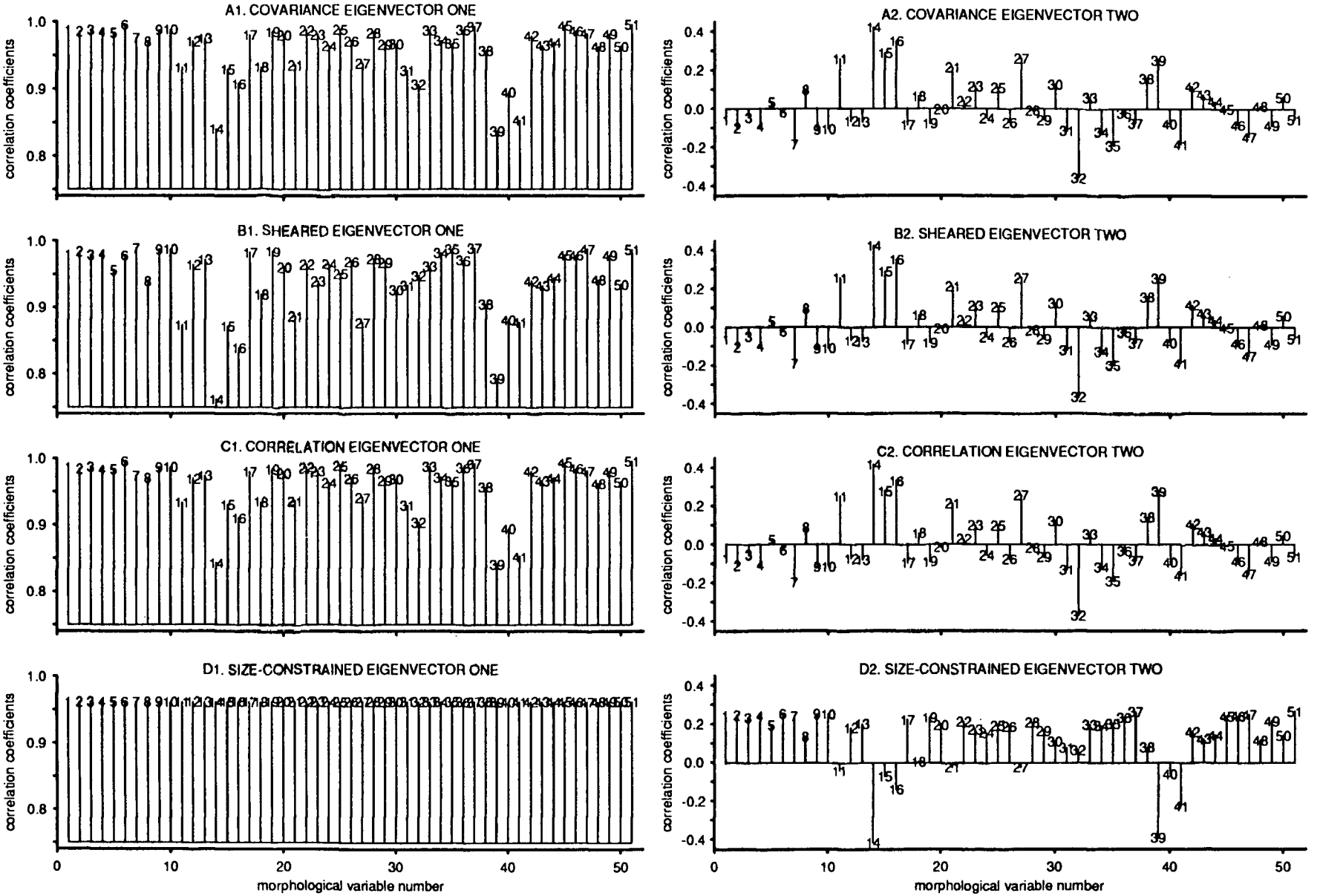


FIGURE 20. Patterns of correlations between morphological variables and their PCA loadings.

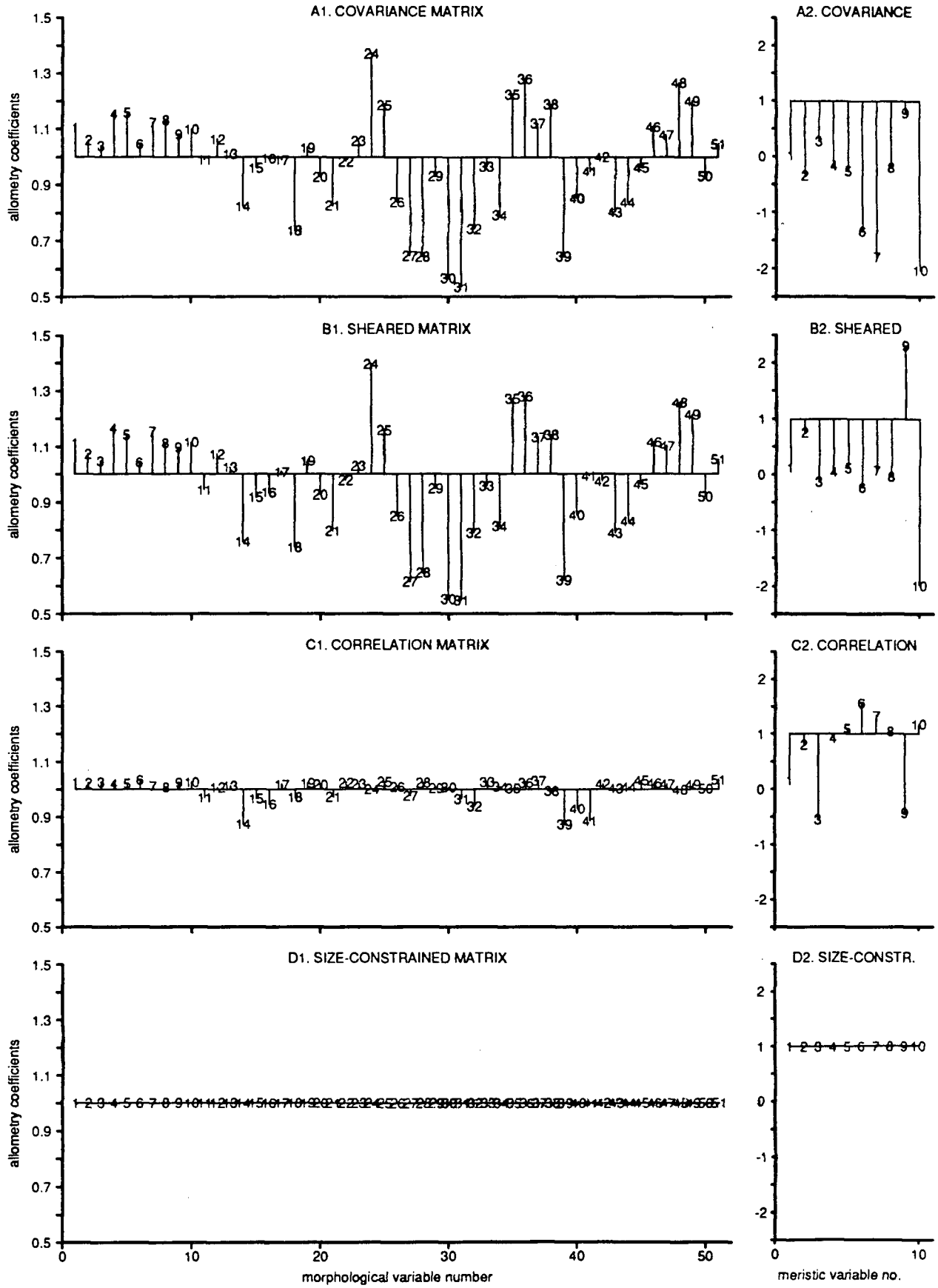


FIGURE 21. Allometry coefficient patterns for the PCA analyses.

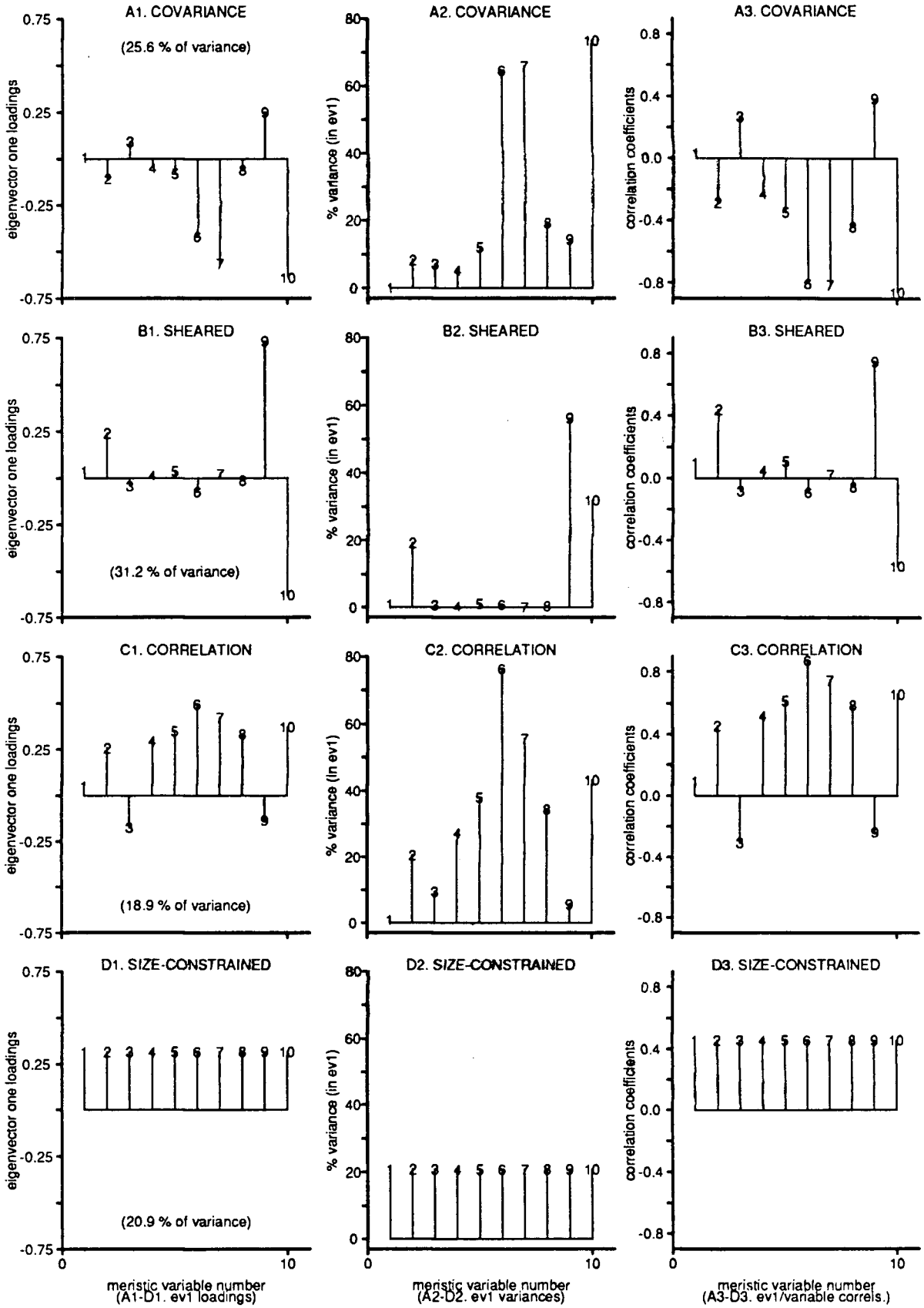


FIGURE 22. All PCA patterns for the meristic variables.

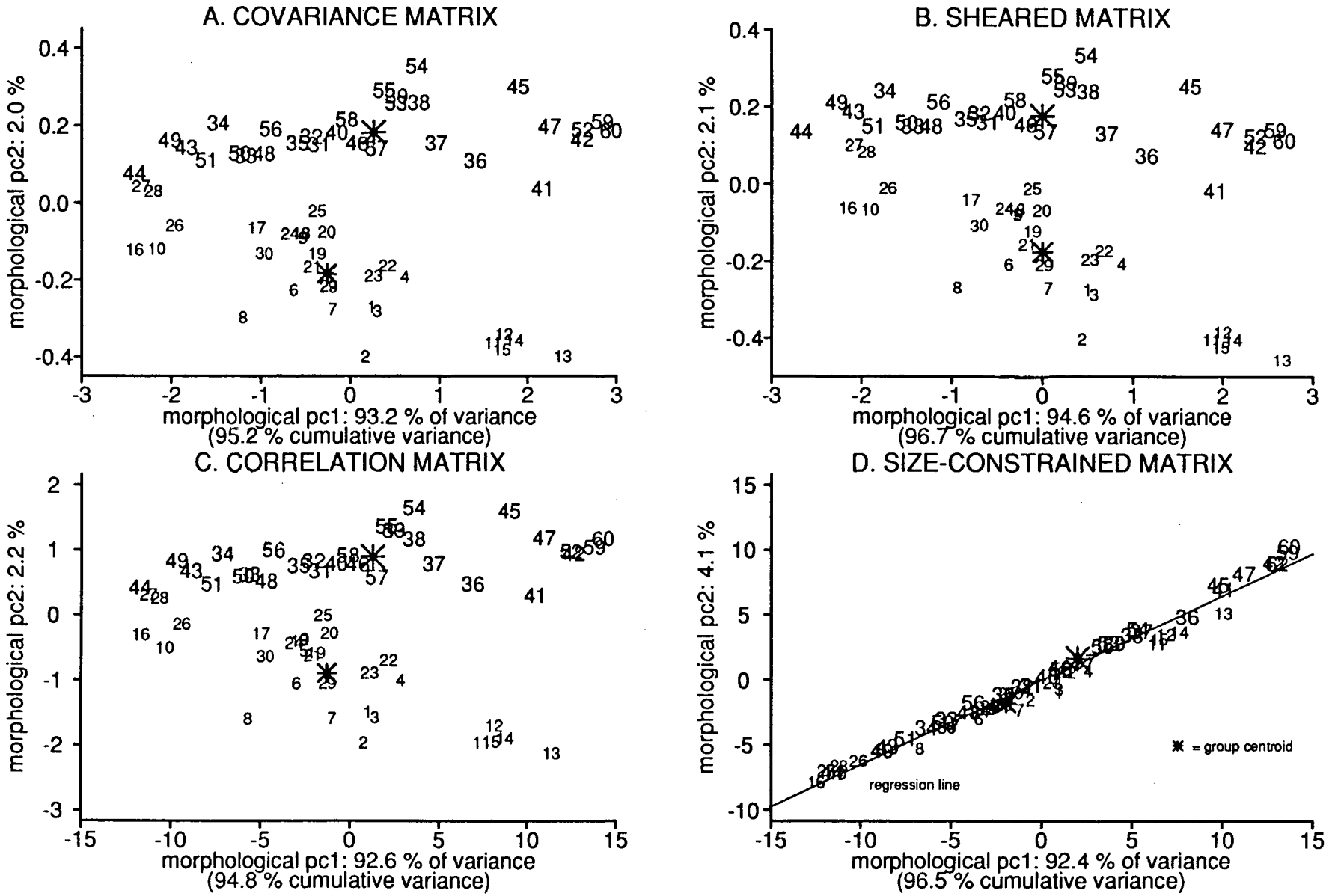


FIGURE 23. PCA patterns for individuals (scatter plots of morphological pc1 and pc2).

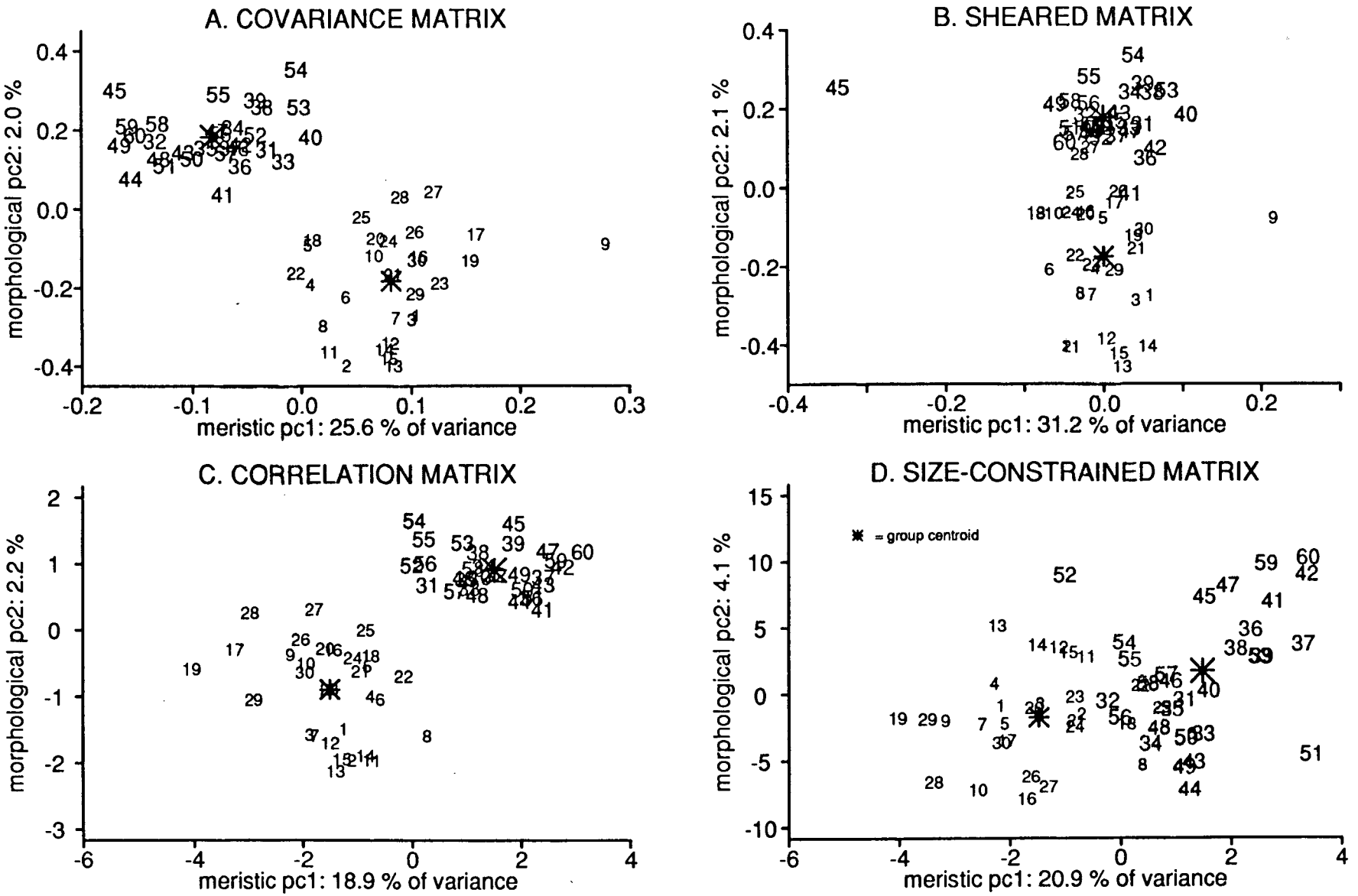


FIGURE 24. PCA patterns for individuals (scatter plots of meristic pc1 and morphological pc2).

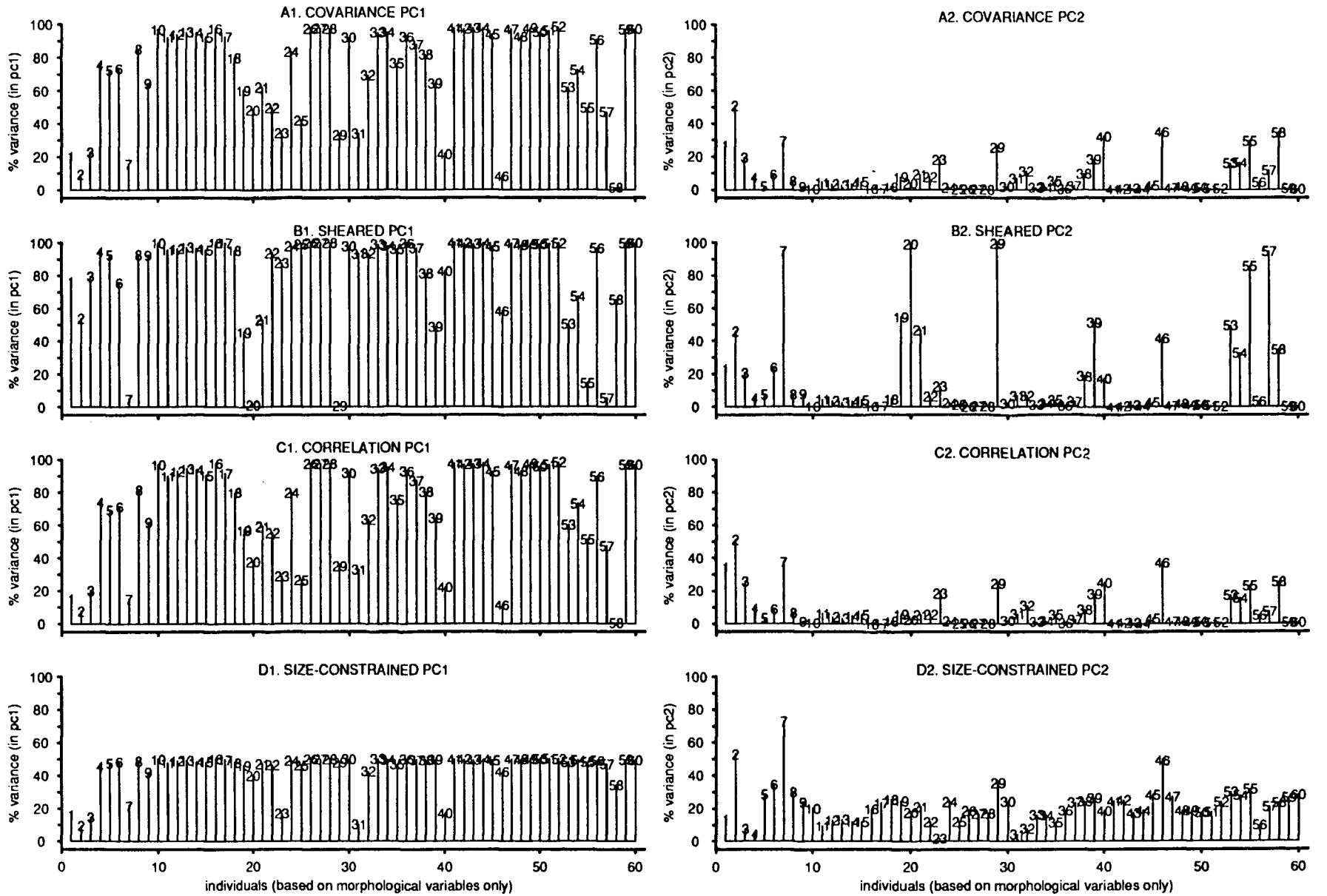


FIGURE 25. Patterns of % variance of each individual accounted for in pc1 and pc2.

Table 2. Isometry statistics for multivariate procedures.

			covariance (A)	shear (B)	correlation (C)	size-constrained (D)
morphology	χ-value	ev1	$p < 0.001$	$p < 0.001$	$p < 0.01$	$p > 0.999$
morphology	χ-value	ev2	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
morphology	degrees	ev1	10.7°	11.0°	2.2°	0.0°
morphology	degrees	ev2	87.8°	89.2°	89.2°	50.0°
meristics	χ-value	ev1	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p > 0.999$
meristics	degrees	ev1	119.9°	83.9°	44.2°	0.0°

orthogonal in **A–C**, but are not in **D**. This is unexpected in **D** because its EV1 is isometric and thus its EV2 should be orthogonal.

The scatter plots in figure 23 show that **A–C** provide effective group and centroid separation, and that **D** does not. This separation in **A–C** is exclusively along the shape (PC2) axis. The individuals also are distributed from small to large sizes within each group along the PC1 axis. The best group separation is in **B** where no numbers overlap, but it is only minimally better than in **A** or **C**. **C** gives somewhat better ordination of larger individuals but is the same as **A** in terms of smaller fish. The two axes in **D** are not orthogonal and thus while group separation is evident on either side of the regression line it is not very obvious or effective. **D** reveals no distinct size or shape information.

The punk plots in figure 25 show that the percentage variance of each original individual accounted for in PC1 and PC2 is similar for **A** and **C**. **B** also is not that different from **A** or **C**, but a few individuals do have an exaggerated pattern, and more variation is generally accounted for in PC2. Once again, **D** is odd. While the patterns for individuals in **D**, especially in PC2, are more consistent with **A–C** than for the previous character comparisons, the same type of differences are still present. In **D**, much less variation is accounted for in PC1 and much more in PC2. There is, however, no consistency between the PC1 and PC2 patterns for **D** whereas there is for **A–C**.

Based on Meristic Data

Figure 22 demonstrates that **A–C** are still usually more similar to each other for the EV loadings, percentage variances and correlations than to **D**. This weak comparability to the morphological data results ends here though and much stronger differences exist in the meristic **A–C** than did for morphology. The EV1 loadings in the first column are similar in trend for **A–B** but are still inconsistently different. **C** is very different and **D** is isometric. There is no indication of a large, general EV1 in **A–C** and the most variance accounted for is by **B** and is only 31.2 %. **D** is a general vector but also accounts for only 20.9 % of the variance. All the procedures account for a small amount of variance for an EV1.

The second column of plots in figure 22 shows that the percentage variance of each of the original characters accounted for in EV1 is generally not large and is inconsistent (is not surprisingly

consistent in isometric **D**). Again, this is unusual for morphometric EV1's. **A** and **C** have the most similar patterns here but this is only relative.

The correlations between the original variables and the EV1 loadings in column three of figure 22 are all unusual. There are both positive and negative correlations in **A–C** for EV1 and there is no consistent pattern between these three matrices. All the correlations are also very small. **D** shows all positive correlations but these too are weak.

The allometry coefficients in figure 21 reveal a similar incongruent pattern between the four matrices. The allometry coefficients in **A–B** are unusual and all are either near zero, negative or very positive. They also do not resemble those of the original data (fig. 15; **chp.5**). The coefficients for **C** are somewhat more realistic but their overall pattern is often inconsistently opposite to that in **A–B**. As well, two of the coefficients in **C** are negative. **C** is most similar to the raw data but this similarity is only relative and not strong. The allometry coefficients for **D** are of course all one since it is isometric.

The isometric comparisons in table 2 demonstrate that all the EV1's in **A–C** are still allometric. **C** is again the closest of these three matrices to being isometric. Naturally, **D** is isometric.

The scatter plots of morphological PC2 against meristic PC1 in figure 24 reveal effective group and centroid separation in **A–C** and partial separation in **D**. In **A** and **C** this separation is on both axes. In **B** the separation is only on the morphological axis, and in **D** is on the meristic axis. **B** is unusual in that it has two outlying individuals which are not readily apparent in the other three plots (except maybe individual no. 9 in **A**).

Multivariate Procedures — Assessment Discussion

Based on Morphological Data

The similar overall results for the covariance (labelled **A**), sheared (**B**) and correlation (**C**) matrices procedures are discussed together first. Specific deliberation is then devoted to the size-constrained matrix technique (**D**) as it is obviously different. The general similarity between the standard covariance and correlation PCA procedures based on allometric data is well known (Boratynski and Davies 1971, Holmes 1975, Leamy and Bradley 1982, Pimentel 1979, Shea 1985),

but in this study both techniques have certain specific effects. Furthermore, the general agreement found here between **A–C** suggests that these PCA procedures are producing realistic output.

The effects of **A–C** on the ordination of individuals and on the separation of size and shape (fig. 23) are quite similar. This is true of their individual variance patterns (fig. 25) as well. All three approaches result in very effective group and centroid separation that is based entirely on shape information. The size information on PC1 is also realistic since it follows that expected by the actual fish size distribution. The shear matrix may be a slightly more effective technique, at least for smaller individuals, but this potential improvement is offset by its more complicated and non-standard matrix algebra approach to PCA. The shear technique also produces some slightly exaggerated individual variance patterns (fig. 25). These patterns may not be overly important but should still be considered in evaluating an allometric procedure. The use of covariance and correlation matrices provides virtually the same group ordination results without invoking more complicated mathematics. Their procedures also are standard, readily available and already programmed.

The effect of **A–C** on characters is also often quite similar, but some differences exist. Their EV1's are large and general (fig. 18) and definitely represent size. This is confirmed by the PCA scatter plots (fig. 23). Their EV1's, however, do not appear to contain all the available size information that is in the data as none of them are statistically isometric (table 2). This leftover size information ends up in EV2 and subsequent EV's.

The consequence of this remaining size is the central issue regarding the effectiveness of these morphometric techniques (Archie 1987, Bookstein et al. 1985, Pimentel 1979, Reyment et al. 1984, Rohlf and Bookstein 1987, Somers 1986). In my opinion, this remaining size is not a problem since it appears that this leftover size information contributes to that of shape. This may sound a little confusing, but shape should still possess some size information if it is to have any meaning itself. Therefore, arguments that EV1 in standard PCA (**A** and **C**) is not isometric and that EV2 thus represents shape information that is confounded by size seem wrong. The PCA scatter plots (figs. 23 and C) reveal that all the individuals are correctly ordinated to their groups based on this standard "confounded shape" vector. Any size information remaining in EV2 is not displayed as smaller and larger individuals but rather as individuals of different shapes.

Isometry is an ideal that is hard to envision and most likely rarely exists. Isometry would exist if organisms of the same size in the same groups have exactly the same growth rates (Burnaby 1966, Shea 1985) and shape. This seems improbable and is not present in this study. Indeed, the allometry coefficients in figures 15 (**chp. 5**) and 21 seem to indicate that it does not even exist for measures which are strongly size-related (eg. body depth (morphology no. 1) and body width (morphology no. 2)). It certainly does not exist for characters which are not strongly size-related (eg. eye size (morphology no. 30)). This lack of isometry is even more pronounced in multivariate procedures (fig. 21) than with bivariate techniques (fig. 15; **chp. 5**). There will nearly always be some size remaining after EV1 and this size will end up in the shape EV. It does not confound that EV though, but rather contributes to the shape information and individual ordination in a more realistic way than if it is removed through further manipulations.

The EV2's in **A–C** (fig. 18) are all smaller than their EV1's and are bipolar, yet they still contain significant variation. They are thus definitely shape vectors. This shape assessment is supported by the PCA scatter plots (fig. 23) as there is no relationship to size in the shape PC2 axis. There also is direct correspondence between EV1 and EV2. When one of these two EV's has a strong relationship in terms of loadings (fig. 18), percentage variance accounted for (fig. 19) or with original character correlations (fig. 20), the other EV is weak and vice-versa.

The allometry coefficients (fig. 21) reveal a consistent pattern in **A–C** but the correlation matrix is much closer to isometry. The proximate isometry in the correlation matrix can also be seen in its table 2 isometry statistics. However, the covariance and shear matrices are more similar to the allometry coefficients for the original data (fig. 15; **chp. 5**) and thus also are more realistic and informative of the allometric relationships which exist.

The size-constrained matrix (**D**) attempts to force all the size information to be isometric in EV1. This obviously results in a very different outcome for both individuals and characters. The PCA scatter plots for **D** (fig. 23) are very ineffective in individual ordination and group/centroid separation. Furthermore, figure 23 and table 2 demonstrate that the first two PC's and EV's are not orthogonal. The little information that can be derived from this procedure cannot then be distinctly prescribed to uncorrelated size or shape.

This non-orthogonality results from the first isometric vector being removed before the standard PCA (Rohlf and Bookstein 1987, Somers pers. comm.). Consequently, the size-constrained procedure does not extract orthogonal or maximum information in EV1 because only isometric size is removed. All the allometric size, notably that which confounds the shape information and is correlated to isometric size, remains. This leftover size ends up in the second and subsequent vectors, and they are then not orthogonal to the first EV. The second EV's and PC's, however, are orthogonal because they are the result of standard PCA. This difficulty with isometric size in **D** is further evidence against the existence of isometry in this data.

The absence of good size and shape information in **D** is also evident in the size-constrained EV2 loadings (fig. 18). They are not strikingly bipolar and it is possible that they do not represent a true shape EV. The size-constrained matrix results in data which often are very inconsistent with the other three procedures (figs. 18–20). It also is not possible to tell whether there is any consistency between the EV1 and EV2 character patterns for **D**, but there is no relationship present between the PC1 and PC2 scores for individuals (fig. 25). As well, the allometry coefficients for **D** (fig. 21) are uninformative because of the forced isometry.

The most appropriate multivariate morphometric procedure for morphological data is PCA on a covariance matrix (procedure **A**) usually based on logarithmic transformed data (Bookstein et al. 1985, Corruccini 1983, Humphries et al. 1981, Jolicoeur 1963a, Pimentel 1979). The first two eigenvectors resulting from **A** definitely correspond to size and shape, and are effective in ordinating individuals and in separating groups and their centroids. The character information it results in is realistic and portrays the allometric relationships that exist in the original data. This technique is also standard and readily available.

PCA on a covariance matrix may only be the best morphometric procedure if it is ased on data that has sufficiently standardized variances. Different character variances will adversely affect PCA results because some characters may load inordinately heavily because of their high variances and not as a result of any important biological features (Eisenbis et al. 1973, McKay and Campbell 1982b, Neff and Marcus 1980, Pimentel 1979, Reyment et al. 1984, Weiner and Dunn 1966). Sufficiently equal character variances should be present in morphometric data, however, because similar measurements usually have proportional variability (Thorpe 1983a). If morphometric characters

have unequal variances, logarithmic transformation of the data usually effectively standardizes it (Bookstein et al. 1985, Humphries et al. 1981, Jolicoeur 1963a). If sufficiently unequal variances still exist after transformation, remove the offending characters or analyze them separately (Pimentel 1979, Thorpe 1983a).

If this recommendation is insufficient, PCA on a correlation matrix (**C**) usually based on logarithmic transformed data may produce better, but still comparable, results. It standardizes the data more effectively (to mean zero and unit standard deviations) and the character contributions are then more equal (Davis and Baker 1974, Pimentel 1981, Reyment et al. 1984, Thorpe 1976, 1980, 1983b). Their allometric relationships, however, will probably not be portrayed as realistically. A suggestion here may be to try using both the covariance and correlation matrices in the analysis. If the correlation matrix produces more reasonable results use it, otherwise use the covariance matrix (Brown and Davies 1974, Pearce and Holland 1960). PCA on a correlation matrix also is a standard and readily available procedure.

The other argument sometimes offered in support of PCA on a correlation matrix is that it is supposedly more effective at removing size (Boratynski and Davies 1971, Brown and Davies 1974, Marriott 1974, Rohwer 1972, Somers 1986, Teissier 1960). This study supports this conclusion insofar as isometric size is concerned, but this additional removal of isometric size seems unnecessary and has no apparent advantage (Bookstein et al. 1985, Jolicoeur 1963a). The correlation matrix procedure in this study does not lead to better individual ordination or group/centroid separation. It also provides less character information than the covariance matrix technique and does not portray the data relationships as realistically.

A final consideration is that this study is based on homoscedastic groups analyzed in a total matrix PCA (see **chp. 4**). If heteroscedastic groups are present, they may require a pooled within-group PCA instead of this total matrix analysis. If a pooled within-group approach is necessary, the shear procedure (**B**) may be more appropriate. The shear procedure is based on a $\log_{10}$ transformed covariance matrix but accounts for within-group size. Therefore, it has the advantages of the covariance approach yet permits within-group size differences to be dealt with. This may be of value in an analysis of heteroscedastic groups.

Based on Meristic Data

There is a weak similarity between the multivariate analyses of the meristic and morphological data. There again are some consistent results between the covariance (**A**), shear (**B**) and correlation (**C**) matrices, with the size-constrained matrix (**D**) producing different results. There also are many more differences in the meristic **A–C**, however, than there are with the morphological data. These differences result from the meristic data and suggest that in this study their suitability for multivariate morphometric procedures is not as good as morphological data. The multivariate analyses of this meristic data cannot be interpreted in the same manner as morphological data. PCA can only adjust the meristic characters for allometry and define their size/shape relationships if these characters have size information (Reist 1985, Somers 1986). The meristic data in this study do not have size information (fig. 13(vi); **chp. 4**).

The meristic data, therefore, cannot be interpreted like the morphological data because PCA based on them will not reveal size and shape information. Individual or group ordination based on this meristic data occurs on PC1 (figs. 13 and 24) because it contains the greatest amount of variation present in the original data and does not represent size. PC1 is now analagous to a "shape" vector for meristics since much of the important ordination information is summarized in it and EV1 is no longer large or general (fig. 22). The first EV and PC are forced to be general in the size-constrained method but they are still not large. This forced general vector effect is seen throughout the size-constrained method analyses but does not generate a size vector for this meristic data.

While EV1 and PC1 do of course contain the maximum possible amount of variation present in the original data, this level is still greatly reduced (figs. 18 and 20). The reduced variance accounted for also causes many more eigenvectors to be significant. This complicates the interpretation of the PCA results and decreases the advantage of data sysnthesis in multivariate morphometrics. In short, multivariate morphometrics based on meristic data may effectively ordinate individuals and separate groups, but their effects on meristic characters are difficult to interpret and are unusual. Accurate portrayal of characters requires that something of their true nature exist after multivariate manipulations but this is not easy to assess in the meristic case. Furthermore, the

ordination that does take place with meristic data does not require multivariate statistics (fig. 17; see **chp. 5**).

The percentage variance of each of the original variables accounted for is low and inconsistent (fig. 22). The correlations between the original characters and EV1 are also very weak, and most are both positive and negative (fig. 22). These character effects can be seen for **A–C** and are very atypical for multivariate morphometrics. They suggest that the character information resulting from the PCA is not well synthesized or realistic. Moreover, each technique still results in additional specific differences and this causes further suspicion of their utility.

As well, the allometry coefficients (fig. 21) are all quite unusual. The correlation matrix provides the most believable alloemtry coefficients and these are the most similar to the original data (fig. 15; see **chp. 5**). This resemblance is still not complete though, but the other three matrix procedures result in allometry coefficients that do not resemble those of the original data values at all.

The plots of morphological PC2 against meristic PC1 (fig. 24) provide the only highlight in the meristic multivariate analysis. While these scatter plots still not provide realistic size information for characters or individuals, they do result in effective ordination. If ordination is the only objective, these scatter plots would work well. The meristic component makes a good discriminating axis, and the morphological PC2 shape variables are still authentic and can be interpreted as morphological shape data.

If meristics are still to be entered into PCA, use a correlation matrix (**C**) for their analysis. The correlation matrix will better standardize the data and this study demonstrates that **C** has the least negative effects on meristic data. This recommendation is relative, however, because even the correlation matrix results are not completely realistic or similar to the usual size/shape interpretation of multivariate morphometrics. Keep these effects in mind if meristic characters are analyzed in PCA and analyze them separately from the morphological variables. The best advice for morphometric studies, however, may still be not to use morphometric procedures on characters which contain no size information (also see **chp. 5**).

Multivariate Procedures — Summary

1. Individual ordination and group/centroid separation is effective and very similar for the covariance (**A**), shear (**B**) and correlation (**C**) matrices.
2. The character information resulting from these three procedures is generally similar but specific differences exist.
    a) These character differences result from how the procedures remove the size information present.
3. PCA based on a covariance matrix (usually of $\log_{10}$ transformed data) is the best multivariate morphometric procedure.
    a) Its first principal component and eigenvector represent size, but this size is not isometric.
    b) The second principal component and eigenvector represent shape. They still has some of the leftover size information present but this size is not confounding.
    c) PCA based on a covariance matrix of $\log_{10}$ transformed data results in character and individual data which are more realistic and representative of the original data relationships than the other three multivariate morphometric techniques.
    d) It also is standard, readily available and already programmed.
4. If data variances are not standardized, even after logarithmic or other transformations, verify that PCA on a correlation matrix (**C**) is not a better procedure. Alternatively, remove or separately analyze the unstandardized characters.
5. If desired or heteroscedastic groups exist in the data and require a pooled within-group analysis the shear procedure (**B**) may be better.
6. The size-constrained matrix (**D**) is unrealistic and produces very different results.
    a) Its first two vectors are not orthogonal.
    b) They also result in inconsistent and uniformative data for characters and individuals, and do not represent size and shape components.

7. If meristic or other characters are not size-related, do not enter them into PCA for morphometric interpretation or allometric adjustment.

    a) Such a PCA does not result in size and shape vectors.

    b) It still results in excellent individual ordination and group separation but only along the first vector. Its character information is also not completely representative.

8. Meristic characters can still be entered into PCA for data synthesis but even here several additional resultant vectors will still be significant and the synthesis is thus less useful.

    a) Use PCA on a correlation matrix for such an analysis.

    b) Analysis of the original meristic data is much easier, however, and often it may be just as useful.

CHAPTER SEVEN

Principal Component Analysis on Bivariate Adjusted Data

Introduction

Many morphometric studies use or recommend (eg. Rohlf and Bookstein 1987) multivariate statistical analyses on data which are already at least partially adjusted for allometry through previous bivariate manipulations. This practise is suggested in order to remove size information from the data before entering it into a multivariate analysis. This is generally said to help minimize size differences in the data and maximize the resultant ordination.

This removal of size does not seem to be an appropriate or sufficient reason to enter bivariate adjusted data into multivariate morphometric techniques. The multivariate procedures already adjust the original data for size, and ordinate the individuals and groups, without previous bivariate morphometric manipulations. Moreover, previous aspects of this study (**chps. 5–6**) suggest that multivariate methods do so more effectively and realistically than bivariate procedures. This is, however, unclear since the practise of entering bivariate adjusted data into multivariate morphometric techniques continues.

The application of multivariate procedures to bivariate adjusted data results in problems which otherwise do not exist if the multivariate techniques are carried out on the unadjusted raw or transformed data. These difficulties will be demonstrated here.

PCA on Regression Data — Assessment Methods

The problems of applying multivariate statistics to previously bivariate adjusted data are demonstrated with principal component analysis (PCA) based on a covariance matrix of $\log_{10}$ transformed regression adjusted data. This study (**chps. 5–6**) has demonstrated that these two morphometric procedures are the best multivariate (chp. 6 technique **A**) and bivariate (chp. 5 technique **e**) approaches. They should therefore provide the best opportunity for a multivariate analysis of bivariate adjusted data to produce the results expected of such an approach.

This analysis is based only on the fity-one morphological characters. This character choice again follows from the previous analyses (**chps. 5–6**) which suggest that the meristic data are not

confounded by size and thus do not require adjustment for allometry. These meristic characters can be better analyzed without morphometric procedures and confound the morphometric interpretations if they are combined with morphological data (see **chp. 5**) All the morphological characters are numbered 1–51. The sixty individual fish analyzed form two groups which are each respectively represented by the numbers 1–30 set in small type and the numbers 31–60 set in large type. The centroids (group means) are also plotted in small and large sizes. PCA, and its terminology and formulas, are discussed in detail in **chapter 6**. The regression technique is similarly explained in **chapter 5**. Figure 26 presents the results for individuals and figure 27 for variables.

Figure 26(i) is a scatter plot of the second principal component (PC2) against the first (PC1). It demonstrates on which axis and how effectively the individuals are ordinated and the groups and centroids are separated. Each axis also gives the PC scores for each individual, and these can be visually assessed by looking at each axis from the perspective of a punk plot (for term see **chp. 5**). Traditional morphometric interpretation of PC1 is that it is a size vector and that PC2 is a shape vector (see **chp. 6**). This plot also reveals whether this interpretation is valid here, and whether both size and shape information are still obtained from a multivariate analysis of bivariate adjusted data.

Figure 26(ii) is a plot of a Scree test (see **chp. 4**) which qualitatively reveals how many eigenvalues are significant. Eigenvalues represent the overall variation accounted for by each eigenvector and PC. If an insufficient amount of variance is accounted for by the initial vectors then more vectors probably require interpetation in order to completely understand the multivariate results. Significance is usually assigned to those eigenvalues above the plot region where the curve asymptotes and stabilizes. If an eigenvalue is significant, it is assumed to represent non-random, relevent information. Significant eigenvectors should thus form part of the interpretation of the study and not simply be discarded in a traditional multivariate morphometric analysis of only the first two or three vectors. Bartlett's test of sphericity (see **chp. 4**) is also applied here to quantitatively assess and verify how many eigenvectors are significant.

Figures 26(iii)-(iv) shows how much of and where the variance of the original individuals is accounted for in each of the first two PC's. These plots demonstrate whether all the individuals are being sufficiently represented in certain PC's, and also whether a PC has significance and should be

analyzed. If the representation of individuals is very inconsistent, they are not being accounting for in the same PC's. An analysis of individuals based on any one or two PC's is then not completely representative of the individuals or the groups they belong to.

Figure 27(i) presents the isometry statistics for the first two eigenvectors (Pimentel 1979; also see Leamy and Bradley 1982, Somers 1986, Thorington 1972). The "degrees" from isometry show how many degrees each of the eigenvectors are from a theoretical isometric size vector. The p-values presented are the chi-square values derived from Anderson's (1963) test. A statistically significant value ($p < 0.05$) indicates that isometry is not achieved for that eigenvector and that some allometric size information remains in that eigenvector. In this multivariate analysis of bivariate adjusted data, size information is supposedly already removed by the bivariate technique so all the eigenvectors should be isometric. Both these values also demonstrate how much size is actually removed by the bivariate technique as shape eigenvectors (eigenvector 2) should be nearly orthogonal ($\approx$90 degrees) to the size vector even if isometry is not present (eg. table 2; **chp. 6**).

Figures 27(ii)–(iv) respectively present the first two eigenvector loadings, the percentage variance of each original character accounted for in the first two eigenvectors, and the correlation of each original character with the first two eigenvectors. These punk plots are all extrapolated from or centred about zero and share equivalent axes within each dual set of figures. This permits assessment of the overall and individual character patterns. These plots demonstrate whether all characters are sufficiently represented in the first two eigenvectors and also whether these eigenvectors have significance and should be analyzed. If the representation of characters is inconsistent, then they are not accounted for in the same eigenvectors. If the loadings, variances or correlations are low, then they are not fully represented in those eigenvectors and more eigenvectors may need interpretation to explain the results of the multivariate analysis. The designation of a multivariate size vector as a large general eigenvector, and of a multivariate shape vector as a smaller bipolar eigenvector, can also be assessed through this set of figures.

The homoscedasticity of the groups based on the $\log_{10}$ transformed regression adjusted characters is assessed using Box's (Box 1954, Pimentel 1979) test. The univariate normality of this data

is tested using the probability (quantile) plot correlation test and the multivariate normality by a qualitative probability plot. These tests are further discussed in the **chp. 4.**

These analyses and graphics are based on computer macros I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Centre, University of British Columbia. These programs are available from me.

PCA on Regression Data — Assessment Results

Figure 26(i) reveals reasonably effective individual ordination and group/centroid separation along PC1. There is some overlap of the groups in the centre of PC1 but these individuals are still separated on the PC2 axis. There is no real separation of either groups or centroids on PC2.

The Scree plot (fig. 26(ii)) suggests that at least the first 6 eigenvalues are likely significant. Bartlett's test quantiatively verifies this estimate.

Figures 26(iii)-(iv) show how much of the variance of the original individuals is accounted for in each of the first two PC's. PC1 accounts for 0–73 % of the variance and PC2 for 0–55 %. There is no corresponding relationship between the two PC's. Here, when one PC accounts for much variation, the other PC does not necessarily account for proportionately less or vice-versa. Some individual variances are also almost unaccounted for in either PC and in the two PC's combined no individuals are cumulatively accounted for above 75 %.

Figure 27(i) demonstrates that neither of the first two eigenvectors are isometric. In multivariate morphometrics, the first eigenvector usually represents size information, but in this case it is not even close to isometric size. The second eigenvector in multivariate morphometrics is usually associated with orthogonal shape information, and should be approximately ninety degrees off a size vector from which confounding allometric shape has been removed. This is not the case here. Since the first eigenvector is previously adjusted for size it may represent shape. It also is not orthogonal to the isometric size vector, and therefore probably does not represent shape unconfounded by size information.

The variable loadings and overall percentage variance that each of the eigenvectors in figure 27(ii) accounts for indicate that they both are small and bipolar. The loadings for eigenvector

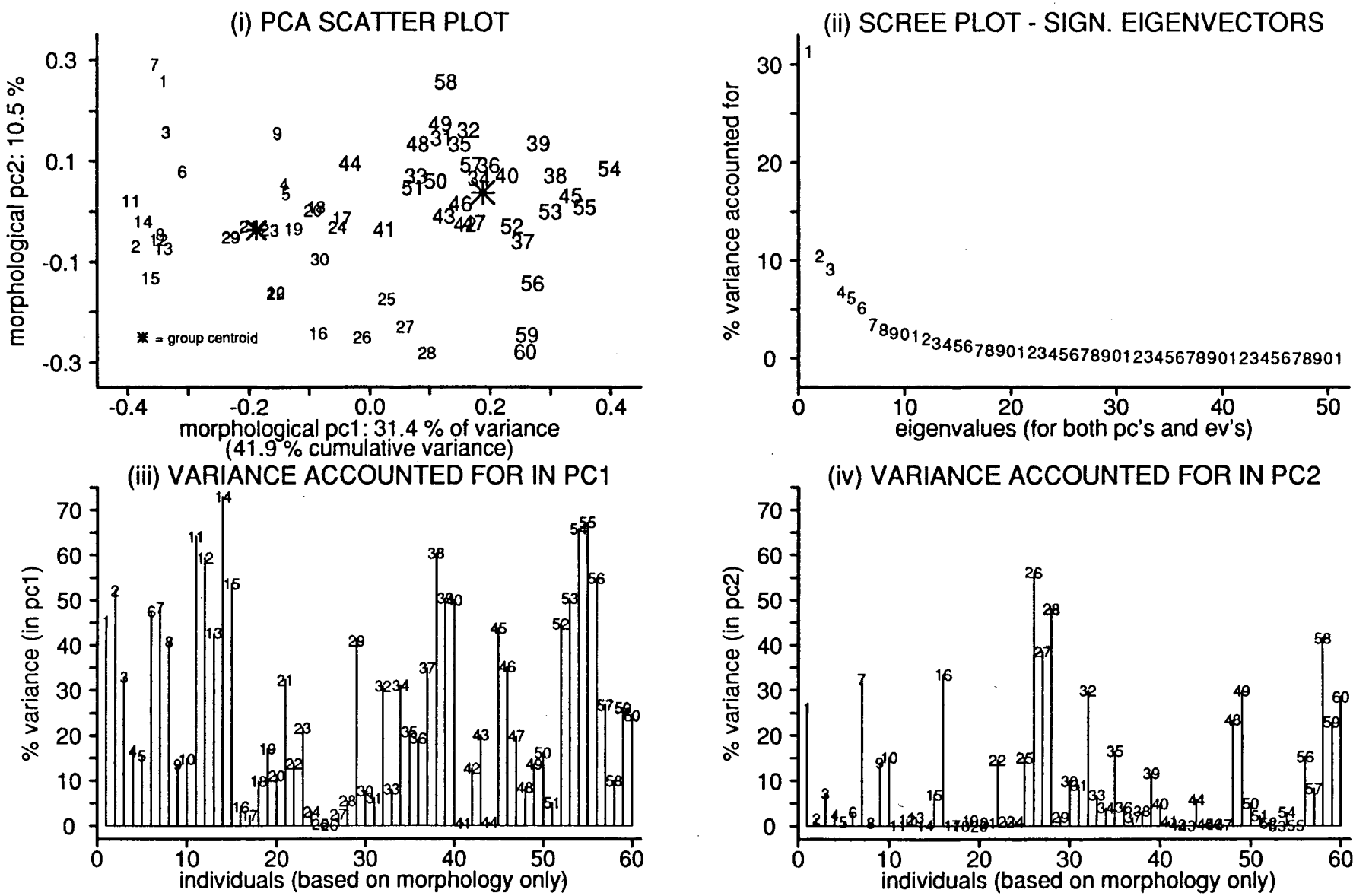


FIGURE 26. PCA patterns for individuals based on regression adjusted variables.

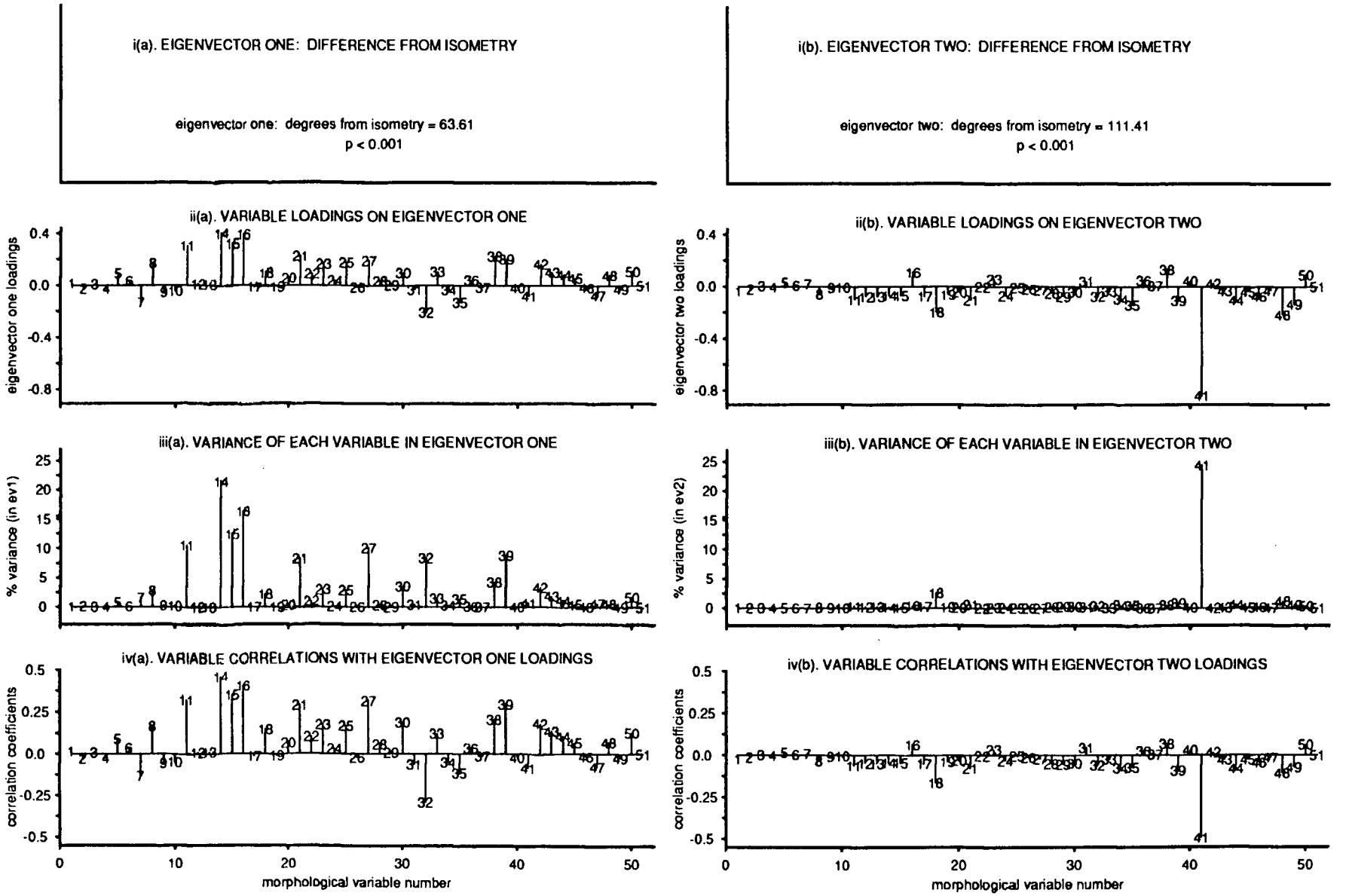


FIGURE 27. PCA patterns for regression adjusted variables (eigenvectors one and two).

one are also not consistent and this is unusual for the first eigenvector in a multivariate morphometric analysis. The second eigenvector loadings also are negative and near zero, and this is also odd.

The percentage variance of each of the original characters accounted for in the first two eigenvectors (fig. 27(iii)) demonstrate a similar picture to the loadings. The first eigenvector does not account for much variance, and this pattern is inconsistent. The second eigenvector accounts for almost no variance, except for one very unusual outlying character (no. 41). The amount of variance accounted for is generally very low and there is no corresponding pattern between the two eigenvectors.

Figure 27(iv) again reveals a similar picture. The correlations of the original variables with the first eigenvector are both positive and negative, and none are very strong. The second eigenvector correlations are all near zero and mostly negative, except again for the same outlying character. There is no consistent pattern within or correspondence between either eigenvector.

The groups are homoscedastic based on the $\log_{10}$ transformed regression data ($p > 0.5$). The $\log_{10}$ transformation also results in marginally improved univariate normality (from 37 untransformed (fig. 13e; **chp. 4**) to 41 transformed normal variables) and reasonable multivariate normality is only attained with this transformed data. The untransformed regression data is otherwise multivariately non-normal. Other conditions of the data such as scale and variance (see **chp. 4**) are also standardized by this $\log_{10}$ transformation.

PCA on Regression Data — Assessment Discussion

The two main results of this PCA on previously bivariate adjusted data are that the synthesis of data by the multivariate analysis is greatly reduced and that the resultant character output is no longer interpretable in a traditional morphometric manner. Neither of these features are desirable. This undesirability will be further demonstrated by comparing these results with those for standard PCA on a covariance matrix of $\log_{10}$ transformed original data. The results for the standard PCA are presented in the **chapter 6.**

At least six of the vectors (fig. 26(ii); and Bartlett's test) resulting from the PCA on bivariate data are significant whereas the standard PCA has only two significant vectors. This six vector

significance here is further confirmed by the consistently lower eigenvalues (percentage variances) of the PCA on bivariate data (fig. 26(i) and 15(ii)) and by how the variances of the original individuals are inconsistently accounted for (figs. (26(iii)-(iv)). The analysis of the PCA on bivariate data should then actually be on six of the resultant vectors if all the significant information resulting from this PCA is to be used and accounted for. If the analysis of the results is on fewer vectors, the interpretation is probably incomplete since statistically significant features of the data are being left out. The standard PCA only requires the analysis of two vectors. This standard analysis is therefore much simpler, and is also easier to accomplish graphically.

The results of PCA on bivariate data can also no longer be interpreted in a traditional multivariate morphometric manner. The first vectors no longer correspond to size information and the second vectors do not contain all the significant shape information (figs. 26(i) and 27). In fact, the second vector does not seem to contain much definitive information at all and it definitely does not correspond to good shape data (fig. 27). These effects greatly diminish the advantages of multivariate morphometrics because such a size/shape interpretation is of great value. As well, the standard PCA procedures for attaining such data are effective.

The individuals, groups and centroids are reasonably well separated by PCA on bivariate adjusted data but this separation is only on PC1 (fig. 26). The PC1 ordination, however, is not complete since there is overlap, and PC2 is thus required for complete group separation. This is further evidenced by the inconsistent individual variance patterns in figures 26(iii)-(iv). In these figures, bivariate PC1 resembles a shape vector since univariate size has already been removed by the bivariate procedure, but its shape representation is not as effective as that resulting from standard PCA (figs. 23 (**chp. 6**) and 15(i); table 2 (**chp. 6**)). The standard PCA ordination is better and more interpretable. This is probably because some of the leftover confounding multivariate size information remains in that bivariate PC1 (fig. 27(i)) since that is where this size information would have ended up in standard PCA.

Since bivariate PC1 is a confounded shape vector, no size information is directly obtained from this PCA on bivariate adjusted data. There is no multivariate size factor present and thus there also is no size information for each individual variable. Standard PCA places all the confounding size information in this first multivariate size factor, and thus provides complete and uncorrelated

size and shape information. The output of PCA on bivariate adjusted data can then no longer be characterized as size and shape, and these two factors remain confounded and have their information spread throughout several additional vectors.

There appears to be little reason to use PCA on data that has already undergone bivariate adjustment. PCA on the original data is simpler, and results in more readily interpretable and informative output. Nothing appears to be gained by the additional step and complication of size-compensating the characters prior to using a multivariate technique that will do it already. The reasons suggested by others for carrying out this prior bivariate adjustment are to help with large size differences in the data and to enhance group ordination. Neither of these objectives are achieved. In fact, less information is obtained and it is not as good or as useful.

A somewhat more justifiable use of PCA on bivariate data is to assess whether the bivariate technique is effectively removing the univariate size measure (but see **chp. 5**). It obviously is in this regression procedure as much of the univariate size is removed from the traditional PC1 size vector (fig. 26(i)). This removal can also be seen by how the univariate size measure (no. 51) is minimally represented on the eigenvectors in figure 27. However, some additional confounding multivariate size information is obviously not being removed because these scatter plots do not reveal complete group separation and the eigenvectors are still far from being isometric (fig. 27(i)). PCA was also performed on covariance matrices of the ratio and $\log_{10}$ ratio data from the previous bivariate analyses (see **chp. 5**). These PCA's revealed that size is not being effectively removed from the data at all by these 2 procedures (Atchley et al. 1976, Reist 1985, Shea 1985).

PCA on Regression Data — Summary

1. For morphometric analyses, use PCA on a covariance matrix usually of $\log_{10}$ transformed data, and not PCA on data previously adjusted by bivariate techniques.
2. PCA on bivariate adjusted data appears to have no benefits and many disadvantages.
    a) Synthesis of data into two or three significant vectors which correspond to size and shape does not occur. No multivariate size vector is obtained, and only a confounded multivariate shape vector is.

c) Complete interpretation of any resultant vectors is also confounded because of their inconsistent representation of characters and individuals.

d) Individual ordination and group/centroid separation still result, but require more than one vector and are not as effective as those based on standard PCA.

CHAPTER EIGHT

Principal Component Analysis of a Mixed Character Data Set

Introduction

Mixed character analysis in multivariate morphometric procedures involves the use of both morphological (continuous) and meristic (discontinuous) characters in the same analysis. The morphological and meristic variables are analyzed together, and their combined effects on the outcome may be quite different than if they are investigated separately (Thorpe 1976, 1980). Generally, separate multivariate analyses of morphological and meristic characters are recommended (Bookstein et al. 1985, Humphries et al. 1981, Pimentel 1979, Thorpe 1983a). However, this recommendation is often not followed. For instance, when only a few characters have been measured their low number may suggest that they be collectively analyzed.

In spite of the recommendation against mixed character analysis in multivariate morphometric procedures, their effects are really only theoretical. An empirical assessment of such a multivariate mixed character analysis is necessary to quantify any effects and determine how significant they are. Realistic recommendations regarding this practise can then be made.

In order to assess these possible effects, principal component analysis (PCA) based on a correlation matrix of $\log_{10}$ transformed data was carried out. This analysis is based on a combination of the fifty-one morphological and ten meristic characters used in **part II** of this thesis. No bivariate morphometric assessment mixed character effects is made because any effects would only occur in such cumulative bivariate measures as mean individuals (see **chp. 5**). Since each character is analyzed separately in bivariate morphometric procedures there will be no mixed character effects on the individually adjusted characters themselves. Therefore, the only bivariate assessment undertaken here is of mean individuals based on mixed characters adjusted by regression analysis.

Mixed Character Analysis — Assessment Methods

The effects of applying multivariate morphometric procedures to mixed character data are analyzed with PCA based on a correlation matrix of $\log_{10}$ transformed mixed data (technique C in **chp. 6**). The correlation matrix is used because it standardizes the data better than a covariance

matrix (see **chp. 6**), and it is recommended for mixed character analyses (Pimentel 1979, Thorpe 1976, 1980). This standardization of the data by the correlation matrix is to zero mean and unit standard deviation for each character and thus better accomodates the analysis of different mixed characters. The correlation matrix PCA should provide the best opportunity for a multivariate analysis of mixed data to produce good results. A full discussion of the multivariate morphometric procedures and their terminology is in **chapter 6.**

The effect of mixed character sets on bivariate morphometric procedures are assessed using regression analysis (technique **e** in **chp. 5**). This regression procedure was determined to be the best bivariate morphometric approach (**chp. 5**), and thus it should provide the best chance for a bivariate analysis of mixed characters to succeed. Since only cumulative bivariate measures can be affected by mixed character analysia, only mean individuals are analyzed. The mean individuals are calculated as the mean of all sixty-one regression-adjusted characters for each individual fish. The mean individuals represent an effective and appropriate shape measure for bivariate analyses, and are further justified in **chapter 5**. Further discussion of the bivariate morphometric techniques is in **chapter 5.**

The morphological characters are numbered 1–51, and the meristic characters are still numbered 1–10 for consistency and comparison within the thesis. The morphological variables are plotted first, followed by the meristic characters. The sixty individual fish form two groups which are each respectively represented by the numbers 1–30 set in small type and the numbers 31–60 set in large type. Their centroids (group means) are also plotted in small and large sizes. Figure 28 presents the results for individuals and figure 29 for characters.

All sixty-one variables are used in spite of this violating a multivariate rule-of-thumb that the number of characters should not exceed the sample size (see **chp. 6**). To ensure that this character number is not a problem, a jackknife procedure (see **chp. 4**) was used. Sets of the variables were removed and the subsets were reanalyzed in the PCA to see if this changed any of the results. No differences in the results were noted because of this character reduction so the full set is used to maintain consistency within the thesis. The sample size was also verified as being large enough for this full character set (see **chp. 4**).

Figure 28(i) is a bivariate scatter plot of the mean individuals for all sixty-one measurements plotted against the $\log_{10}$ transformed univariate size measure (standard length). This figure demonstrates where and how effectively the individuals are ordinated and their groups and centroids are separated. It also reveals whether the use of a mixed character set has any effect on the removal of size from the mean individual shapes based on these characters. Standard length is employed as the univariate size measure because it is used in the bivariate regression technique to derive these bivariate shape measures. Standard length is $\log_{10}$ transformed to make more effective use of the graph space. Untransformed size produces a similar plot but with the individuals clumped and less evenly distributed.

Figure 28(ii) is a PCA scatter plot of the second principal component (PC2) against the first (PC1). Size information is usually summarized in PC1 and shape in PC2, and thus these plots reveal whether effective size/shape separation is occuring with the analysis of a mixed character data set. The plot also demonstrates if the individuals are correctly ordinated, and if their groups and centroids are effectively separated. The percentage variance accounted for by each pc will also give some indication of the effectiveness of both the mixed character analysis and the segregation of the information into size and shape vectors. The arrangement of the data ellipses for each group permits further assessment of the type of shape and size information that results from this mixed data analysis. How mixed character patterns differ from those patterns based only on the morphological variables (fig. 23; **chp. 6**) indicates some of the effects which mixed characters have on PCA.

Figures 28(iii)-(iv) show how much and where the variance of the original individuals is accounted for in each of the first two PC's. These plots demonstrate whether all the individuals are being sufficiently represented in the PC's and also whether a PC has significance and should be analyzed. If the representation of individuals is very inconsistent, the individuals are not being accounted for in the same fashion. An analysis of individuals based on the first two PC's is then not representative of the individuals or groups they belong to. If an insufficient amount of variance is accounted for by the initial PC's, more PC's probably need to be interpreted in order to understand the multivariate morphometric results.

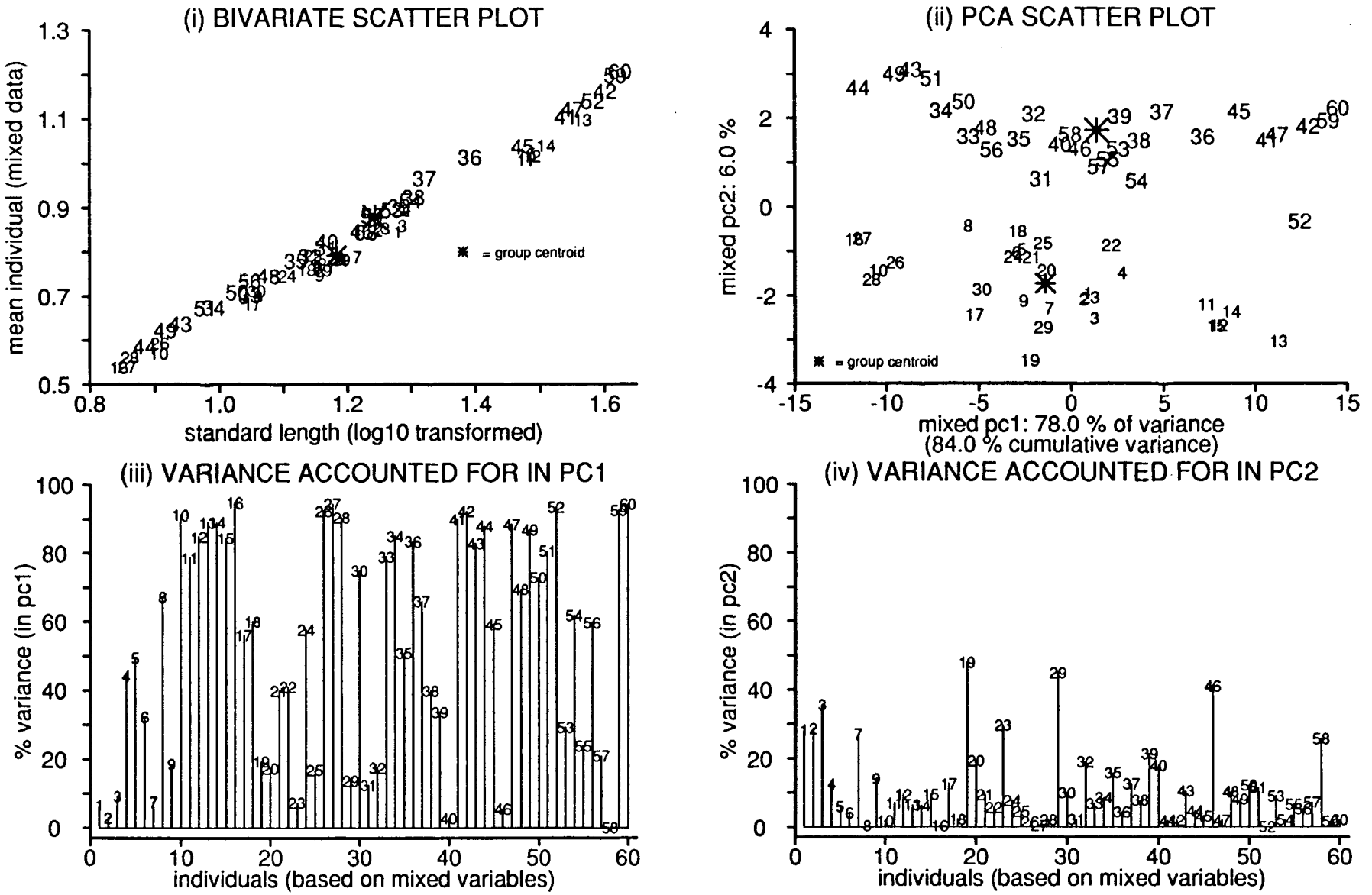


FIGURE 28. Patterns for individuals based on mixed variables (morphology and meristics).

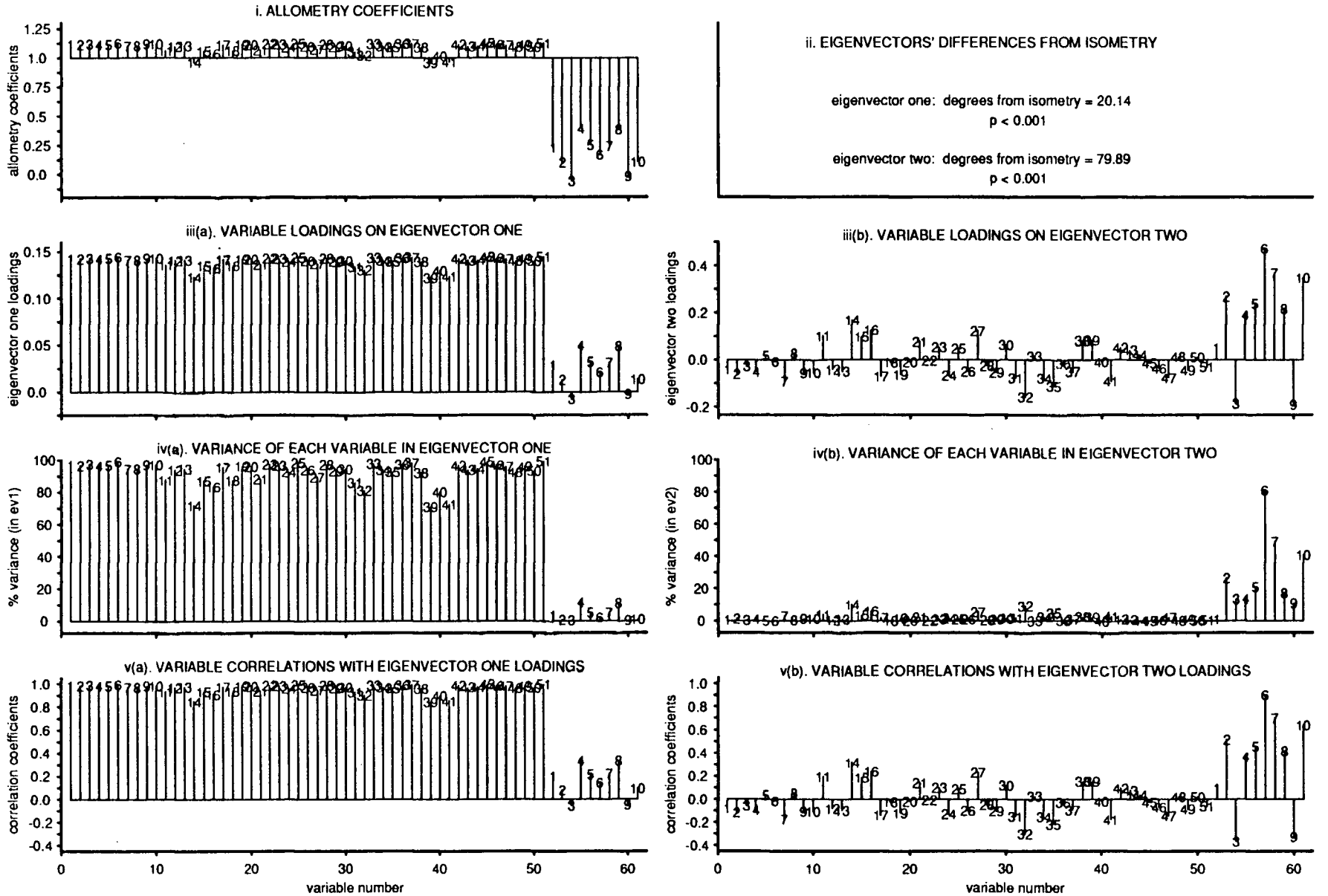


FIGURE 29. PCA patterns for mixed variables (morphology and meristics).

Figure 29(i) gives the allometry coefficients calculated from the PCA. These reveal how well the mixed variable PCA retains the important allometric information, both for individual characters and overall. The allometry coefficients are compared amongst themselves and are also contrasted with the allometry coefficients calculated previously for PCA's done separately on the morphological and meristic character sets (fig. 21C; **chp. 6**). The multivariate allometry coefficients and their calculations are explained in **chapter 6**.

Figure 29(ii) presents the isometry statistics for the first two eigenvectors (Pimetnel 1979; also see Leamy and Bradley 1982, Somers 1986, Thorington 1972). The "degrees" show how many degrees each of the eigenvectors are from a theoretical vector of isometric size. They also demonstrate how much size is actually removed because the shape eigenvectors should be nearly orthogonal ($\approx$90 degrees) to the size vector even if complete isometry does not exist in the data (eg. table 2; **chp. 6**). The p-values presented are the chi-square values derived from Anderson's (1963) test. A statistically significant value ($p < 0.05$) indicates that isometry is not achieved for that eigenvector and that some allometric size information is present in it.

Figures 29(iii)–(v) respectively present the first two eigenvector loadings, the percentage variance of each original character accounted for in the first two eigenvectors, and the correlations of each original character with the first two eigenvector loadings. These punk plots are all extrapolated from or centred about zero and share equivalent axes within each dual set of figures. This permits the assessment of the overall and individual character patterns. These plots indicate whether the characters are sufficiently represented in specific eigenvectors, and also whether an eigenvector has significance and should be analyzed. If the representation of characters is very inconsistent, they are not accounted for in the same vectors. If the loadings, variances or correlations are very low then they are being not being fully represented in those eigenvectors and more eigenvectors may need interpretation to explain all of the multivariate analysis. The traditional designation of a multivariate size vector as a large general eigenvector, and of a multivariate shape vector as a smaller bipolar eigenvector, is also examined through this set of figures.

The homoscedasticity of the groups based on the total mixed character set is assessed using Box's test. The multivariate normality of this data is tested using the qualitative probability plot. Both of these tests are fully discussed in **chapter 4**. Univariate normalities are not tested since

they will remain the same as that presented in figure 25 (**chp. 4**). The number of significant eigenvectors resulting from the PCA are tested using a Scree test and Bartlett's test of sphericity (see **chp. 6**).

The analyses and results are based on computer macros I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Centre, University of British Columbia. These programs are available from me.

Mixed Character Analysis — Assessment Results

The use of a mixed character set in PCA and in a cumulative bivariate shape estimate (mean individuals) has some definite effects on both individuals and characters. While many of these are not too insidious, they nonetheless must be recognized if such mixed data analyses are to be used. Even minimal effects can result in some differences which may influence the final interpretation of a data set.

The effect on the bivariate mean individual shape estimate (fig. 28(i)) is severe. The mean individuals no longer seem to correspond to shape at all, even though the mean individuals based only on morphology do at least approximate shape (fig. 16; **chp. 5**). These mixed data mean individuals are tightly correlated to the univariate size measure and do not ordinate individuals or separate their groups or centroids effectively. In bivariate analyses of mixed characters, there is no effect on the characters themselves but there is a definite problem if these variables are to be cumulatively summarized in some fashion.

On the other hand, the PCA scatter plot in figure 28(ii) demonstrates completely effective group and centroid separation. The PC1 axis definitely corresponds to size information and the PC2 axis to shape. The PC2 axis represents all the important information in regards to the groups, but the shape of each group's plotted ellipse is different than those based on morphology alone (fig. 23; **chp. 6**).

While the individuals are still correctly ordinated to their groups in the mixed character analysis, the actual placement of some individuals and their cumulative range and patterns are different from the PCA scatter plots based only on morphological variables (fig. 23; **chp. 6**). The first mixed character group is now more tightly clumped along the shape axis whereas the second

mixed character group is more spread out. The second mixed character group has shifted its shape designation of smaller fish to a different set of PC2 scores. The mixed character PC1 axis also accounts for a much lower overall percentage variance, and the PC2 axis for a much higher percentage.

These differences in how individuals are portrayed in terms of shape and size is further demonstrated in figures 28(iii)–(iv). Here there is a much less consistent representation of the individuals in PC1 than in the PCA based only on morphology (fig. 25C; **chp. 6**). Also, those mixed character individuals which score strongly on PC1 and weakly on PC2 are the larger fish whereas they are the intermediate size fish in the morphology PCA. It makes intuitive sense that the intermediate size fish should score more strongly since they are probably closer to the group means and PCA is essentially producing results that are mean related. As well, the cumulative variance of each individual accounted for in the first two PC's is slightly lower in comparison to the PCA on morphology alone.

The allometry coefficients in figure 29(i) demonstrate that the allometric relationships in the original data (fig. 15a/b; **chp. 5**) and those represented by PCA on morphology or meristics alone (fig. 21C; **chp. 6**) are lost in the mixed character analysis. Nearly all the morphological variables in the mixed character PCA load positively and at approximately the same level. The meristic characters in this data set have no size information (see **part II**) so their allometry coefficients are not realistic and can only be used for procedure assessment and comparison. The meristic allometry coefficients for mixed character data are very different from those based on the meristic data alone. The allometric size information presented by the mixed character analysis is therefore quite different from that based on an analysis of only the morphological or meristic data.

However, the mixed character loadings (fig. 29(iii)), their percentage variances accounted for (fig. 29(iv)), and the correlations of the original variables with these loadings (fig. 29(v)) strongly resemble those of the analyses based only on the morphological or meristic data (figs. 18–20; **chp. 6**). The morphological variables are almost identical in both cases and the meristic characters from the mixed character PCA are only slightly different (characters 5–7). In the mixed data set, the patterns for morphological variables are slightly higher in the first eigenvector and lower in the second, and the opposite is true of the meristic characters. The morphological characters

dominate the first eigenvector and the meristic ones the second eigenvector. This lower second morphological eigenvector representation for the mixed data set is particularly noticeable in the percentage variance results.

The morphological variables have a high general character loading pattern in the first eigenvector so this eigenvector still corresponds to size at least for the morphological variables. This is offset, however, by the meristic characters which are not strongly loaded and are even slightly bipolar in the first eigenvector. The second eigenvector is smaller and bipolar for the morphological variables so it probably corresponds to a shape vector, again at least for them. The meristic characters have a bipolar pattern in the second eigenvector as well, but here they definitely dominate and load heavily in comparison to the morphological variables. These problems with size and shape designation are supported by the isometry statistics (fig. 29(ii)). The first mixed character eigenvector is now further from isometry than it is in the PCA based only on morphological variables (table 2; **chp. 6**). The second mixed character eigenvector is also not nearly as orthogonal to an isometric eigenvector and its shape is thus not as well separated from size.

The multivariate normality of this mixed character set is not as good as either the separate morphological or meristic character sets. Other transformations of the mixed data, aside from logarithmic, did not help in this regard. The mixed character analysis also resulted in three vectors being significant instead of the two vectors which previously were significant for the morphological and meristic data alone (**chp. 6**).

Mixed Character Analysis — Assessment Discussion

The overall effects of a mixed character PCA are subtly different from those of a PCA based only on a morphological or meristic character set. Individual ordination and group/centroid separation (fig. 28(ii)) are improved in the mixed character analysis, but the representation of individuals and characters is not as effective or realistic. The extent to which this should be of concern depends on the research objectives, but if possible analyze morphological and meristic characters separately. If mixed character PCA is undertaken be aware of its effects on the results.

The improvement in classification of the mixed character PCA is also not better than the scatter plots of meristic PC2 against morphological PC1 (fig. 24; **chp. 6**). In these other PCA's, the

meristic and morphological data are analyzed separately. The morphological data interpretation is thus realistic and effective, and excellent individual ordination and group separation is also possible. The classification advantages of this other approach to mixed character analysis are obtained without the disadvantages of poor character and individual representation resulting from a single mixed character PCA.

While the first PC and eigenvector still seem to correspond to size in the mixed character analysis, these relationships are no longer as strong or clean (fig. 29). This is probably the result of the addition of the meristic characters to the analysis. Since these meristic characters contain no size information, the only information they add to the first vectors must relate to a meristic equivalent of shape and only confounds the morphological size vector (Reist 1985, Somers 1986). The size vectors also no longer represent the original allometric relationships of the data, are not even close to being isometric, account for less overall variance and do so inconsistently for the variances of each individual. Most importantly, this also affects the shape vectors.

The second PC and eigenvector still correspond to shape, but this shape is now also somewhat different. Some of the variation that is size-related in the morphological data set and much of the meristic character information now ends up in these second vectors. The result is that they account for more variation, but that they also do not portray the traditional shape relationships in the data as effectively.

This effect on shape is worse for the cumulative bivariate assessment of shape using mean individuals (fig. 28(i)). The mean individuals no longer represent shape at all and are tightly correlated to the individual size measure. This is opposite to the previous mean bivariate individual plots based only on either the morphological or meristic data sets.

The mixed character analysis also results in less effective multivariate data synthesis as three eigenvectors are now significant. This suggests that the first three vectors, and not just the first two, should be analyzed if all the significant information resulting from this multivariate analysis is to be used. The results for the third vector are not graphically portrayed here so that the mixed character PCA results could be consistently compared with the traditional morphometric interpretation of the first two PCA vectors. When analyzed, the third vectors seem to correspond

to further shape information. Nonetheless, this shape information in the third vector from the mixed character PCA only contains variation that is usually summarized in the first two vectors of a PCA based only on morphological data.

These problems may also partly result from the rather poor multivariate normal distribution this mixed data set has. The multivariate normality problem could not be corrected with any data transformations tried. Such multivariate non-normality may often be a problem in the multivariate analysis of mixed character data sets. This multivariate non-normality probably results from the poor correlations between the morphological and meristic characters.

This poor correlation of morphological and meristic characters also suggests that this data set may be underestimating the potential effects of mixed variable analyses. The large number of morphological variables analyzed in this study may be minimizing the potential influence of the meristic characters on the results of the PCA on a mixed data set. If the meristic characters composed a larger proportion of the data set they might exert stronger effects and distort the analysis further.

Mixed Character Analysis — Summary

1. PCA based on mixed character data sets should only be done if necessary, and with the negative effects of this procedure kept in mind.
    a) Group/centroid separation is improved, but this improvement can be obtained without the problems generated by mixed character analysis.
    b) The representation of individuals, and of characters and their allometric relationships, is confounded and unrealistic.
2. The traditional multivariate morphometric interpretation of the first vectors as size and the second vectors as shape is still possible in a mixed character PCA, but it also somewhat confounded and confused.
    a) This is especially true of the shape relationships of individuals reflected in the second PC.

b) There is also a reduction in data synthesis since an additional eigenvector in this study now contains statistically significant shape information.

3. The mixed character effects on a cumulative bivariate shape measure such as mean individuals are more severe than in the multivariate case.

   a) This bivariate mean individual shape measure is now completely size-confounded.

   b) Mixed character data sets can be analyzed with bivariate morphometric procedures without affecting their results for individuals and characters, but any cumulative representation of these results will probably be strongly confounded.

CHAPTER NINE

Back-Transformation of Principal Component Analysis

Introduction

One of the most confusing aspects of principal component analysis (PCA) is the numerical output. The numbers in the eigenvectors and principal components (PC's) bear no resemblance to the original data or their scale, and consequently are often difficult to understand. While their interpretation is obviously possible and informative (see **chp. 6**), their relationships to the biological phenomena under study can only be assessed through inference.

This inferential relationship has added to the difficulty in understanding multivariate morphometric procedures, and has thus also often resulted in an avoidance of the use of these techniques. After all, multivariate procedures are already mathematically complex and these methods will be ignored or used incorrectly if their output is not simple, direct and understandable. This anathema is unfortunate since multivariate morphometric procedures provide useful syntheses of the significant variation in morphological data into components of size and shape. The techniques thus realistically arrange each character into uncorrelated information pertaining to size and shape, and state how much variance each is accounting for. The multivariate morphometric ordination of individuals, and separation of groups and their centroids, is also excellent.

Multivariate procedures should, therefore, be used for morphometrics. What is needed then is a simple technique for back-transforming the eigenvector loadings and PC scores into numbers that intuitively resemble the original data. The back-transformed numbers should form two matrices, one for size information and one for shape. These matrices should be in the original data matrix format and dimensions, and resemble the original data and their scale.

No back-transformation method has ever been used in morphometrics. Somers (1986) discusses such a technique and has a possible program for it in his size-constrained PCA procedure based on a correlation matrix. He does not, however, actually use the technique or realize any predecessors. Further literature review revealed only brief discussions of back-transformation and

its possible practicality in Chan and Dunn (1975), Lachenbruch (1975), Phillips et al. (1973) and Veitch (1965).

This study (**chps. 5–6**) has revealed that PCA based on a covariance matrix of $\log_{10}$ transformed data is the best multivariate morphometric procedure. A possible back-transformation method for this procedure is therefore developed and presented here. This back-transformation method has benefited greatly from the back-transformation procedure for a PCA on a correlation matrix presented in Somers (1986) and through personal communication with him. His equivalent back-transformation formula for PCA on a correlation matrix is also given here because its presentation in Somers (1986) is verbal and not altogether clear. His formula also acted as a reference against which this PCA on a covariance matrix back-transformation formula could be assessed.

Formulas for PCA Back-Transformation

Covariance Matrix

For a back-transformed size matrix from PCA on a total covariance matrix:

$$halfback1 = pc1 \cdot ev1',$$

$$fullback_{size} = halfback1_p + \bar{x}_p$$

For a back-transformed shape matrix from PCA on a total covariance matrix:

$$halfback2 = pc2 \cdot ev2',$$

$$fullback_{shape} = halfback2_p + \bar{x}_p$$

Correlation Matrix

For a back-transformed size matrix from PCA on a total correlation matrix (Somers 1986, pers. comm.):

$$halfback1 = pc1 \cdot ev1',$$

$$fullback_{size} = halfback1_p \cdot S_p + \bar{x}_p$$

For a back-transformed shape matrix from PCA on a total correlation matrix (Somers 1986, pers. comm.):

$$halfback2 = pc2 \cdot ev2',$$

$$fullback_{shape} = halfback2_p \cdot S_p + \bar{x}_p$$

where:

$pc1$ = principal component one (size vector for individuals);

$pc2$ = principal component two (shape vector for individuals);

$ev1'$ = transposed first eigenvector (size vector for characters);

$ev2'$ = transposed second eigenvector (shape vector for characters);

$halfback_p$ = each p^{th} character column of the halfway back transformed matrix;

$fullback$ = fully back transformed matrix of size or shape data;

$\bar{x}_p$ = vector of untransformed original p^{th} character means;

S_p = vector of untransformed original p^{th} character standard deviations.

Formula Discussion

These formulas are correct for matrices based on untransformed or transformed (any transformation) data. This is because the back-transformation uses the eigenvectors and PC's derived from the PCA and these numbers will be scaled regardless of the original data transformation. The original untransformed character means (and their standard deviations for the correlation matrix method) will be needed, however, for the last step of the back-transformation.

Since a correlation matrix standardizes the data to character means of zero and unit (one) standard deviations, the back-transformation procedure for a PCA based on a correlation matrix requires that both these dimensions be put back into the data. A covariance matrix does not standardize the variances of the data so only the character means must be re-established. The covariance matrix must be mean-centred in the PCA because its back-transformation is based on mean-centred matrices. A standard set of means could also be added back onto either data matrix in order to compare several populations on some predetermined size scale.

Since these data are homoscedastic for both groups (see **chp. 4**), the back-transformed procedure works effectively with the total matrix sample means (and standard deviations). If groups are pooled for deliberate reasons or to deal with heteroscedasticity, their means (and standard deviations) returned to the PCA data values should be derived from each group independently (Somers pers. comm.). If group means are added back on, the groups' size and shape data may become more disparate than if total means are used. This is because any differences in the group means will be put returned to the data, whereas this does not happen with total means.

A final suggestion is that the back-transformation will be best if the first two PCA vectors are statistically significant (see **chp. 6**). If both sets of vectors are significant, then all the important size and shape information will be used in the back-transformation. When the initial three PCA vectors are significant, the back-transformation procedure will probably still be effective because the information that is being used still represents size and shape. The significant third vectors almost certainly correspond to shape information (Pimentel 1979) and may require interpretation then as well. If more than three vectors are significant, however, this back-transformation procedure will probably not work well unless all the significant vectors are back-transformed. Unfortunately, back-transformation of all these vectors will likely be difficult to interpret and may even be meaningless since their information is spread across many vectors (eg. **chp. 7**). Only the first two PCA vectors are significant in this data (see **chp. 6**).

PCA Back-Transformation — Assessment Methods

Fifty-one morphological and ten meristic characters are independently analyzed here (see **chp. 4**). The sixty individual fish in these plots compose two groups. Group one is represented by the numbers 1–30 set in small type and group two by the numbers 31–60 set in large type. heir centroids (group means) are also plotted in small and large sizes. The morphological and meristic back-transformed size and shape variables are tested for homoscedasticity using Box's test (see **chp. 4**).

Figure 30 is a scatter plot of mean individuals against the $\log_{10}$ transformed univariate size measure of standard length. The mean individuals were calculated by taking the mean of all the measurements for each individual and are done separately for the morphological and meristic data

sets. Their use is novel yet can be justified by their effectiveness and consistency on figure 30, and because they have very uniform, if high, standard deviations (for further justification see **chp. 5**).

Figure 30 permits an assessment of how well the back-transformed morphological and meristic characters correspond to size and shape. It also shows how well and on which axes the individuals are ordinated, and the groups and their centroids are separated. PC1 and eigenvector one should correspond to size, and PC2 and eigenvector two to shape. These size and shape features of PCA have long been accepted but are supported only by much inferential evidence (see chp. 6). This back-transformation analysis provides an additional test of their validity by seeing how the back-transformed size and shape matrices relate to a univariate size measure. If PC1 and eigenvector one are size-related they should correlate strongly to the univariate size measure, and vice-versa for shape-related PC2 and eigenvector two.

The analyses and graphics are based on computer programs I wrote within the S facility (Becker and Chambers 1984) used in the UNIX operating system (McGilton and Morgan 1983) at the Biological Data Centre, University of British Columbia. These programs are available from me.

PCA Back-Transformation — Assessment Results

Figure 30 presents all the individual ordination and group/centroid separation results of the back-transformation of a PCA based on a covariance matrix of my same $\log_{10}$ transformed morphological data set (see **chp. 4**). These back-transformed character values closely resemble those of the original data and their scale. The data back-transformed from the covariance matrix do not, however, have any variance because only the means are returned to the PCA values.

Figure 30(i) demonstrates that the back-transformed morphological size data is tightly correlated to the univariate size measure of standard length. The relationship between multivariate size and this univariate size measure is strong. The back-transformed morphological shape data in figure 30(ii) reveal the same individual ordination and group/centroid separation patterns as the original PCA scatter plot (fig. 23; **chp. 6**). No confounding size information is present in the shape data, and this shape information is not related to the univariate size measure.

The back-transformed meristic characters confirm that the meristic data does not contain size information. Figure 30(iii) shows excellent individual ordination and effective group/centroid

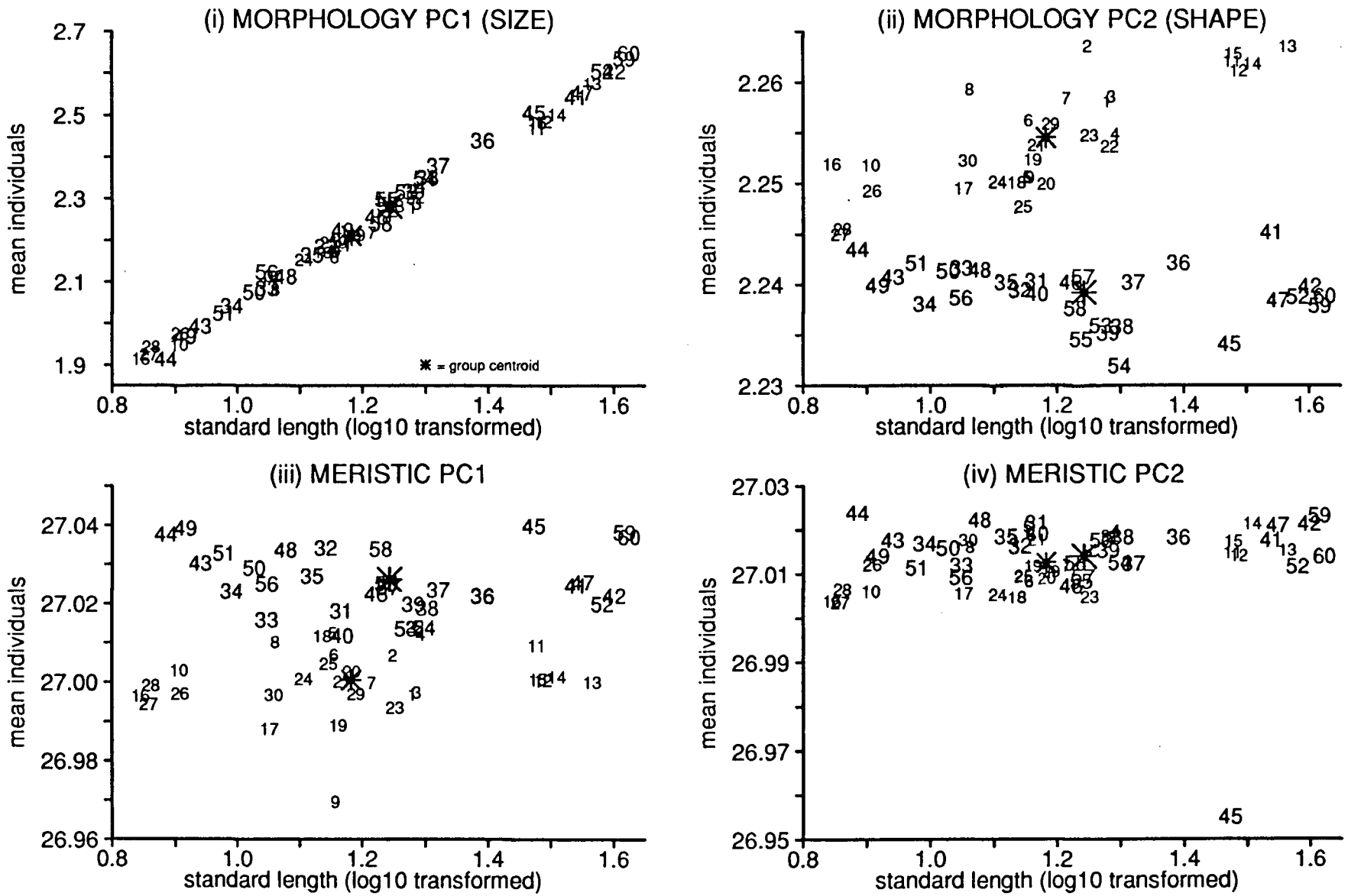


FIGURE 30. Bivariate scatter plots of back-transformed individuals (from PCA).

separation in PC1. This pattern would normally represent size in a morphological data set. It does not appear to here. The meristic PC2 plot in figure 30(iv) does not ordinate individuals or separate their groupings. It does not seem to correspond to any size or shape information. This is true even if the outlying individual (no. 45) is removed from the scatter plot. This outlying individual is also quite an unusual PCA result, especially since it has not appeared as an outlier in any previous analyses (**chps. 5–8**).

The groups are still homoscedastic (Box's test : $p > 0.5$) in the back-transformed size and shape matrices for both the morphological and meristic data sets. Both morphological matrices are also not singular despite their large number of characters (see **chp. 4**). As well, the variables still have the same original data distributions as represented in figure 14 (**chp. 4**), except for standard deviations in the covariance matrix case.

Scatter plots of mean individuals based on shape against the mean individuals based on size produce identical ordination results to the mean shape individual plotted against univariate size. The back-transformed data from PCA on a correlation matrix are also similar to the covariance matrix results. Some of the actual character values resulting from the correlation back-transformation differ slightly, but this has no effect on any general patterns. There are some slight character changes in the back-transformed data from the correlation matrix, however, and these further indicate the subtle differences which can result from using PCA on either the covariance or correlation matrix.

PCA Back-Transformation — Assessment Discussion

The back-transformation of the PCA output is simple, and it appears to be realistic and effective. The back-transformed numbers have none of the conceptual difficulties of the numerical PCA output and all of their advantages. The characters regain their original distribution and scale, and the morphological variables now form distinct categories of unconfounded size and shape information. The individuals based on morphology are correctly ordinated, and their groups and centroids are effectively separated (figs. 30(i)–(ii)). The back-transformed meristic characters reveal that they contain no size information and that they should not be corrected for allometry.

Figures 30(iii)-(iv) show that the mean meristic individual ordination is only on the first PC and that this PC does not seem to correspond to size. The second meristic PC does not seem to correspond to any biological feature. PCA back-transformation on these meristics, or on other data without size information, are reasonably acceptable as a technique for data synthesis but not for allometric adjustment. In this case, a PC1 is meristic equivalent of "shape" and not size, and further PC's may be biologically meaningless.

The creation of a representative mean shape (or size) individual for each group in an analysis is another possible use of back-transformation. The analysis of multiple groups in single PCA does not provide direct character information for the individuals in those groups. The eigenvectors give character loadings but these correspond to how significant these variables are and what relationships they have to other characters. Back-transformation still permits this recognition of significant characters and their allometric relationships, but it also does so for the individuals and groups not just the characters themselves.

Such representative mean individuals for a group have previously only been accessible through bivariate morphometric procedures. Often, however, these bivariate procedures are not as effective in allometric adjustment as multivariate methods are. In addition, these bivariate groups are usually based on *a priori* assignment and this prior recognition is not necessary with PCA (see **chp. 4**).

This back-transformation data also directly verifies the long-held assumption that the first morphological PC and eigenvector correspond to size, and that the second morphological PC and eigenvector contain shape information. The size information that is leftover in the second PC and eigenvector is also not confounding the shape relationships present. The size information in the first PC and eigenvector also represent the actual allometric relationships and overall sizes most effectively. Therefore, the argument that standard PCA on morphological variables does not effectively remove all the size information in the first component (Archie 1987, Bookstein et al. 1985, Pimentel 1979, Reyment et al. 1984, Rohlf and Bookstein 1987, Somers 1986) is misleading. This argument may well be true in terms of isometric size, but the allometric size information that remains in the second eigenvector and PC contributes to that of shape and the relationships it portrays.

The relationship between the multivariate morphological size factor and the univariate size measure is strong. However, a comparison of these multivariate back-transformed scatter plots (fig. 30) with similar plots based on bivariate morphometric procedures (fig. 16; **chp. 5**) further demonstrates that univariate size measure compensation is insufficient in this study. The multivariate size factor is effectively removing the univariate size measure, but it is also accounting for some other size information that is not present in the univariate case (Baumgartner et al. 1988, Rohlf and Bookstein 1987).

PCA Back-Transformation — Summary

1. PCA back-transformation procedure is simple and appears to be effective.
    a) Only the significant variation is accounted for.
    b) Character values are realistic and are in the same dimensions, scale and distribution as the original values.
    c) Individuals are realistically ordinated.
    d) Groups and centroids are effectively portrayed and separated.
2. The back-transformed morphological data produce distinct and uncorrelated size and shape matrices. The variance each matrix accounts for is also known from their PCA eigenvalues.
3. The first morphological principal component and eigenvector are definitely size.
    a) In this study, the multivariate size factor is better for the adjustment of confounding allometric size than the univariate size measure.
4. The second morphological principal component and eigenvector are definitely shape.
    a) While some size information remains here it is not confounding the shape parameters and should not be removed through further manipulations. This leftover size is contributing to the shape information.
5. Representative mean individuals of all the measurements for any groups can be calculated from the back-transformed data matrices and used for inter-group comparisons.

References

Abe, S. and J. Muramoto. 1974. Differential staining of chromosomes of two salmonid species, *Salvelinus leucomaenis* and *Salvelinus malma*. Proc. Jap. Acad. 50:507–511.

Aho, A.V., B.W. Kernighan and P.J. Weinberger. 1988. **The AWK programming language.** Addison-Wesley Publ. Co., Reading, MA.

Alberch, P. 11980. Ontogenesis and morphological diversification. Amer. Zool. 20:653–667.

Alberch, P. 1982. Developmental constraints in evolutionary processes, pp.313–332. **In:** J.T. Bonner (ed.), Evolution and development. Springer—Verlag, Berlin.

Alberch, P. 1985. Problems with the interpretation of developmental sequences. Syst. Zool. 34:46–58.

Alberch, P. and J. Alberch. 1981. Heterochronic mechanisms of morphological diversification and evolutionary change in the neotropical salamander, *Bolitoglossa occidentalis* (Amphibia; Plethodontidae). J. Morph. 167:249–264.

Alberch, P. and E.A. Gale. 1983. Size dependence during development of the amphibian foot. Colchicine—induced digital loss and reduction. J. Embryol. Exp. Morph. 76:177–197.

Alberch, P. and E.A. Gale. 1985. A developmental analysis of an evolutionary trend: Digital reduction in amphibians. Evolution 39:8–23.

Alberch, P., S.J. Gould, G.F. Oster and D.B. Wake. 1979. Size and shape in ontogeny and phylogeny. Paleobiology 5:296–317.

Albrecht, G.H. 1978. Some comments on the use of ratios. Syst. Zool. 27:67–71.

Allen, J.E., M. Burns and S.C. Sargent. 1986. **Cataclysms on the Columbia: A layman's guide to the features produced by the catastrophic Bretz floods in the Pacific Northwest.** Timber Press, Portland, OR.

Allendorf, F.W. and F.M. Utter. 1979. Population genetics, pp. 407–454. **In:** W.S. Hoar, D.S. Randall and J.R. Brett (eds.), Fish physiology, vol.8. Academic Press, New York, NY.

Alley, N.F. and S.C. Chatwin. 1979. Lake Pleistocone history and geomorphology, southwestern Vancouver Island, British Columbia. Can. J. Earth Sci. 16:1645–1657.

Anderson, A.J.B. 1971. Numeric examination of multivariate soil samples. Math. Geol. 3:1–14.

Anderson, D.E. and R. Lydic. 1977a. Ratio data and the quantification of drug effects. Biobehav. Rev. 1:55–57.

Anderson, D.E. and R. Lydic. 1977b. On the effect of using ratios in the analysis of variance. Biobehav. Rev. 1:225–229.

Anderson, T.W. 1963. Asymptotic theory for principal component analysis. Ann. Math. Stat. 34:122–148.

Andersson, L., N. Ryman and G. Ståhl. 1983. Protein loci in the Arctic charr, *Salvelinus alpinus* L.: Electrophoretic expression and genetic variability patterns. J. Fish Biol. 23:75–94.

Andrews, D.F., R. Gnanadesikan and J.L. Warner. 1973. Methods for assessing multivariate normality, pp. 95–116. **In:** Proc. Int. Symp. Multivariate Analysis. vol. 3. Academic Press, New York, NY.

Andrusak, H. and T.G. Northcote. 1971. Segregation between adult cutthroat trout(*Salmo clarki*) and Dolly Varden (*Salvelinus malma*) in small coastal British Columbia lakes. J. Fish. Res. Bd. Can. 28:1259–1268.

Archie, J.W. 1987. Summary of the Twentieth International Numerical Taxonomy Conference. Syst. Zool. 36:216–223.

Arkhipchuk, V.V. and G.D. Berdyshev. 1987. Relationship between karyotypic and morphological variation in fishes. J. Ichthyol. 27:158–161.

Armstrong, J.E. 1981. Post-Vashon Wisconsin glaciation, Fraser Lowland, British Columbia. Bull. Geol. Surv. Can. 322:1–34.

Armstrong, R.H. and J.E. Morrow. 1980. The Dolly Varden charr, *Salvelinus malma*, pp. 99–140. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Armstrong, R.H. and W.M. Morton. 1969. Revised annotated bibliography on the Dolly Varden char. Alaska Dept. Fish Game Res. Rep. 7.

Atchley, W.R. 1978. Ratios, regression intercepts, and the scaling of data. Syst. Zool. 27:78–83.

Atchley, W.R. 1980. M-statistics and morphometric divergence. Science 208:1059–1060.

Atchley, W.R. 1984. Ontogeny, timing of development, and genetic variance-covariance structure. Amer. Nat. 123:519–540.

Atchley, W.R. and D. Anderson. 1978. Ratios and the statistical analysis of biological data. Syst. Zool. 27:71–78.

Atchley, W.R., C.T. Gaskins and D. Anderson. 1976. Statistical properties of ratios. I. Empirical results. Syst. Zool. 25:137–148.

Atchley, W.R., B. Riska, L.A.P. Kohn, A.A. Plummer and J.J. Rutledge. 1984. A quantitative genetic analysis of brain and body size associations, their origin and ontogeny: Data from mice. Evolution 38:1165–1179.

Atchley, W.R. and J.J. Rutledge. 1980. Genetic components of size and shape. I. Dynamics of components of phenotypic variability and covariability during ontogeny in the laboratory rat. Evolution 34:1161–1173.

Atchley, W.R., J.J. Rutledge and D.E. Cowley. 1982. A multivariate statistical analysis of direct and correlated response to selection in the rat. Evolution 36:677–698.

Baker, A.J. 1980. Morphometric differentiation in New Zealand populations of the house sparrow, (*Passer domesticus*). Evolution 34:638–653.

Baker, R.J., W.R. Atchley and V.R. McDaniel. 1972. Karyology and morphometrics of Peters' tent-making bat, *Uroderma bilobatum* Peters (Chiroptera, Phyllostamatidae). Syst. Zool. 21:414–429.

Baker, A.J., R.L. Peterson, J.L. Eger and T.H. Manning. 1978. Statistical analysis of geographic variation in the skull of the Arctic hare (*Lepus arcticus*). Can. J. Zool. 56:2067–2082.

Baker, V.L. 1973. Paleohydrology and sedimentology of Lake Missoula flooding in eastern Washington. Geol. Soc. Am. Spec. Pap. 144.

Baker, V.L. 1988. Book review: Cataclysms on the Columbia. Amer. Sci. 76:187–188.

Balinsky, B.I. 1981. **An introduction to embryology**. (5th edition). Saunders College Publ., Philadelphia, PA.

Ball, G.H. and D.J. Hall. 1970. Some implications of interactive graphic computer systems for data analysis and statistics. Technometrics 12:17–31.

Ball, I.R. 1975. Nature and formation of biogeographical hypotheses. Syst. Zool. 24: 407–430.

Balon, E.K. 1979. The juvenilization process in phylogeny and the altricial to precocial forms in the ontogeny of fishes. Env. Biol. Fishes 4:193–198.

Balon, E.K. 1980a. Early ontogeny of the lake charr, *Salvelinus (Cristivomer) namaycush*, pp.485–562. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Balon, E.K. 1980b. Early ontogeny of the North American landlocked Arctic charr – sunapee, *Salvelinus (Salvelinus) alpinus oquassa*, pp. 563–606. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Balon, E.K. 1980c. Early ontogeny of the European landlocked Arctic charr – altricial form, *Salvelinus (Salvelinus) alpinus alpinus*, pp. 607–630. **In**: E.K. Balon (ed.), Charrs; Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Balon, E.K. 1980d. Early ontogeny of the brook charr, *Salvelinus (Baione) fontinalis*, pp. 631–666. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Balon, E.K. 1980e. Comparative ontogeny of charrs, pp. 703–720. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Balon, E.K. (ed.). 1980f. **Charrs: Salmonid fishes of the genus** *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Balon, E.K. 1981. Saltatory processes and altricial to precocial forms in the ontogeny of fishes. Amer. Zool. 21:573–596.

Balon, E.K. 1983. Epigenetic mechanisms: reflections on evolutionary processes. Can. J. Fish. Aquat. Sci. 40:2045–2058.

Balon, E.K. 1984. Life histories of Arctic charrs: An epigenetic explanation of their invading ability and evolution, pp. 109–141. **In**: L. Johnson and B.L. Burns (eds.), Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Baltz, D.M. and P.B. Moyle. 1981. Morphometric analysis of tule perch (*Hysterocarpus traski*) populations in three isolated drainages. Copeia 1981:305–311.

Barbour, S.E. 1984. Food size and jaw shape in Arctic charr, *Salvelinus alpinus*, pp.571–574. **In**: L. Johnson and B.L. Burns (eds.), Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Barraclough, R.M. and R.E. Blackith. 1962. Morphometric relationships in the genus *Ditylenchus*. Nematologica 8:51–58.

Bartlett, M.S. 1949. Fitting a straight line when both variables are subject to error. Biometrics 5:207–212.

Baumgartner, J.V., M.A. Bell and P.H. Weinberg. 1988. Body form differences between the Enos Lake species pair of threespine sticklebacks (*Gasterosteus aculeatus* complex). Can. J. Zool. 66:467–474.

Becker, R.A. and J.M. Chambers. 1984. **S. An interactive environment for data analysis and graphics**. Wadsworth Statistics / Probability Series, Wadsworth, Inc., Belmont, CA.

Behnke, R.J. 1972. The systematics of salmonid fishes of recently glaciated lakes. J. Fish. Res. Bd. Can. 29:639–671.

Behnke, R.J. 1980. A systematic review of the genus *Salvelinus*, pp. 441–481. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Behnke, R.J. 1984. Organizing the diversity of the Arctic charr complex, pp. 3–21. **In**: L. Johnson and B.L. Burns (eds.), Biology of the Arctic charr: Proceedings of the international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Behnke, R.J. and J. Shimizu. 1962. Book review. Studies on the chars found in Japanese waters, by M. Oshima. Copeia 1962:674–675.

Bell, M.A. 1981. Lateral plate polymorphism and ontogeny of the complete platemorph of the three spine sticklebacks (*Gasterosteus aculeatus*). Evolution 35:67–74.

Berg, L.S. 1948. Ryby presnykh vod SSSR: Sopredelnykh stran. Edition 4. Zool. Inst. Akad. Navk. SSSR, No. 27. [In Russian].

Best, T.L. 1978. Variation in kangaroo rats (genus *Dipodomys*) of the it Heermanni group in Baja California, Mexico. J. Mammol. 59:160–175.

Bigelow, R.S. and C. Reimer. 1954. An application of the linear discriminant function to insect taxonomy. Can. Ent. 86:69–73.

Bissell, A.F. and R.A. Ferguson. 1975. The jackknife — Toy, tool or two-edged weapon? Statistician 24:79–100.

Bisson, P.A. and C.E. Bond. 1971. Origin and distribution of the fishes of Harney Basin, Oregon. Copeia 1971:268–281.

Blackett, R.F. 1968. Spawning behavior, fecundity, and early life history of anadromous Dolly Varden, *Salvelinus malma* (Walbaum) in southeastern Alaska. Alaska Dept. Fish Game Res Rept. no.6.

Blackith, R.E. and R.A. Reyment. 1971. **Multivariate morphometrics.** Academic Press, London, UK.

Blackstone, N.W. 1986. Relative growth and specific growth rates in crustaceans. Growth 50:118–127.

Blackstone, N.W. 1987a. Specific growth rates of parts in a hermit crab: A reductionist approach to the study of allometry. J.Zool. 211A:531–545.

Blackstone, N.W. 1987b. Allometry and relative growth: Pattern and process in evolutionary studies. Syst. Zool. 36:76–78.

Blackstone, N.W. 1987c. Size and time. Syst. Zool. 36:211–216.

Blackwelder, R.E. 1964. Phyletic and phenetic versus omnispective classification, pp. 17–28. **In:** V.H. Heywood and J. McNeill (eds.), Systematics Assoc. Publ. no. 6., London, UK.

Bloom, W.L. 1976. Multivariate analysis of introgressive replacement of *Clarkia nitens* by *Clarkia speciosa polyantha* (Onagraceae). Evolution 30:412–424.

Bond, C.E. 1973. Keys to Oregon freshwater fishes. Oregon St. Univ., Agric. Exp. Stat., Tech. Bull. 58 (revised).

Bonner, J.T. 1965. **Size and cycle, an essay on the structure of biology.** Princeton Univ. Press, Princeton, NJ.

Bonner, J.T. (ed.). 1982. **Evolution and development.** Springer—Verlag, Berlin.

Bookstein, F.L., B. Chernoff, R. Elder, J. Humphries, G. Smith and R.Strauss. 1985. **Morphometrics in evolutionary biology.** Acad. Natural Sci. of Philadelphia, Special Publication 15, Philadelphia, PA.

Bookstein, F.L., R.E. Strauss, J.M. Humphries, B. Chernoff, R.L. Elder and G.R. Smith. 1982. A comment upon the uses of Fourier methods in systematics. Syst. Zool. 31:85–92.

Boratynski, K. and R.G. Davies. 1971. The taxonomic value of male Coccoidae (Homoptera) with an evaluation of some techniques. Biol. J. Linn. Soc. 3:57–102.

Borton, S.A., C.R. Gilbert, R.E. Jenkins and J.L. Oglesby. 1982. Ichthyofaunal cluster analysis of the western North Atlantic River drainages. Bull. Assoc. Southeast. Biol. 29:53.

Bostock, H.S. 1969. Kluane Lake, Yukon Territory, its drainage and applied problems (115G, and 115FE). Geol. Surv. Can. Pap. 69–28.

Box, G.E.P. 1954. Some theorems on quadratic forms applied in the study of analysis of variance problems. I. Effect of inequality of variance in the one-way classification. Ann. Math. Stat. 25:290–303.

Boyce, A.J. 1969. Mapping diversity: a comparative study of some numerical methods, pp. 1–30. **In:** A.J. Cole (ed.). Numerical taxonomy. Academic Press, London, UK.

Brooks, D.R. 1985. Historical ecology: A new approach to studying the evolution of ecological associations. Ann. Missouri Bot. Gard. 72:660–680.

Brooks, D.R. and E.O. Wiley. 1986. **Evolution as entropy. Toward a unified theory of biology.** Univ. Chicago Press, Chicago, IL.

Brower, J.C. and J. Veinus. 1978. Multivariate analysis of allometry using point coordinates. J. Paleontol. 52:1037–1053.

Brown, C.J.D. 1971. **Fishes of Montana**. Montana State Univ. Press, Bozeman, MT.

Brown, J.H. and A.C. Gibson. 1983. **Biogeography**. Mosby, St.Louis, MO.

Brown, K. 1983. Do life history tactics exist at the intraspecific level? Data from freshwater snails. Amer. Nat. 121:871–879.

Brown, V. and R.G. Davies. 1972. Allometric growth in two species of *Ectobius* (Dictyoptera: Blattidae). J. Zool. 166:97–132.

Brown, W.L. and E.O. Wilson. 1954. The case against the trinomen. Syst. Zool. 3:174–176.

Brundin, L. 1972. Evolution, causal biology and classification. Zool. Scripta 1:107-120.

Bryant, E.H. 1986. On the use of logarithms to accomodate scale. Syst. Zool. 35:552–559.

Bunker, R.C. 1982. Evidence of multiple late-Wisconsin floods from Glacial Lake Missoula in Badger Coulee, Washington. Quat. Res. 18:17–31.

Burnaby, T.P. 1966. Growth-invariant discriminant functions and generalized distances. Biometrics 22:96–110.

Calder, W.A. ,III. 1983. **Size, function and life history**. Harvard Univ. Press, Cambridge, MA.

Cain, S.A. 1944. **Foundations of plant geography**. Harper and Row, New York, NY.

Calder, J.A. and R.S. Taylor. 1968. FLora of the Queen Charlotte Islands. Part I. Systematics of the vascular plants. Can. Dept. Agric. Monogr. 4, part I.

Campbell, N.A. 1976. A multivariate approach to variation in microfilariae: Examination of the species *Wuchereria lewisi* and demes of the species *W. bancrofti*. Aust. J. Zool. 24:105–114.

Campbell, N.A. 1980. Robust procedures in multivarial analysis. I. Robust covariance estimation. Appl. Stat. 29:231–237.

Carl, G.C., W.A. Clemens and C.C. Lindsey. 1977. **The fresh-water fishes of British Columbia**. B.C. Prov. Mus. Handbook no. 5., Victoria, B.C..

Cattell, R.B. 1966. The scree test for the number of factors. Multivar. Behav. Res. 1:245–276.

Cavender, T.M. 1970. A comparison of coregonines and other salmonids with the earliest known teleost fishes, pp.1–32. **In**: C.C. Lindsey and C.S. Woods (eds.), Biology of coregonid fishes, Univ. Manitoba Press, Winnipeg.

Cavender, T.M. 1978. Taxonomy and distribution of the bull trout *Salvelinus confluentus* (Suckley), from the American Northwest. Calif. Fish Game 64:139–174.

Cavender, T.M. 1980. Systematics of *Salvelinus* from the North Pacific Basin, pp. 295–322. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Cavender, T.M. 1984. Cytotaxonomy of North American *Salvelinus*, pp. 431–445. **In**: L. Johnson and B.L. Burns (eds.), Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Cavender, T.M. 1986. Review of the fossil history of North American freshwater fishes, pp. 699–724. **In**: C.H. Hocutt and E.O. Wiley (eds.), Zoogeography of North American freshwater fishes. J. Wiley and Sons, New York, NY.

Chambers, J.M., W.S. Cleveland, B. Kleiner and P.A. Tukey. 1983. **Graphical methods for data analysis**. Wadsworth International Group, Duxbury Press, Belmont, CA.

Chan, L.S. and O.J. Dunn. 1972. The treatment of missing values in discriminant analysis — 1. The sampling experiment. J.Amer. Stat. Assoc. 67:473–477.

Chang, H.S. and H.G. Gauch, Jr. 1986. Multivariate analysis of plant communities and environmental factors in Ngari, Tibet. Ecology 67:1568-1575.

Chatfield, C. and A.J. Collins. 1980. **Introduction to multivariate analysis.** Chapman and Hall, New York, NY.

Chayes, F. 1949. On ratio correlation in petrography. J. Geol. 57:239–245.

Chen, E.H. 1971. The power of the Shapiro-Wilk W test for normality in samples from contaminated normal distributions. J. Amer. Stat. Assoc. 66:760–762.

Chereshnev, I.A. 1982. The taxonomic status of sympatric diadromous charrs of the genus *Salvelinus* (Salmonidae) from eastern Chukota. J. Ichthyol. 22(6):22–38.

Chernenko, Ye.U. and R.M. Viktorovsky. 1971. Chromosome sets of the masu salmon, the Siberian char and the southern Dolly Varden char. Nauch. soobshch. in-ta Biol. Morya 2:232–235 [In Russian].

Chernoff, B. 1982. Character variation among populations and the analysis of biogeography. Amer. Zool. 22:425–439.

Cheverud, J.M. 1982a. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

Cheverud, J.M. 1982b. Relationships among ontogenetic static and evolutionary allometry. Amer. J. Phys. Anthr. 59:139–149.

Cheverud, J.M., M.M. Dow and W. Leutenegger. 1985. The quantitive assessment of phylogenetic constraints in comparative analyses: Sexual dimorphism in body weight among primates. Evolution 39:1335–1351.

Cheverund, J.M., J.J. Rutledge and W.R. Atchley. 1983. Quantitative genetics of development: Genetic correlations among age-specific trait values and the evolution of ontogeny. Evolution 37:895–905.

Christiansen, E.A. 1979. The Wisconsin deglaciation of southern Saskatchewan and adjacent areas. Can. J. Earth Sci. 16:913–938.

Christensen, K. 1954. Ratios as a means of specific differentiation in Collembola. Ent. News 65:176–177.

Clague, J.J., J.E. Armstrong and W.H. Mathews. 1980. Advance of the Late Wisconsin Cordilleran ice sheet in southern British Columbia since 22,000 B.P.. Quat. Res. 13:322–326.

Clague, J.J. and V.N. Rampton. 1982. Neoglacial Lake Alsek. Can. J. Earth Sci. 19:94–117.

Clark, D.W. and J.E. McInerney. 1974. Emigration of the peamouth chub, *Mylocheilus caurinus*, across a dilute seawater bridge: an experimental zoogeographic study. Can. J. Zool. 52:457–469.

Clarke, M.R.B. 1980. The reduced major axis of a bivariate sample. Biometrika 67:441–446.

Clayton, J.W. and P.E. Ihssen. 1980. Dehydrogenase isozymes in *Salvelinus*: Genetics and interspecies phenotypic comparisons, pp. 339–356. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Claytor,R.R. and H.R. MacCrimmon. 1987. Partitioning size from morphometric data: A comparison of five statistical procedures used in fisheries stock identification research. Can. Tech. Rep. Fish. Aquat. Sci. 1531.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. J. Amer. Stat. Assoc. 74:829–836.

Clifford, H.T. and F.E. Binet. 1954. A quantitative study of a presumed hybrid swarm between *Eucalyptus elaeophora* and *E. goniocalyx*. Aust. J. Bot. 2:325–336.

Clutton-Brock, T.H. and P.H. Harvey. 1977. Primate ecology and social organization. J. Zool. 183:1–39.

Clutton-Brock, T.H. and P.H. Harvey. 1979. Comparison and adaptation. Proc. Roy. Soc. London 205(B):547–565.

Clutton-Brock, T. and P. Harvey. 1984. Comparative approaches to investigating adaptation, pp.7–29. **In**: J. Krebs and N. Davies (eds.), Behavioural ecology: An evolutionary approach. 2nd edition. Sinauer Press, New York, NY.

Cochran, W.G. 1977. **Sampling techniques**. 3rd ed. John Wiley and Sons, Toronto, Ont.

Cock, A.G. 1966. Genetical aspects of metrical growth and form in animals. Quart. Rev. Biol. 41:131–190.

Cooley, W.W. and P.R. Lohnes. 1971. **Multivariate data analysis**. John Wiley and Sons, New York, NY.

Corruccini, R.S. 1973. Size and shape in similarity coefficients based on metric characters. Amer. J. Phys. Anthr. 38:743–753.

Corruccini, R.S. 1975. Multivariate analysis in biological anthropology: Some considerations. J. Hum. Evol. 4:1–19.

Corruccini, R.S. 1977. Correlation properties of morphometric ratios. Syst. Zool. 26:211–214.

Corruccini, R.S. 1978. Morphometric analyses: Uses and abuses. Yearb. Phys. Anthrop. 21:134–150.

Corruccini, R.S. 1983. Principal components for allometric analysis. Am. J. Phys. Anthrop. 60:451–453.

Corruccini, R.S. 1987. Univariate versus multivariate morphometric variation: An alternate viewpoint. Syst. Zool. 36:396–397.

Cox, C.B. and P.D. Moore. 1985. **Biogeography: An ecological and evolutionary approach**. 4th edition. Blackwell Scientific Publs., Palo Alto, CA.

Cox, D.R. 1968. Notes on some aspects of regression analysis. J. Roy. Stat. Soc. 131A:265–279.

Cox, D.R. and N.J.H. Small. 1978. Testing multivariate normality. Biometrika 65(2):263–272.

Crain, I.K. and K. Bhattacharyya. 1967. Treatment of non-equispaced two-dimensional data with a digital computer. Geoexploration 5:173–194.

Cracraft, J. 1983. Species concepts and speciation analysis, pp. 159–187. **In**: R.F. Johnston (ed.), Current ornithology. Volume 1. Plenum Press, New York, NY.

Cracraft, J. 1987. Species concepts and the ontology of evolution. Biol. Philos. 2:329–346.

Craw, R.C. and P. Weston. 1984. Panbiogeography: A progressive research program? Syst. Zool. 33:1–13.

Creighton, G.K. and R.E. Strauss. 1986. Comparative patterns of growth and development in cricetine rodents and the evolution of ontogeny. Evolution 40:94–106.

Croizat, L. 1962. **Space, time and form: The biological synthesis**. Privately published by the author; Caracas, Venezuela.

Croizat, L., G. Nelson and D.E. Rosen. 1974. Centers of origin and related concepts. Syst. Zool. 23:265–287.

Crossman, E.J. and D.E. McAllister. 1986. Zoogeography of freshwater fishes of the Hudson Bay drainage, Ungave Bay and the Arctic Archipelago, pp. 53–104. **In**: C.H. Hocutt and E.O. Wiley (eds.), Zoogeography of North American freshwater fishes. J. Wiley and Sons, New York, NY.

Crovello, T.J. 1970. Analysis of character variation in ecology and systematics. Ann. Rev. Ecol. Syst. 1:55–98.

Crovello, T.J. 1981. Quantitative biogeography: An overview. Taxon 30:563–575.

Croy, C.D. and R.L. Dix. 1984. Notes on sample size requirements in morphological plant ecology. Ecology 65:662–666.

D'Agostino, R. and E.S. Pearson. 1973. Tests for departure from normality. Empirical results for the distribution of b_2 and $\sqrt{b_1}$. Biometrika 60:613–622.

Darwin, C. 1859. **On the origin of species.** Fascimile of the first edition. Harvard Univ. Press, Cambridge, MA.

Davis, B.L. and R.J. Baker. 1974. Morphometrics, evolution and cytotaxonomy of mainland bats of the genus *Macrotus* (Chiroptera: Phyllostomatidae). Syst. Zool. 23:26–39.

Day, N.E. 1969. Divisive cluster analysis and a test for multivariate normality. Bull. Inst. Int. Stat. 43:110–112.

de Candolle, A.P. 1820. Geographie botanique. **In**: Dictionnaire des sciences naturelles, vol. 18. Strasbourg and Paris, France.

de Beer, G.R. 1958a. Darwin's views on the relations between embryology and evolution. J. Linn. Soc. London. 44:15–23.

de Beer, G.R. 1958b. **Embryos and ancestors.** Oxford Univ. Press, Oxford.

DeLacy, A.C. and W.M. Morton. 1943. Taxonomy and habits of the charrs *Salvelinus malma* and *Salvelinus alpinus* of the Karluk drainage system. Trans. Amer. Fish. Soc. 72:79–91.

Delaney, M.J. and M.J.R. Healy. 1964. Variation in the long — tailed field mouse (*Apodemus sylvaticus* (L.)) in north-west Scotland. II. Simultaneous examination of all characters. Proc. Roy. Soc. B. 161:200–207.

Dempson, J.B. 1984. Identification of anadromous Arctic charr stocks in coastal areas of northern Labrador, pp. 143–162. **In**: L.Johnson and B.L. Burns (eds.), Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Dempster, A.P. 1971. An overview of multivariate data analysis. J. Multiv. Analysis 1:315–346.

Desarbo, W.S. 1981. Canonical/redundancy factoring analysis. Psychometrika 46:307–329.

Dobson, F.S. 1985. The use of phylogeny in behavior and ecology. Evolution 39:1384–1388.

Dodson, P. 1978. On the use of ratios in growth studies. Syst. Zool. 27:62–67.

Dougenik, J.A. and D.E. Sheehan. 1979. **SYMAP user's reference manual.** Laboratory for computer graphics and spatial analysis, Harvard Univ. Graduate School of Design, Cambridge, MA.

Draper, N.R. and H. Smith. 1981. **Applied regression analysis.** John Wiley and Sons, New York, NY.

Dudzinski, M.L., J.M. Norris, J.T. Chmura and C.B.H. Edwards. 1975. Repeatability of principal components in samples: Normal and non-normal data sets compared. Multiv. Behav. Rev. 10:109–117.

Duellmann, W.E. 1985. Reproductive modes in anuran amphibians: Phylogenetic significance of adaptive strategies. South Afr. J. Sci. 8:174–178.

Dunham, A.E. and D.B. Miles. 1985. Patterns of covariation in the life history traits of squamate reptiles: The effects of size and phylogeny reconsidered. Amer. Nat. 126:231–257.

Dunn, O.J. 1971. Some expected values for probabilities of correct classification in discriminant analysis. Technometrics 13:345.

Dunn, O.J. and P.D. Vardy. 1966. Probablities of correct classification in discriminant analysis. Biometrics 22:908–924.

Dymond, J.R. 1932. The trout and other game fishes of British Columbia. Bull. Biol. Bd. Can. 32:1–51.

Ehrlich, P. 1961. Has the biological species concept outlived its usefulness? Syst. Zool. 10:167–176.

Eickwort, K. 1969. Differential variation of males and females in *Polistes exclamans*. Evolution 23:391–405.

Eisen, E.J. 1975. Results of growth curve analysis in mice and rats. J. Anim. Sci. 42:1008–1023.

Eisenbis, R.A., G.G. Gilbert and R.B. Avery. 1973. Investigating the relative importance of individual variables and variable subsets in discriminant analysis. Comm. Stat. 2:205–219.

Eldredge, N. and S.J. Gould. 1972. Punctuated equilibria: An alternative to phyletic gradualism, pp. 82–115. **In**: T.J.M. Schopf (ed.), Models in paleontology. Freeman, Cooper and Co., San Francisco, CA.

Emerson, S.B. 1986. Heterochrony and frogs: The relationship of a life-history trait to morphological form. Amer. Nat. 127:167–183.

Endler, J.A. 1977. **Geographic variation, speciation, and clines**. Monogr. Pop. Biol. 10. Princeton Univ. Press, Princeton, NJ.

Endler, J.A. 1982a. Alternative hypotheses in biogeography: Introduction and synopsis of the symposium. Amer. Zool. 22:349–354.

Endler, J.A. 1982b. Problems in distinguishing historical from ecological factors in biogeography. Amer. Zool. 22:441–452.

Everitt, B. 1978. **Graphical techniques for multivariate data**. Heinemann Educational Books, London, UK.

Eyles, A.C. and R.E. Blackith. 1965. Studies on hybridization in *Scolopostethus* Fieber (Heteroptera Lygeidae). Evolution 19:465–479.

Facchin, A. and G. King. 1980. Lake survey and stocking record for the lower mainland region of British Columbia. B.C. Fish and Wildlife Branch Techn. Circ. no. 47.

Falconer, D.S. 1981. **Introduction to quantitative genetics**. 2nd ed. Longman House, London, UK.

Farris, J.S. 1973. On the use of the parsimony criterion for inferring evolutionary trees. Syst. Zool. 22:250–256.

Farver, T.B. and O.J. Dunn. 1979. Stepwise variable selection in classification problems. Biometrical J. 21:145–153.

Felsenstein, J. 1978. Cases in which parsimony or compatability methods will be positively misleading. Syst. Zool. 27:401–410.

Felsenstein, J. 1983. Parsimony in systematics: Biological and statistical issues. Ann. Rev. Ecol. Syst. 14:313–333.

Felsenstein, J. 1985. Phylogenies and the comparative method. Amer. Nat. 125:1–15.

Felsenstein, J. and E. Sober. 1986. Parsimony and likelihood: An exchange. Syst. Zool. 35:617–626.

Ferguson, A. 1981. Systematics of Irish charr as evidenced by electrophoretic analysis of tissue proteins. Biochem. System. Ecol. 9:225–232.

Filliben, J.J. 1975. The probability plot correlation coefficient test for normality. Technometrics 17:111–117.

Fink, W.L. 1982. The conceptual relationship between ontogeny and phylogeny. Paleobiology 8:254–264.

Fink, W.L. and S.H. Weitzman. 1982. Relationships of the stomiiform fishes (Teleostei), with a description of *Diplophos*. Bull. Mus. Comp. Zool. Harv. Univ. 150:31–93.

Fisher, D.R. 1968. A study of faunal resemblance using numerical taxonomy and factor analysis. Syst. Zool. 17:48–63.

Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. Annals Eugenics 7:179–188.

Fox, D.J. and K.E. Guire. 1976. **Documentation for MIDAS**. Statistical Research Laboratory, Univ. of Michigan, Ann Arbor, MI.

Fraser, N.C. and G. Power. 1984. The interactive segregation of landlocked Arctic charr (*Salvelinus alpinus*) from lake charr (*S. namaycush*) and brook charr (*S. fontinalis*) in two lakes of subarctic Quebec, Canada, pp. 163–181. **In**: L. Johnson and B.L. Burns (eds.), Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Friedman, J.H. and L.C. Rafsky. 1981. Graphics for the multivariate two-sample problem. J. Amer. Stat. Assoc. 76:277–295.

Frohne, I.V. 1973. Statistical analyses of discrete morphology in northern populations of the fish genus *Salvelinus*. Biol. Pap. Univ. Alaska 13:10–20.

Frost, W.E. 1965. Breeding habits of Windermere charr, *Salvelinus willughbii* (Günther), and their bearing on speciation of these fish. Proc. R. Soc. Ser. B. Biol. Sci. 163:232–284.

Fulton, R.J. 1969. Glacial lake history, southern interior plateau, British Columbia. Geol. Surv. Can. Pap. 37–69.

Funk, V.A. 1985. Phylogenetic patterns and hybridization. Ann. Missouri Bot. Garden. 72:681–715.

Gabriel, K.R. 1985. Multivariate graphics, pp. 66–79. **In**: S. Kotz and N.L. Johnson (eds.), Encyclopedia of Statistical Sciences. vol. 6. John Wiley and Sons, New York, NY.

Garstang, W. 1922. The theory of recapitulation: A critical restatement of the biogenetic law. J. Linn. Soc. Zool. 35:81–101.

Gerhart, J.C., S. Berking, J. Cooke, G.L. Freeman, A. Hildebrandt, H. Jokusch, P.A. Lawrence, C. Nusslein-Volhard, G.F. Oster, K. Sander, H.W. Sauer, G.S. Stent, N.K. Wessells and L. Wolpert. 1982. The cellular basis of morphogenetic change , pp. 87–114. **In**: J.T. Bonner (ed.), Evolution and development. Springer-Verlag, Berlin.

Gibson, A.R., A.J. Baker and A. Moeed. 1984. Morphometric variation in introduced populations of the common myna (*Acridotheres tristis*): An application of the jackknife to principal component analysis. Syst. Zool. 33:408–421.

Gilbert, E.S. 1968. On discrimination using qualitative variables. J. Amer. Stat. Assoc. 63:1399.

Gilbert, E.S. 1969. The effect of unequal variance covariance matrices on Fisher's linear discriminant function. Biometrics 25:505–516.

Girard, C.F. 1856. Notice upon the species of the genus *Salmo* of authors, observed chiefly in Oregon and California. Proc. Acad. Nat. Sci. Philadelphia 8:217–220.

Gittens, R. 1968. Trend-surface analysis of ecological data. J. Ecol. 56:845–869.

Gittins, R. 1979. Ecological applications of canonical analysis, pp. 309–535. **In**: L. Orloci, C.R. Rao and W.M. Stiteler (eds.), Multivariate methods in ecological work. Int. Co-op. Publ. House, Fairland, MD.

Glahn, H.R. 1968. Canonical correlation and its relationship to discriminant analysis and multiple regression. J. Atmos. Sci. 25:23–31.

Glubokovsky, M.K. and I.A. Chereshnev. 1981. Unresolved problems concerning the phylogeny of chars (*Salvelinus*) of the Holarctic: I. Migratory chars of the East Siberian Sea basin. J. Ichthyol. 21(6):1–15.

Gnanadesikan, R. 1977.**Methods for statistical data analysis of multivariate observations**. John Wiley and Sons, New York, NY.

Gnanadesikan, R. and J.R. Kettering. 1972. Robust estimates, residuals and outlier detection with multiresponse data. Biometrics 28:81–124.

Gnanadesikan, R. and M.B. Wilk. 1969. Data analytic methods in multivariate statistical analysis, pp. 593–638. **In**: P.R. Krishnaiah (ed.), Multivariate analysis II. Academic Press, New York, NY.

Goldschmidt, R. 1940. **The material basis of evolution.** Yale Univ. Press, New Haven, CT.

Goodwin, B.C. 1982. Development and evolution. J. Theor. Biol. 97:43–55.

Goodwin, B.C., N. Holder and C.G. Wylie (eds.). 1983. **Development and evolution.** Cambridge Univ. Press, Cambridge, UK.

Gould, S.J. 1966. Allometry and size on ontogeny and phylogeny. Biol. Rev. 41:587–640.

Gould, S.J. 1968. Ontogeny and the explanation of form: An allometric analysis. **In:** D.B. Macurda (ed.), Paleobiological aspects of growth and development, a symposium. Paleont. Soc. Mem. 2:81–98.

Gould, S.J. 1971. Geometric similarity in allometric growth: A contribution to the problem of scaling in the evolution of size. Amer. Nat. 105:113–136.

Gould, S.J. 1977. **Ontogeny and phylogeny.** Harvard Univ.Press, Cambridge, MA.

Gould, S.J. and R.F. Johnston. 1972. Geographic variation. Ann. Rev. Ecol. Syst. 3:457–498.

Gould, S.J., D.F. Woodruff and J.P. Martin. 1974. Genetics and morphometrics of *Cerion* at Pongo Carpet: A new systematic approach to this enigmatic snail. Syst. Zool. 23:518–535.

Gould, W.R. 1987. Features in the early development of bull trout (*Salvelinus confluentus*). Northwest Sci. 61:264–268.

Gower, J.C. 1967. Multivariate analysis and multidimensional geometry. Statistician 17:13–28.

Gower, J.C. 1972. Measures of taxonomic distance and their analysis, pp. 1–24. **In:** J.S. Weiner and J. Huizinga (eds.), The assessment of population affinities in man. Clarendon, Oxford.

Gower, J.C. 1976. Growth-free canonical variates and generalized inverses. Bull. Geol. Inst. Univ. Uppsala, New Series 7:1–10.

Gower, J.C. and G.J.S. Ross. 1969. Minimum spanning trees and single linkage cluster analysis. Appl. Stat. 18:54–64.

Grady, J.M., R.C. Cashner and J.S. Rogers. 1983. Fishes of the Bayou Sara drainage, Louisiana and Mississippi, with a discriminant function analysis of factors influencing species distribution. Tulane Stud. Zool. Bot. 24:83–100.

Green, P.E. 1978. **Analyzing multivariate data.** Dryden Press, Hinsdale, Illinois, IL.

Green, R.H. 1971. A multivariate data statistical approach to the Hutchinson niche: bivalve molluscs of central Canada. Ecology 52:543–556.

Green, R.H. 1979. **Sampling design and statistical methods for environmental biologists.** John Wiley and Sons, New York, NY.

Guerrant, E.O., Jr. 1982. Neotenic evolution of *Delphinium nudicaule* (Ranunculaceae): A hummingbird-pollinated larkspur. Evolution 36:699–712.

Guptill, S.C. and L.E. Starr. 1988. Making maps with computers. Amer. Sci. 76:136–145.

Haeckel, E. 1866. **Generelle Morphologie der Organismen: Allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie.** Volumes 1 and 2. G. Reimer, Berlin, Germany.

Habbema, J.D.F. and J. Hermans 1977. Selection of variables in discriminant analysis by F-statistic and error rate. Technometrics 19:487–493.

Hagmeier, E.M. 1966. Numerical analysis of distributional patterns of North American mammals. 2. Re-evaluation of the provinces. Syst. Zool. 15:279–299.

Hall, B.K. 1982. How is mandibular growth controlled during development and evolution? J. Craniofacial Gen. Dev. Biol. 2:45–49.

Hall, B.K. 1984. Developmental processes underlying heterochrony as an evolutionary mechanism. Can, J. Zool. 62:1–7.

Hamburger, V. 1980. Embryology and the modern synthesis in evolutionary biology, pp. 96–112. **In**: E. Mayr and W.B. Provine (eds.), The evolutionary synthesis: Perspectives on the unification of biology. Harvard Univ. Press, Cambridge, MA.

Hammar, J. 1984. Ecological characters of different combinations of sympatric populations of Arctic char in Sweden, pp. 35–63. **In**: L. Johnson and B.L. Burns (eds.), Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Harris, R.J. 1975. **A primer of multivariate statistics.** John Wiley and Sons, New York, NY.

Hartley, S.E. 1987. The chromosomes of salmonid fishes. Biol. Rev. 62:197–214.

Harvey, P.H. 1982. On rethinking allometry. J. Theor. Biol. 95:37–41.

Hatheway, W.H. 1962. A weighted hybrid index. Evolution 16:1–10.

Hayami, I. and A. Matsukuma. 1970. Variation of bivariate characters from the standpoint of allometry. Paleontology 13:588–605.

Healey, J.R. 1968. Multivariate normal plotting. Appl. Stat. 17:157–161.

Healy, M.J.R. 1969. Rao's paradox concerning multivariate tests of significance. Biometrics 25:411–413.

Henderson, M.A. and T.G. Northcote. 1985. Visual prey detection and foraging in sympatric cutthroat trout (*Salmo clarki clarki*) and Dolly Varden (*Salvelinus malma*). Can. J. Fish. Aqu. Sci. 42:785–790.

Hennig, W. 1966. **Phylogenetic systematics.** Univ. Illinois Press, Urbana, IL.

Henricson, J.and L. Nyman. 1976. The ecological and genetical seregation of two sympatric species of dwarfed char (*Salvelinus alpinus* (L.) species complex), Rep. Inst. Freshw. Res. Drottningholm 55:15–37.

Herbert, J.G., J.F. Kidwell and H.B. Chase. 1979. The inheritance of growth and form in the mouse. IV. Changes in the variance components of weight, tail length and tail width during growth. Growth 43:36–46.

Hertwig, O. 1894. **The biological problem of today: preformation or epigenesis? The basis of a theory of organic development.** MacMillan Press, New York, NY.

Heusser, C.J. 1960. Late Pleistocene environments of north Pacific North America. Amer. Geogr. Soc. Spec. Publ. 35.

Hiernaux, J. 1972. The analysis of multivariate biological distances between human populations: Principles and applications of sub-Saharan Africa, pp. 96–114. **In**: J.S. Weiner and J. Huizinga (eds.), The assessment of population affinities in man. Clarendon Press, Oxford.

Hill, M.O. and H.G. Gauch, Jr. 1980. Detrended correspondence analysis: an improved ordination technique. Vegetatio 47:47–58.

Hills, M. 1969. On looking at large correlation matrices. Biometrika 56:249–253.

Hills, M. 1978. On ratios — a response to Atchley, Gaskins, and Anderson. Syst. Zool. 27:61–62.

Hills, M. 1982. Bivariate versus multivariate allometry: A note on a paper by Jungers and German. Amer. J. Phys. Anthrop. 59:321–322.

Hindar, K. and B. Jonsson. 1982. Habitat and food segregation of dwarf and normal Arctic charr (*Salvelinus alpnius*) from Vangsvatnet Lake, Western Norway. Can¿ J. Fish. Aquat. Sci. 39:1030–1045.

Hindar, K., B. Jonsson, J.H. Andrew, and T.G. Northcote. 1988. Resource utilization of sympatric and experimentally allopatric cutthroat trout and Dolly Varden charr. Oecologia 74:481–491.

Hindar, K., N. Ryman and G. Stahl. 1986. Genetic differentiation among local populations and morphotypes of Arctic charr, *Salvelinus alpinus*. Biol. J. Linn. Soc. 27:269–285.

Ho, M.W. and P.T. Saunders. 1979. Beyond neo-Darwinism: an epigenetic approach to evolution. J. Theor. Biol. 78:573–591.

Hodkinson, I.D. 1980. Present-day distribution patterns of the Holarctic Psylloidea (homoptera: Insects) with particular reference to the origin of the Nearctic fauna. J. Biogeogr. 7:127–146.

Hoffman, R.S. 1981. Different voles for different holes: Environmental restrictions on refugial survival of mammals, pp. 25–45. **In**: G.G.E. Scudder and J.L. Reveal (eds.), Evolution today, Proceedings of the 2nd Int. Congress of Syst. Evol. Biol.. Hunt Inst. Bot. Doc., Carnegie-Mellon Univ., Pittsburgh, PA.

Hoffman, R.S., J.W. Koeppl and C.F. Nadler. 1979. The relationships of the Amphiberigian marmots (Mammalia: Sciuridae). Occ. Pap. Mus. Nat. Hist. Univ. Kansas 83:1–56.

Holčik, J. 1982. Towards the characteristics of the genera *Hucho* and *Brachymystax* (Pisces, Salmonidae). Folia Zool. 31:368-380.

Holland, D. A. 1968. Component analysis: An aid to the interpretation of data. Exp. Agric. 5:151–164.

Holland, S.S. 1964. Landforms of British Columbia: A physiography outline. Bull. Brit. Col. Dept. Mines Petr. Res. 48:1–138.

Holloway, J.D. and N. Jardine. 1968. Two approaches to zoogeography: A study based on the distributions of butterflies, birds and bats in the Indo-Australian region. Proc. Linn. Soc. London 179:153–188.

Holloway, L.N. and O.J. Dunn. 1967. The robustness of Hotelling's T^2. J. Amer. Stat. Assoc. 62:124–136.

Holmes, J.M.C. 1975. A comparison of numerical taxonomic techniques using measurements on the genera *Gammarus* and *Marinogammarus* (Amphipoda). Biol. J. Linn. Soc. 7:183–214.

Holsinger, K.E. 1984. The nature of the biological species. Philos. Sci. 51:293–307.

Hopkins, D.M. 1959. Cenozoic history of the Bering land bridge. Science 129:1519–1528.

Hopkins, D.M. 1972. The paleogeography and climate history of Beringia during late Cenozoic time. Internord 12:121–150.

Hopkins, D.M. 1973. Sea level history in Beringia during the past 250,000 years. Quat. Res. 3:520–540.

Horn, J.L. and R. Engstrom. 1979. Cattell's scree test in relation to Bartlett's chi-square test and other observations on the number of factors problem. Multivar. Behav. Res. 14:283–300.

Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. J. Educ. Psych. 24:417–441.

Hotelling, H. 1936. Relations between two sets of variates. Biometrika 28:321–377.

Howes, D.E. 1982. Late Quaternary sediments and geomorphic history of northern Vancouver Island, British Columbia. Can. J. Earth Sci. 20:57–65.

Hubbs, C.L. 1926. The structural consequences of the developmental rate in fishes, considered in reference to certain problems in evolution. Amer. Nat. 60:57–81.

Hubbs, C.L. 1943. Criteria for subspecies, species and genera, as determined by researches on fishes. Ann. N.Y. Acad. Sci. 44:109–121.

Hubbs, C.L. and K.F. Lagler. 1958. **Fishes of the Great Lakes Region.** Univ. Michigan Press, Ann Arbor, MI.

Hubbs, C.L. and R.R. Miller. 1948a. Correlation between fish distribution and hydrographic history in the desert basins of western United States, pp. 17–144. **In**: The Great Basin, with emphasis on glacial and postglacial times. Bull. Univ. Utah 38, Biol. Ser. 10(7).

Hubbs, C.L. and R.R. Miller. 1948b. Two new relict genera of cyprinid fishes from Nevada. Occ. Pap. Mus. Zool. Univ. Michigan 507:1–30.

Hughes, R.M., E. Rexstad and C.E. Bond. 1987. The relationship of aquatic ecoregions, river basins and physiographic provinces to the ichthyogeographic regions of Oregon. Copeia 1987:423–432.

Huheey, J.E. 1966. A mathematical method of analyzing biogeographical data. I. Herpetofauna of Illinois. Amer. Midl. Nat. 73:490–500.

Hull, D.L. 1970. Contemporary systematic philosophies. Ann. Rev. Ecol. Syst. 1:19–54.

Hull, D.L. 1978. A matter of individuality. Phil. Sci. 45:335–360.

Hume, J.M.B. and T.G. Northcote. 1985. Initial changes in use of space and food by experimentally segregated populations of Dolly Varden (*Salvelinus malma*) and cutthroat trout (*Salmo clarki*). Can. J. Fish. Aquat. Sci. 42:101–109.

Humphries, C.J. and L.R. Parenti (eds.). 1986. **Cladistic biogeography.** Oxford Monographs on Biogeography 2. Oxford Univ. Press, Oxford, UK.

Humphries, J.M., F.L. Bookstein, B. Chernoff, G.R. Smith, R.L. Elder and S.G. Poss. 1981. Multivariate discrimination by shape in relation to size. Syst. Zool. 30:291–308.

Huxley, J.S. 1932. **Problems of relative growth.** The Dial Press, New York, NY.

Huxley, J.S. and G. Teissier. 1936. Terminology of relative growth. Nature 137:780.

Imbrie, J. 1956. Biometrical methods in the study of invertebrate fossils. Bull. Amer. Mus. Nat. Hist. 108:215–252.

Imbrie, J. and N.G. Kipp. 1971. A new micropaleontological method for quantitative paleoclimatology: Application to a Late Pleistocene Caribbean core, pp. 71–181. **In:** K.K. Turekian (ed.), The Late Cenozoic glacial ages. Yale Univ. Press, New Haven, CT.

Ishigaki, K. 1969. **Ecology and morphology of genus *Salvelinus* in Hokkaido.** Ph.D. thesis. Hokkaido Univ., Sapporo, Hokkaido. [In Japanese].

Ito, K. 1969. On the effect of heteroscedasticity and nonnormality upon some multivariate test procedures, pp. 87–120. **In:** P.R. Krishnaiah (ed.), Multivariate Analysis — II. Proc. Second. Int. Symp. Multiv. Analysis. Academic Press, New York, NY.

Ito, K. and W.J. Schull. 1964. On the robustness of the T^2 test in multivariate analysis of variance when variance-covariance matrices are not equal. Biometrika 51:71–82.

Jain,A.K. and W.G. Waller. 1979. On the optimal number of features in the classification of multivariate Gaussian data. Pattern Recognition 10:365–374.

Jamison, P.L. and S.L. Zegura. 1974. A univariate and multivariate examination of measurement error in anthropometry. Amer. J. Phys. Anthr. 40:197–204.

Jeffers, J.N.R. 1967. The study of variation in taxonomic research. Statistician 17:29–43.

Johannson, J.K. 1981. An extension of Wollenberg's redundancy analysis. Psychometrika 46:93–103.

Johnson, L. 1980. The Arctic char, *Salvelinus alpinus*, pp. 15–98. **In:** E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Johnson, L. and B.Burns (eds.). 1984. **Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr.** Univ. Manitoba Press, Winnipeg, Manitoba.

Johnston, R.F. 1969. Character variation and adaptation in European sparrows. Syst. Zool. 18:206–231.

Johnston, R.F. 1973. Evolution in the house sparrow. IV. Replicate studies in phenetic covariation. Syst. Zool. 22:219–226.

Jolicoeur, P. 1959. Multivariate geographical variation in the wolf *Canis lupus* L. Evolution 13:283–299.

Jolicoeur, P. 1963a. The multivariate generalization of the allometry equation. Biometrics 19:497–499.

Jolicoeur, P. 1963b. The degree of generality of robustness in *Martes americana*. Growth 27:1–27.

Jolicoeur, P. 1965. Calcul d'un intervalle de confiance pour la pente de l'axe de la distribution normale de deux variables. Biometrie-Praximetrie 6:31–35.

Jolicoeur, P. and J.E. Mosimann. 1960. Size and shape variation in the painted turtle: A principal component analysis. Growth 24:339–354.

Jolliffe, I.T. 1972. Discarding variables in a principal component analysis. I. Artificial data. Appl. Stat. 21:160–173.

Jolliffe, I.T. 1973. Discarding variables in a principal component analysis. II. Real data. Appl. Stat. 22:21–31.

Jonsson, B., K. Hindar, and T.G. Northcote. 1984. Optimal age at sexual maturity of sympatric and experimentally allopatric cutthroat trout and Dolly Varden charr. Oecologia 61:319–325.

Jordan, D.S. 1879. Notes on a collection of fishes from the Clackamas River, Oregon. Proc. U.S. Nat. Mus. 1878 1:69–85.

Jordan, D.S. and B.W. Evermann. 1896. A check-list of the fishes and fish-like vertebrates on North and Middle America. Rep. U.S. Fish. Comm. 1895 (1896):207–584.

Jordan, D.S., B.W. Evermann and H.W. Clark. 1930. Check list of the fishes of North and Middle America. Rep. U.S. Comm. Fish (for 1928)—reprint 155, pp. 59–61.

Jordon, D.S. and C.H. Gilbert. 1882. Synopsies of fishes of North America. U.S. Nat. Mus. Bull. 16.

Jöreskog, K.G., J.E. Klovan and R.E. Reyment. 1976. **Geological factor analysis: Methods in Geomathematics I.** Elsevier, Amsterdam.

Jungers, W.L. and R.Z. German. 1981. Ontogenetic and interspecific skeletal allometry in nonhuman primates: Bivariate versus multivariate analyses. Amer. J. Phys. Anthr. 55:195–202.

Kaiser, H.F. 1960. The application of electronic computers to factor analysis. Educ. Psych. Measur. 20:141–151.

Karlstrom, T.N.V. 1961. The glacial history of Alaska: Its bearing on paleoclimatic theory. Ann. N.Y. Acad. Sci. 95:290–340.

Katz, M.J. 1980. Allometry formula: A cellular model. Growth 44:89–96.

Katz, M.J. 1982. Ontogenetic mechanisms: The middle ground of evolution, pp. 207–212. **In:** J.T. Bonner (ed.), Evolution and development. Springer-Verlag, Berlin.

Kauffmann, S.A. 1983. Developmental constraints: Internal factors in evolution, pp. 195–225. **In:** B.C. Goodwin, N. Holder and C.G. Wylie (eds.), Development and Evolution. Cambridge Univ. Press, Cambridge, UK.

Kendall, M.G. and A. Stuart. 1961. **The advanced theory of statistics.** vol. 2. Griffin and Co., London, UK.

Kennedy, M.L. and G.D. Schnell. 1978. Geographic variation and sexual dimorphism in Ord's kangaroo rat, *Dipodomys ordii*. J. Mammol. 59:45–59.

Kermack, K.A. and J.B.S. Haldane. 1950. Organic correlation and allometry. Biometrika 37:30–41.

Key, K.H.L. 1981. Species, parapatry and the morabine grasshoppers. Syst. Zool. 30:425–428.

Khan, N.Y. and S.U. Qadri. 1971. Intraspecific variations and postglacial distribution of lake char (*Salvelinus namaycush*). J. Fish. Res. Bd. Can. 28:465–476.

Kidwell, J.F. and H.B. Chase. 1967. Fitting the allometric equation — A comparison of ten methods by computer simulation. Growth 31:165–179.

Kidwell, J.F., J.G. Herbert and H.B. Chase. 1979. The inheritance of growth and form in the mouse. V. Allometric growth. Growth 43:47–57.

Kimball, B.F. 1960. On the choice of plotting positions on probability paper. J. Amer. Stat. Assoc. 55:546–560.

Kircheis, F.W. 1976. Reproductive biology and early life history of the Sunapee trout of Floods Plain, Maine. Trans. Amer. Fish. Soc. 105:615–619.

Klein, D.R. 1965. Postglacial distribution patterns of mammals in the southern coastal regions of Alaska. Arctic 18:7–20.

Klemetsen, A. and P.E. Grotnes. 1975. Food and habitat segregation by two sympatric Arctic char populations. Verh. Internat. Verein. Limnol. 19:2521–2528.

Klemetsen, A. and P.E. Grotnes. 1980. Coexistence and immigration of two sympatric Arctic charr, pp. 757–763. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Kluge, A.G. and R.E. Strauss. 1985. Ontogeny and systematics. Ann. Rev. Ecol. Syst. 16:247–268.

Kolyushev, A. 1971. Some osteological characters of chars (genus *Salvelinus*) in connection with the problem of their systematic position. J. Ichthyol. 11:464–473.

Kowalski, C.J. 1972. A commentary on the use of multivariate statistical methods in anthropomorphic research. Amer. J. Phys. Anthrop. 36:119–132.

Krumbein, W.C. 1955. The statistical analysis of facies maps. J. Geol. 63:452–470.

Krumbein, W.C. 1959. Trend-surface analysis of contour-type maps with irregular control-point spacing. J. Geophys. Res. 64:823–834.

Krzanowski, W.J. 1977. The performance of Fisher's linear discriminanat function under nonoptimal conditions. Technometrics 19:191–200.

Kuhry, B. and L.F. Marcus. 1977. Bivariate linear models in biometry. Syst. Zool. 26:201–209.

Lachenbruch, P.A. 1975. **Discriminant analysis**. Hafner, New York, NY.

Lachenbruch, P.A. and M. Goldstein. 1979. Discriminanat analysis. Biometrics 35:69–85.

Lachenbruch, P.A., C. Sneeringer and L.T. Revo. 1973. Robustness of the linear and quadratic discriminanat functions to certain types of non-normality. Comm. Stat. 1:39–56.

Laird, A.K. 1965. Dynamics of relative growth. Growth 29:249–263.

Laird, A.K., S.A. Tyler and A.D. Barton. 1965. Dynamics of normal growth. Growth 29:233–248.

Laird, A.K., A.D. Barton and S.A. Tyler. 1968. Growth and time: Interpretation of allometry. Growth 32:347–354.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to the brain: Body size allometry. Evolution 33:402–416.

Lande, R. 1982. A quantitative genetic theory of life history evolution. Ecology 63:607–615.

Lande, R. 1985. Genetic and evolutionary aspects of allometry, pp. 21–32. **In**: W.L. Jungers (ed.), Size and scaling in primate biology. Plenum Press, New York, NY.

Lande, R. and S.J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Larson, A. 1980. Paedomorphosis in relation to rates of morphological and molecular evolution in the salamander *Aneides flavipunctatus* (Amphibia: Plethodontidae). Evolution 34:1–17.

Larsen, D.P., J.M. Omernik, R.M. Hughes, C.M. Rohm, T.R. Whittier, A.J. Kinney, A.L. Gallant and D.R. Dudley. 1986. Correspondence between spatial patterns in fish assemblages in Ohio streams and aquatic ecoregions. Env. Manage. 10:815–828.

Lauder, G.V. 1982. Historical biology and the problem of design. J. Theor. Biol. 97:57–67.

Lawrence, B. and W.H. Bossert. 1969. The cranial evidence for hybridization in New England Canis. Breviora 330:1–13.

Leamy, L. and D. Bradley. 1982. Static and growth allometry of morphometric traits in random-bred house mice. Evolution 36:1200–1212.

Leary, R.F., F.W. Allendorf and K.L. Knudsen. 1983. Consistently high meristic counts in natural hybrids between brook trout anf bull trout. Syst. Zool. 32:369–376.

Leary, R.F., F.W. Allendorf and K.L. Knudsen. 1985. Developmental instability and high meristic counts in interspecific hybrids of salmonid fishes. Evolution 39:1318–1326.

Lee, D.S., C.R. Gilbert, C.H. Hocutt, R.E. Jenkins, D.E. McAllister and J.R. Stauffer, Jr. 1980. **Atlas of North American freshwater fishes.** N.C. State Mus. Nat. Hist., Raleigh, NC.

Lee, P.J. 1969. The theory and application of canonical trend surfaces. J.Geol. 77:303–318.

Legendre, P. 1986. Reconstructing biogeographic history using phylogenetic-tree analysis of community structure. Syst. Zool. 35:68–80.

Legendre, P. and V. Legendre. 1983. Postglacial dispersal of freshwater fishes in the Quebec peninsula. Can. J. Fish. Aquat. Sci. 41:1781–1802.

Leggett, J.W. 1980. Reproductive ecology and behaviour of Dolly Varden charr in British Columbia, pp. 721–737. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Leland, H.V., J.L. Carter and S.V. Fend. 1986. Use of detrended correspondence analysis to evaluate factors controlling spatial distribution of benthic insects. Hydrobiologia 132:113–123.

Lerner, I.M. 1954. **Genetic homeostasis.** Oliver and Boyd, London, UK.

Lindsey, C.C. 1956. Distribution and taxonomy of fishes in the MacKenzie drainage of British Columbia. J. Fish. Res. Bd. Can. 13:759–789.

Lindsey, C.C. 1964. Problems in zoogeography of the lake trout, *Salvelinus namaycush*. J. Fish. Res. Bd. Can. 21:977–994.

Lindsey, C.C. 1975. Proglacial lakes and fish dispersal in southwestern Yukon Territory. Verh. Int. Verein. Limnol. 19:2364–2370.

Lindsey, C.C. and J.D. McPhail. 1986. Zoogeography of fishes of the Yukon and MacKenzie basins, pp. 639–674. **In**: C.H. Hocutt and E.O. Wiley (eds.), Zoogeography of North American freshwater fishes. J. Wiley and Sons, New York, NY.

Lohmann, G.P. 1983. Eigenshape analysis of microfossils: A general morphometric procedure for describing changes in shape. Math. Geol. 15(6):659–672.

Loudenslager, E.J. and G.H. Thorgaard. 1979. Karyotypic and evolutionary relationships of the Yellowstone (*Salmo clarki bouvieri*) and west-slope (*S. c. lewisi*) cutthroat trout. J. Fish. Res. Bd. Can. 36:630–635.

Løvtrup, S. 1974. **Epigenetics: A treatise on theoretical biology.** Wiley Publishers Ltd. London.

Lubischew, A. 1962. On the use of discriminant functions in taxonomy. Biometrics 18:455–477.

MacArthur, R.H. 1972. **Geographical ecology.** Princeton Univ. Press, Lawrenceville, NJ.

MacArthur, R.H. and E.O. Wilson. 1967. **The theory of island biogeography.** Princeton Univ. Press, Princeton, NJ.

MacDonald, D.D. (ed.). 1985. **Proc. of the Flathead River basin bull trout biology and population dynamics modelling information exchange.** Fisheries Branch, B.C. Min. of Env., Cranbrook, B.C.

Madansky, A. 1959. The fitting of straight lines when both variables are subject to error. J.Amer. Stat. Assoc. 54:173–205.

Maekawa, K. 1977. Studies on the variabliity of land-locked Miyabe char, *Salvelinus malma miyabei*. III. Geographical variation of the Dolly Varden, *Salvelinus malma*, and morphological characters of the Miyabe char. Jap. J. Ichthyol. 24:49–56. [In Japanese].

Maekawa, K. 1984. Life history patterns of the Miyabe charr in Shikaribetsu Lake, Japan, pp. 233–250. **In**: L. Johnson and B.L. Burns (eds.), Biology of the Arctic charr, proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Malde, H.E. 1965. The Snake River Plain, pp. 255–264. **In**: H.E. Wright, Jr. and D.G. Frey (eds.), The Quaternary of the United States. Princeton Univ. Press, Princeton, NJ.

Manaster, B. J. and S. Manaster. 1975. Techniques for estimating allometric equations. J. Morph. 147:299–308.

Manly, B.F. 1987. **Multivariate statistical methods: A primer.** Chapman and Hall, New York, NY.

Marcus, L.F. and J.H. Vandermeer. 1966. Regional trends in geographic variation. Syst. Zool. 15:1–13.

Mardia, K.V. 1970. Measures of multivariate skewness and kurtosis with applications. Biometrika 57:519–530.

Mardia, K.V. 1971. The effect of nonnormality on some multivariate tests and robustness to nonnormality in the linear model. Biometrika 58:105–121.

Mardia, K.V. 1975. Assessment of multinormality and the robustness of Hotelling's T^2 test. Appl. Stat. 24:163–171.

Marks, S. and O.J. Dunn. 1974. Discriminant functions when covariance matrices are unequal. J. Amer. Stat. Assoc. 69:555–559.

Marriott, F.H.C. 1974. **The interpretation of multiple observations.** Academic Press, New York, NY.

Mathews, W.H., J.G. Fyles and H.W. Nasmith. 1970. Postglacial crustal movements in southwestern British Columbia and adjacent Washington State. Can. J. Earth Sci. 7:690–702.

Matthews, W.J. 1985. Distribution of midwestern fishes on multivarite environmental gradients, with emphasis on *Notropis lutrensis*. Amer. Midl. Nat. 113:225–237.

Matthews, W.J. and H.W. Robison. 1988. The distribution of the fishes of Arkansas: A multivariate analysis. Copeia 1988:358–374.

Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup and L. Wolpert. 1985. Developmental contraints and evolution. Quart. Rev. Biol. 60:265–287.

Mayr, E. 1963. **Animal species and evolution.** Harvard University Press, Cambridge, MA.

Mayr, E. 1969. The biological meaning of species. Biol. J. Linn. Soc. 1:311–320.

McAllister, D.E. and C.C. Lindsey. 1959. Systematics of the freshwater sculpins (*Cottus*) of British Columbia. Nat. Mus. Can. Bull. 172. Contrib. Zool.

McCart, P.J. 1980. A review of the systematics and ecology of Arctic char, *Salvelinus alpinus*, in the western Arctic. Can. Tech. Rep. of Fish. Aquat. Sci. no. 935.

McCrossen, P.G. and R.P. Glaister (eds.) 1964. **Geological history of western Canada.** Alberta Soc. Petr. Geol., Calgary, Alta.

McGlade, J.M. and E.G. Boulding. 1986. The truss: A geometric and statistical approach to the analysis of form in fishes. Can. Tech. Rep. Fish. Aquat. Sci. 1457.

McKay, R.J. and N.A. Campbell. 1982a. Variable deletion techniques in discriminant analysis. II. Allocation. Br. J. Math. Stat. Psychiatr. 35:30–41.

McKay, R.J. and N.A. Campbell. 1982b. Variable deletion techniques in discriminant analysis. I. Description. Br. J. Math. Stat. Psychiatr. 35:1–29.

McKee, B. 1972. **Cascadia — The geologic evolution of the Pacific Northwest.** McGraw-Hill Inc., New York, NY.

McKitrick, M.C. and R.M. Zink. 1988. Species concepts in ornithology. The Condor 90:1–14.

McLachlan, G.J. 1976. A criterion for selecting variables for the linear discriminant function. Biometrics 32:529–534.

McLennan, D.A., D.R. Brooks and J.D. McPhail. 1988. The benefits of communication between phylogenetic systematics and comparative ethology: A case study using gasterosteid fishes. Can. J. Zool. 66:(in press).

McMahon, T.A. and J.T. Bonner. 1983. **On size and life.** Sci. Amer. Library, New York, NY.

McPhail, J.D. 1961. A systematic study of the *Salvelinus alpinus* complex in North America. J. Fish. Res. Bd. Can. 18:793–816.

McPhail, J.D. 1986. Book Review. Biology of the Arctic charr: Proceedings of the international symposium on Arctic charr, by L.Johnson and B. Burns (eds.). Copeia 1986:844–845.

McPhail, J.D. and C.C. Lindsey. 1970. **Freshwater fishes of north-western Canada and Alaska.** Bull. Fish. Res. Bd. Can. 173.

McPhail, J.D. and C.C. Lindsey. 1986. Zoogeography of the freshwater fishes of Cascadia (the Columbia system and rivers north to the Stikine), pp. 615–637. **In:** C.H. Hocutt and E.O. Wiley (eds.), Zoogeography of North American Freshwater Fishes. J.Wiley and Sons, Inc. New York, NY.

Mednikov, B.M., V.A. Maksimov and K.A. Savvaitova. 1980. Genetic divergence of Eurasian charrs, pp. 357–363. **In:** E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Medvedeva, K.D. and K.A. Savvaitova. 1980. Intrapopulation and geographic variability of the skull in charrs, pp. 435–440. **In:** E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Menozzi, P., A. Piazza and L. Cavalli-Sforza. 1978. Synthetic maps of human gene frequencies in Europeans. Science 201:786–792.

Meyer, A.W. 1935. Some historical aspects of the recapitulation idea. Quart. Rev. Biol. 10:379–396.

Middleton, G.V. 1962. A multivariate statistical technique applied to the study of sandstone composition. Trans. Roy. Soc. Can. Ser. III 56:119–126.

Miller, J. 1987. Host-plant relationships in the Papilionidae (Lepidoptera): Parallel cladogenesis of colonization? Cladistics 3:105–120.

Miller, R.G. 1974. The jackknife — A review. Biometrika 61:1–15.

Miller, R.R. 1965. Quaternary freshwater fishes of North America, pp. 569–581. **In:** H.E. Wright and D.G. Frey (eds.), The Quaternary of the United States. Princeton Univ. Press, Princeton, NJ.

Minckley, W.L., D.A. Hendrickson and C.E. bond. 1986. Geography of western North American freshwater fishes: Description and relationships to intracontinental tectonism, pp. 519–613. **In:** C.H. Hocutt and E.O. Wiley, Zoogeography of North American freshwater fishes. J. Wiley and Sons, New York, NY.

Misra, R.K. and E.C.R. Reeve. 1964. Genetic variation of relative growth in *Notonecta undulata*. Genet. Res. Camb. 5:384–396.

Mitter, C. and D.R. Brooks. 1983. Phylogenetic aspects of coevolution, pp. 65–98. **In:** D.J. Futuyma and M. Slatkin (eds.), Coevolution. Sinauer Press, New York, NY.

Monmonier, M.S. 1972. A spatially-controlled principal component analysis. Geogr. Anal. 2:192–195.

Moodie, G.G.E. 1972a. Morphology, life-history, and ecology of an unusual stickleback (*Gasterosteus aculeatus*) in the Queen Charlotte Islands, Canada. Can. J. Zool. 50:721–732.

Moodie, G.G.E. 1972b. Predation, natural selection, and adaptation in an unusual threespine stickleback. Heredity 28:155–167.

Moodie, G.E.E. and T.E. Reimchen. 1976. Glacial refugia, endemism and stickleback populations of the Queen Charlotte Islands, British Columbia. Can. Field-Nat. 90:471–474.

Moore, D.H. 1973. Evaluation of five discrimination procedures for binary variables. J. Amer. Stat. Assoc. 71:339–404.

Morishima, H. 1969. Phenetic similarity and phylogenetic relationships among strains of *Oryza perennis*, estimated by methods by methods of numerical taxonomy. Evolution 23:429–443.

Morrow, J.E. 1973. A new species of *Salvelinus* from the Brooks Range, northern Alaska, pp. 1–18. **In**: Studies on Alaskan fishes. Bid. Pap. Univ. Alaska 13:41.

Morrow, J.E. 1980a. Analysis of the Dolly Varden charr, *Salvelinus malma*, of northwestern North America and northeastern Siberia, pp. 323–338. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Morrow, J.E. 1980b. **The freshwater fishes of Alaska.** Alaska Northwest Publishing Co., Anchorage, AK.

Morton, N.E. and J.M. Lalouel. 1973. Topology of kinship in Micronesia. Amer. J. Hum. Gen. 25:422–432.

Morton, W.M. 1955. Charr or char — History of a common name for *Salvelinus*. Science 121:874–875.

Morton, W.M. 1970. On the validity of all subspecific descriptions of North American *Salvelinus malma* (Walbaum). Copeia 1970 (3):581–587.

Morton, W.M. 1980. Charr or char: A history of the English name for members of the salmonid genus *Salvelinus*, pp. 4–6. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Mosimann, J.E. 1970. Size allometry: Size and shape variables with characterizations of the log-normal and generalized gamma distributions. J.Amer. Stat. Assoc. 65:930–945.

Mosimann, J.E. and F.C. James. 1979. New statistical methods for allometry with application to Florida red-winged blackbirds. Evolution 33:444–459.

Mosimann, J.E. and J.D. Malley. 1979. Size and shape variables, pp. 175–189. **In**: L. Orlóci, C.R. Rao, and W.M. Stiteler (eds.), Multivariate methods in ecological work. Int. Co-op. Publ. House, Fairyland, MD.

Mosteller, F. and J.W. Tukey. 1977. **Data analysis and regression: A second course in statistics.** Addison-Wesley, Reading, MA.

MTS. 1976. **MTS, The Michigan Terminal System. Volume 1: Mts and the Computing Centre.** Univ. of Michigan Computing Centre, Ann Arbor, MI.

Muramoto J., J. Azumi and H. Fukuoka. 1974. Karotypes of 9 species of the Salmonidae. Chrom. Info. Serv. 17:20–23.

Nakamura, M. 1963. **Keys to the freshwater fishes of Japan fully illustrated in colours.** Hokuryukan Book Co., Tokyo. [In Japanese].

Neave, F. 1958. The origin and speciation of *Oncorhynchus*. Trans. Roy. SOc. Canada (V), 52:25–38.

Needham, P.R. and T.M. Vaughan. 1952. Spawning of the Dolly Varden, *Salvelinus malma*, in Twin Creek, Idaho. Copeia 1952:197–199.

Neff, N.A. and L.F. Marcus. 1980. **A survey of multivariate methods for systematics.** Privately published, New York, NY.

Neff, N.A. and G.R. Smith. 1979. Multivariate analysis of hybrid fishes. Syst. Zool. 28:176–196.

Nelson, C.D., D.M. Hopkins and D.W. Scholl. 1974. Cenozoic sedimentary and tectonic history of the Bering Sea, pp. 485–516. **In**: D.W. Hood and E.J. Kelley (eds.), Oceanography of the Bering Sea. Univ. Alaska, Inst. Mar. Sci., Fairbanks, Occ. Publ. 2.

Nelson, G.J. 1978a. From Candolle to Croizat: Comments on the history of biogeography. J. Hist. Biol. 11:293–329.

Nelson, G.J. 1978b. Ontogeny, phylogeny, paleontology, and the biogenetic law. Syst. Zool. 27:324–345.

Nelson, G.J. and N. Platnick. 1981. **Systematics and biogeography: cladistics and vicariance**. Columbia Univ. Press, New York, NY.

Nelson, G.J. and D.E. Rosen (eds.). 1981. **Vicariance biogeography: A critique**. Columbia Univ. Press, New York, NY.

Nelson, J.S. 1977. The postglacial invasion of fishes into Alberta. Alberta Nat. 7:129–135.

Newman, K.W. and R.C. Jancey. 1981. Sample size in studies of geographic variation. Can. J. Bot. 59:2158–2159.

Newman, K.W. and R.C. Jancey. 1983. Character selection and data structure in geographic variation in *Pinus contorta*. Silvae Genet. 32:137–141

Neyman, J. and E.L. Scott. 1951. On certain methods of estimating the linear structural relationship. Ann. Math. Stat. 22:352–361.

Nijhout, H.F., G.A. Wray, C.Kremen and C.K. Teragawa. 1986. Ontogeny, phylogeny and evolution of form: An algorithmic approach. Syst. Zool. 35:445–457.

Nilsson, N.-A. 1954. Studies on the feeding habits of trout and charr in north Swedish lakes. Rept. Inst. Freshw. Res. Drottningholm 36:163–225.

Nilsson, N.-A. 1960. Seasonal fluctuation in the food segregation of trout, charr and whitefish in 14 north-Swedish lakes. Rept. Inst. Freshw. Res. Drottningholm 41:185–205.

Nilsson, N.-A. 1963. Interaction between trout and charr in Scandinavia. Trans. Amer. Fish. Soc. 92:276–285.

Nilsson, N.-A. and O. Filipsson. 1971. Characteristics of two discrete populations of Arctic char (*Salvelinus alpinus* L.) in a north Swedish lake. Rep. Inst. Freshw. Res. Drottningholm 51:90–108.

Norcliffe, G.B. 1969. On the use and limitations of trend surface models. Can. Geogr. 13:338–348.

Norden, C.R. 1961. Comparative osteology of representative salmonid fishes, with particular refrence to the grayling (*Thymallus arcticus*) and its phylogeny. J. Fish. Res. Bd. Can. 18:679–791.

Nordeng, H. 1983. Solution to the "char problem" based on Arctic char (*Salvelinus alpinus*) in Norway. Can. J. Fish. Aquat. Sci. 40:1372–1387.

Nyman, L. 1967. Protein variations in Salmonidae. Rept. Inst. Freshw. Res. Drottningholm 47:5–38.

Nyman, L. 1972. A new approach to the taxonomy of the '*Salvelinus alpinus*' species complex. Rep. Inst. Freshw. Res. Drottningholm 52:103–131.

Nyman, L. 1984. Management of allopatric and sympatric populations of landlocked Arctic charr in Sweden, pp. 23–34. **In**: L. Johnson and B.L. Burns (eds.),Biology of the Arctic charr: Proceedings of an international symposium on Arctic charr. Univ. Manitoba Press, Winnipeg, Manitoba.

Nyman, L., J. Hammar and R. Gydemo. 1981. The systematics and biology of land-locked populations of Arctic char from northern Europe. Rep. Inst. Freshw. Res. Drottningholm 59:128–141.

Odeh, R.E. and M.Fox. 1975. **Sample size choice**. Marcel Dekker, Inc., New York, NY.

Odell, G., F.G. Oster, B. Burnside and P. Alberch. 1981. The mechanical basis of morphogenesis. Dev. Biol. 85:446–462.

Oden, N.L. and R.R. Sokal. 1986. Directional autocorrelation: An extension of spatial correlograms to two dimensions. Syst. Zool. 35:608–617.

O'Grady, R.T. 1985. The phylogenetics of parasitic flatworm life cycles. Cladistics 1:159–170.

Olson, E. and R. Miller. 1958. **Morphological integration**. Univ. Chicago Press, Chicago, IL.

Omelchenko, V.T. 1975. Application of protein electropherograms in *Salvelinus* taxonomy. Biol. Morya 4:76–79. [In Russian].

Oppenheimer, J. 1959. An embryological enigma in the **Origin of Species**, pp. 292–322. **In**: B.Glass, O Temkin and W.L. Strauss, Jr. (eds.), Forerunners of Darwin: 1745–1859. John Hopkins Univ. Press, Baltimore, MD.

Orloci, L. 1967. Data centering: A review and evaluation with reference to componenet analysis. Syst. Zool. 16:208–212.

Oshima, M. 1961. Studies on charrs found in Japanese waters. Japan Wildl. Bull. 18:3–70. [In Japanese].

Ouellette, R.P. and S.U. Qadri. 1968. The discriminatory power of taxonomic characteristics in separating salmonid fishes. Syst. Zool. 17:70–75.

Owen, J.G. and M.A. Chmielewski. 1985. On canonical variates analysis and the construction of confidence ellipses in systematic studies. Syst. Zool. 34:366–374.

Page, J.W. 1976. A note on interobserver error in multivariate analyses of populations. Amer. J. Phys. Anthr. 44:521–526.

Parenti, L.R. 1984. Biogeography of the Andean killifish genus *Orestias* with comments on the species flock conept, pp. 85–92. **In**: A.A. Echelle and I. Kornfield (eds.), Evolution of fish species flocks. Univ. Maine at Orono Press, Orono, ME.

Paterson, H.E.H. 1985. The recognition concept of species. Transuaal Mus. Monogr. 4:21–29.

Paterson, H.E.H. and M. McNamara. 1984. The recognition concept of species. S. Afr. J. Sci. 80:312–318.

Patterson, C. 1981. The development of the North American fish fauna: A problem of historical biogeography, pp. 265–281. **In**: P.O. Forey (ed.), The evolving biosphere: Chance, change and challenge. Cambridge Univ. Press, New York, NY.

Pauken, R.J. and D.E. Metter. 1971. Geographic representation of morphologic variation among populations of *Ascaphus truei* Stejneger. Syst. Zool. 20:434–441.

Pearce, S.C. and D.A. Holland. 1960. Some applications of multivariate analysis in botany. Appl. Stat. 9:1–7.

Pearson, K. 1897. On a form of spurious correlation which may arise when indices are used in the measurement of organs. Proc. Roy. Soc. London 60:489–502.

Peters, R.H. 1983. **The ecological implications of body size**. Cambridge Univ. Press, Cambridge, UK.

Phillips, B.F., N.A. Campbell and B.R. Wilson. 1973. A multivariate study of geographic variation in the whelk *Dicathais*. J. Exp. Mar. Biol. Ecol. 11:27–69.

Phillips, R.B. 1983. Shape characters in numerical taxonomy and problems with ratios. Taxon 32:535–544.

Pianka, E.R. 1970. On r and K selection. Amer. Nat. 104:592–597.

Piazza, A., P. Menozzi and L. Cavalli-Sforza. 1981a. The making and testing of geographic gene-frequency maps. Biometrics 37:635–659.

Piazza, A., P. Menozzi and L. Cavalli-Sforza. 1981b. Synthetic gene frequency maps of man and selective effects of climate. Proc. Natl. Acad. Sci. U.S.A. 78:2638–2642.

Pimentel, R.A. 1979. **Morphometrics, the multivariate analysis of biological data.** Kendall / Hunt Co., Dubuque, IA.

Pimentel, R.A. 1981. A comparative study of data and ordination techniques based on a hybrid swarm of sand verbenas (*Abronia* Juss.). Syst. Zool. 30:250–267.

Platnick, N.I. and G.J. Nelson. 1978. A method of analysis for historical biogeography. Syst. Zool. 27:1–16.

Platt, T. and W. Silvert. 1981. Ecology, physiology, allometry and dimensionality. J. Theor. Biol. 93:855–860.

Power, D.M. 1971. Statistical analysis of character correlations in Brewer's blackbirds. Syst. Zool. 20:186–203.

Prim, R.C. 1957. Shortest connection matrix network and some generalizations. Bell System Tech. J. 36:1389–1401.

Qadri, S.U. 1974. Taxonomic status of the *Salvelinus alpinus* complex. J. Fish. Res. Bd. Can. 31:1355–1361.

Quenouille, M.H. 1949. Approximate tests of correlation in time-series. J. Roy. Stat. Soc. 11B:68–84.

Raff, R.A. and T.C. Kaufman. 1983. **Embryos, genes and ancestors.** MacMillan, New York, NY.

Rao, C.R. 1952. bf Advanced statistical methods in biometric research. John Wiley and Sons, New York, NY.

Rao, C.R. 1964. The use and interpretation of principal component analysis in applied research. Sankhya 26:329–358.

Rao, C.R. 1966a. Covariance adjustment and related problems in multivariate analysis, pp. 87–103. **In:** P.R. Krishnaiah (ed.), Multivariate analysis. Academic Press, New York, NY.

Rao, C.R. 1966b. Discriminant function between composite hypotheses and related problems. Biometrika 53:339–345.

Read, D.W. and P.E. Lestrel. 1986. Comment on uses of homologous-point measures in systematics: A reply to Bookstein et al.. Syst. Zool. 35:241–253.

Reeve, E.C.R. and J.S. Huxley. 1945. Some problems in the study of allometric growth, pp. 121–156. **In:** W.E. le Gros Clark and P.B. Medawar (eds.), Essays on growth and form. Oxford Univ. Press, Oxford.

Reeves, B.O.K. 1973. The nature and age of the contact between the Laurentide and Cordilleran ice sheets in the western interior of North America. Arctic Alpine Res. 5:1–16.

Reist, J.D. 1985. An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. Can. J. Zool. 63:1429–1439.

Reist, J.D. 1986. An empirical evaluation of the coefficients used in residual and allometric adjustment of size covariation. Can. J. Zool. 64:1363–1368.

Reist, J.D. and E.J. Crossman. 1987. Genetic basis of variation in morphometric characters as implied by hybrids between subspecies of *Esox americanus* (Pisces: Esocidae). Can. J. Zool. 65:1224–1229.

Rensch, B. 1959. **Evolution above the species level.** Methuen and Co. Ltd., London.

Reyment, R.A. 1961. A note on geographical variation in European *Rana*. Growth 25:219–227.

Reyment, R.A. 1971. Multivariate normality in morphometric analysis. J. Int. Assoc. Math. Geol. 3:357–368.

Reyment, R.A. 1979. On the interpretation of the smallest principal component. Bull Geol. Inst. Univ. Uppsala 8:1–4.

Reyment, R.A. and C.F. Banfield. 1976. Growth-free canonical variates applied to fossil foraminifers. Bull. Geol. Inst. Univ. Uppsala, New Series 7:11–21.

Reyment, R.A., R.E. Blackith and N.A. Campbell. 1984. **Multivariate morphometrics.** 2nd ed. Academic Press, London, UK.

Rhoads, J.G. and E. Trinkaus. 1977. Morphometrics of the Neandertal talus. Amer. J. Phys. Anthr. 46:29–44.

Ricker, W.E. 1973. Linear regressions in fishery research. J. Fish. Res. Bd. Can. 30:409–434.

Ricklefs, R.E. 1979. Adaptation, constraint, and compromise in avain postnatal development. Biol. Rev. 54:269–290.

Ricklefs, R.E. 1987. Community diversity: Relative roles of local and regional processes. Science 235:167–171.

Ridley, M. 1983. **The explanation of organic diversity: The comparative method and adaptations for mating.** Clarendon, Oxford, UK.

Ripley, B.D. 1981. **Spatial statistics.** John Wiley and Sons, New York, NY.

Riska, B., W.R. Atchley and J.J. Rutledge. 1984. A genetic analysis of target growth in mice. Genetics 107:79–101.

Robins, C.R., R.M. Bailey, C.E. Bond, J.R. Brooker, E.A. Lachner, R.N. Lea and W.B. Scott. 1980. **A list of common and scientific names of fishes from the United States and Canada.** (4th ed.), Amer. Fish. Soc. Spec. Publ. no.12. Bethesda, MD.

Rohlf, F.J. 1967. Correlated characters in numerical taxonomy. Syst. Zool. 16:109–126.

Rohlf, F.J. and F.L. Bookstein. 1987. A comment on shearing as a method for "size correction". Syst. Zool. 36:356–367.

Röhrs, M. 1961. Allometrieforschung und biologische Formanalyse. Z. Morph. Anthrop. 51:289–321.

Rohwer, S.A. 1972. A multivariate assessment of interbreeding between the meadowlarks, *Sturnella*. Syst. Zool. 21:313–338.

Rohwer, S.A. and D.L. Kilgore. 1973. Interbreeding in the arid-land foxes, *Vulpes velox* and *V. macrotis*. Syst. Zool. 22:157–165.

Rosen, D.E. 1974. Phylogeny and zoogeography of salmoniform fishes. Bull. Amer. Mus. Nat. Hist. 153:265–326.

Ross, H.H. 1972. The origin of species diversity in ecological communities. Taxon 21:253–259.

Rounsefell, G.A. 1962. Relationships among North American Salmonidae. Fish. Bull. 209:235–270.

Rutter, N.W. 1980. Late Pleistocene history of the western Canadian ice-free corridor. Special AMQUA issue, the ice-free corridor and peopling of the new world. Can. J. Anthrop. 1:1–18.

Ryan, T.A.,Jr. and B.L. Joiner. 1971. Normal probability plots and tests for normality. Tech. Rep., Stat. Dept., Penn. State. Univ., PA.

Ryan, T.A., Jr. and B.L. Joiner. 1974. Normal probability plots and tests for normality. Tech. Rep., Stat. Dept., Penn. State Univ., PA.

Ryan, T.A., Jr., B.L. Joiner and B.F. Ryan. 1976. **Minitab. Student Handbook.** Duxbury Press, Boston, MA.

Sacher, G.A. 1970. Allometric and factorial analysis of brain structure in insectivores and primates, pp. 245–287. **In:** C.R. Noback and W. Montagna (eds.), The Primate Brain. Appleton-Century-Crofts, New York, NY.

SAS. 1982. **SAS user's guide: Statistics.** SAS Institute, Cary, NC.

SAS. 1985. **SAS / STAT guides for personal computers.** SAS Institute, Cary, NC.

Savvaitova, K.A. 1973. Ecology and systematics of freshwater chars of the genus *Salvelinus* (Nilsson) from some bodies of water in Kamchatka. J. Ichthyol. 13:58–68.

Savvaitova, K.A. 1980a. Taxonomy and biogeography of charrs in the Palearctic, pp. 281–294. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague The Netherlands.

Savvaitova, K.A. 1980b. Comments to the "Systematic review of the genus *Salvelinus*", pp. 480–481. **In**: E.K. Balon (ed.), Charrs: Salmonid fishes of the genus *Salvelinus*. Dr. W. Junk Publishers, The Hague, The Netherlands.

Savvaitova, K.A. 1983. The application of the biological species concept to an evaluation of chars of the genus *Salvelinus* (Salmonidae). J. Ichthyol. 23(6):1–12.

Scagel, R.K., Y.A. El-Kassaby and J. Emanuel. 1985. Assessing sample size and variable number in multivariate data, with specific reference to cone morphology variation in a population of *Picea sitchensis*. Can. J. Bot. 63:232–241.

Schaafsma, W. and G.N. van Vark. 1979. Classification and discrimination problems with applications, part IIa. Stat. Neerlandica 33:91–126.

Schindewolf, O.H. 1950. **Grundfragen der Paläontologie**. Schweizerbart, Stuttgart.

Schmidt-Nielsen, K. 1984. **Scaling: Why is animal size so important?** Cambridge Univ. Press, Cambridge, UK.

Schnell, G.D., P.G. Risser and J.F. Helsel. 1977. Factor analysis of tree distribution patterns in Oklahoma. Ecology 58:1345–1355.

Schoener, A. 1974. Experimental zoogeography, colonization and marine mini islands. Amer. Nat. 108:715–738.

Schueler, F.W. and J.D. Rising. 1976. Phenetic evidence of natural hybridization. Syst. Zool. 25:283–289.

Schuessler, K. 1974. Analysis of ratio variables: Opportunities and pitfalls. Amer. J. Sociol. 80:379–396.

Schutz, D.C. and T.G. Northcote. 1972. An experimental study of feeding behaviour and interaction of coastal cutthroat trout (*Salmo clarki clarki*) and Dolly Varden (*Salvelinus malma*). J. Fish. Res. Bd. Can. 29:555–565.

Scott, W.B. and E.J. Crossman. 1973. **Freshwater fishes of Canada**. Bull. Fish. Res. Bd. Can. 184.

Scudder, G.G.E. 1974. Species concept and speciation. Can. J. Zool. 52:1121–1134.

Seberg, O. 1986. A critique of the theory and methods of panbiogeography. Syst. Zool. 35:369–380.

Shaklee, J.B. and C.S. Tamaru. 1981. Biochemical and morphological evolution of Hawaiian bonefishes (*Albula*). Syst. Zool. 30:125–146.

Shapiro, S.S. and R.S. Francia. 1972. An approximate analysis of variance test for normality. J. Amer. Stat. Assoc. 67:215–216.

Shapiro, S.S. and M.B. Wilk. 1965. An analysis of variance test for normality (complete samples). Biometrika 52:591–611.

Shapiro, S.S., M.B. Wilk and H.J. Chen. 1968. A comparative study of various tests for normality. J. Amer. Stat. Assoc. 63:1343–1372.

Shea, B.T. 1983. Paedomorphosis and neoteny in the pigmy chimpanzee. Science 222:521–522.

Shea, B.T. 1985. Bivariate and multivariate growth allometry: Statistical and biological considerations. J. Zool. 206:367–390.

Shumway, W. 1932. The recapitulation theory. Quart. Rev. Biol. 7:93–99.

Siegel, A.F. and R.H. Benson. 1982. A robust comparison of biological shapes. Biometrics 38:341–350.

Sillén-Tullberg, B. 1988. Evolution of gregariousness in aposematic butterfly larvae: A phylogenetic analysis. Evolution 42:293–305.

Simberloff, D.S. 1974. Equilibrium theory of island biogeography. Ann. Rev. Ecol. Syst. 5:161–182.

Simberloff, D.S. 1986. Calculating probabilities that cladograms match: A method of biogeographical inference. Syst. Zool. 36:175–195.

Simberloff, D.S. and E.O. Wilson. 1969. Experimental zoogeography of islands: The colonization of empty islands. Ecology 50:278–296.

Simpson, G.G. and A. Roe. 1939. **Quantitative zoology**. McGraw-Hill, New York, NY.

Simpson, G.G., A. Roe and R.C. Lewontin. 1960. **Quantitative zoology**. 2nd. ed. Harcourt, New York, NY.

Sites, J.W., Jr. and C. Moritz. 1987. Chromosomal evolution and speciation revisited. Syst. Zool. 36:153–174.

Sjøvold, T. 1975. Some notes on the distribution and certain modifications of Mahalanobis' generalized distance (D^2). J. Human Evol. 4:549–558.

Skreslet, S.I. 1973. Group segregation in landlocked Arctic char *Salvelinus alpinus* (L.) of Jan Mayen Island in relation to the char problem. Astarte 6:55–58.

Smith, A.F.M. and D.J. Spiegelhalter. 1981. Bayesian approaches to multivariate structure, pp. 335–348. **In**: V. Barnett (ed.), Interpreting multivariate data. Wiley, Chichester, UK.

Smith, G.R. 1966. Distribution and evolution of the North American catastomid fishes of the subgenus *Pantosteus*, genus *Catastomus*. Misc. Publ. Mus. Zool. Univ. Michigan 129:1–132.

Smith, G.R. 1975. Fish of the Pliocene Glenns Ferry formation, southwest Idaho. Univ. Michigan Pap. Paleon. 14:1–68.

Smith, G.R. 1981. Late Cenozoic freshwater fishes of North America. Ann. Rev. Ecol. Syst. 12:163–193.

Smith, G.R. and D.R. Fisher. 1970. Factor analysis of distribution patterns of Kansas fishes, pp. 259–277. **In**: Pleistocene and recent environments of the central Great Plains. Univ. Kansas Dept. Geol. Spec. Publ. 3.

Smith, R.J. 1980. Rethinking allometry. J. Theor. Biol. 87:97–111.

Sneath, P.H.A. and K.G. McKenzie. 1973. Statistical methods for the study of biogeography, pp. 45–60. **In**: N.F. Hughes (ed.), Organisms and continents through time. Paleontological Assoc., London, UK.

Sneath, P.H.A. and R.R. Sokal. 1973. **Numerical taxonomy**. W.H. Freeman, San Francisco, CA.

Snell, O. 1891. Das Gewicht des Gehirnes und den Hirnmantels der Säugetiere in Beziehung zu deren geistigen Fähigkeiten. Sitzungsberichte Ges. Morph. Physiol. München 7:90–94.

Sober, E. 1983. Parsimony in systematics: Philosophical issues. Ann. Rev. Ecol. Syst. 14:335–357.

Soin, S.C. 1980. Types of development of salmoniform fishes and their taxonomic importance. J. Ichthyol. 20(1):49–56.

Sokal, R.R. 1965. Statistical methods in systematics. Biol. Rev. 40:337–391.

Sokal, R.R. 1974. The species problem reconsidered. Syst. Zool. 22:360–374.

Sokal, R.R. 1979. Testing statistical significance of geographic variation patterns. Syst. Zool. 28:227–232.

Sokal, R.R. and T.J. Crovello. 1970. The biological species concept: A critical evaluation. Amer. Nat. 104:127–153.

Sokal, R.R. and P. Menozzi. 1982. Spatial autocorrelations of HLA frequencies in Europe support demic diffusion of early farmers. Amer. Nat. 119:1–17.

Sokal, R.R. and R.C. Rinkel. 1963. Geographic variation of alate *Pemphigus populitransversus* in eastern North America. Kansas Univ. Sci. Bull. 44:467–507.

Sokal, R.R. and F.J. Rohlf. 1969. **Biometry**. W.H. Freeman and Co., San Francisco, CA.

Somers, K.M. 1986. Multivariate allometry and removal of size with principal components analysis. Syst. Zool. 35:359–368.

Sparholt, H. 1985. The population, survival, growth, reproduction and food of Arctic charr, *Salvelinus alpinus* (L.), in four unexploited lakes in Greenland. J. Fish. Biol. 26:313–330.

Srivastava, M.S. and E.M. Carter. 1983. **An introduction to applied multivariate statistics**. North-Holland Publishing Co., New York, NY.

Stanley, S.M. 1979. **Macroevolution: Patterns and processes**. Freeman and Co., San Francisco, CA.

Stearns, S.C. 1976. Life-history tactics: A review of the data. Quart. Rev. Biol. 51:3–47.

Stearns, S.C. 1977. The evolution of life-history traits: A critique of the theory and a review of the data. Ann. Rev. Ecol. Syst. 8:145–171.

Stearns, S.C. 1983. The influence of size and phylogeny on patterns of covariation among life-history traits in the mammals. Oikos 41:173–187.

Stevenson, M.M., G.D. Schnell and R. Black. 1974. Factor analysis of fish distribution patterns in western and central Oklahoma. Syst. Zool. 23:202–218.

Stewart, D. and W. Love. 1968. A general canonical relation index. Psych. Bull. 70:160–163.

Straight, L. 1982. **B.C. Freshwater Fishing Guide**. D.W. Friesen and Sons Ltd., Cloverdale, B.C..

Strauss, R.E. 1987. On allometry and relative growth in evolutionary studies. Syst. Zool. 36:72–75.

Strauss, R.E. and F.L. Bookstein. 1982. The truss: Body form reconstructions in morphometrics. Syst. Zool. 31:113–135.

Strauss, R.E. and L.A. Fuiman. 1985. Quantitative aspects of body form and allometry in larval and adult Pacific sculpins (Teleostei: Cottidae). Can. J. Zool. 63:1582–1589.

Suckley, G. 1858. Descriptions of sveral new species of Salmonidae, from the north-west coast of America. Ann. Lyc. Nat. Hist. 7:1–10.

Suckley, G. 1860. Report upon Salmonidae, chapter 1, **In**: Pacific Railroad reports, fishes, 12(5):307–349.

Suckley, G. 1861. Notices of certain new species of North American Salmonidae, chiefly in the collection of the N.W. Boundary Commission in charge: Archibald Campbell, Esq., Commissioner of the United States, collected by Doctor C.B.R. Kennerly, Naturalist of the Commission. Ann. Lyc. Nat. Hist. 7:306–313.

Sweet, S.S. 1980. Allometric inference in morphology. Amer. Zool. 20:643–652.

Szij, L.J. 1962. Morphological analysis of the sympatric populations of meadowlarks in Ontario. Proc. Int. Ornith. Congr. XIII:176–188.

Tabachnik, B.G. and L.S. Fidell. 1983. **Using multivariate statistics**. Harper and Row Publs., New York, NY.

Takai, S. 1977. Principal component analysis of the elongation of metacarpal and phalangeal bones. Amer. J. Phys. Anthrop. 47:301–304.

Tanner, J.M. 1963. Regulation of growth in size in mammals. Nature 199:845–850.

Teissier, G. 1948. La relation d'allométrie: sa signification statistique et biologique. Biometrics 4:14–48.

Teissier, G. 1960. Relative growth, pp. 537–560. **In**: T.H. Waterman (ed.), The physiology of the crustacea. I. Metabolism and growth. Academic Press, New York, NY.

Templeman-Kluit, D. 1980. Evolution of physiography and drainage in southern Yukon. Can. J. Earth Sci. 17:1189–1203.

Thomas, P.A. 1968. Variation and covariation in characters of the rabbittick, *Haemaphysalis leporispalustris*. Kansas Univ Sci. Bull. 47:829–862.

Thompson, D'A.W. 1942. **On growth and form.** 2nd ed. MacMillan, New York, NY.

Thomson, K.S. 1988. Ontogeny and phylogeny recapitulated. Amer. Sci. 76:273–275.

Thorington, R.W., Jr. 1972. Proportions and allometry in the gray squirrel, *Sciurus carolinensis*. Nemouria 8:1–17.

Thorpe, R.S. 1975a. Quantitative handling of characters useful in snake systematics with particular reference to intraspecific variation in the ringed snake *Natrix natrix* (L.). Biol. J. Linn. Soc. 7:27–43.

Thorpe, R.S. 1975b. Biometric analysis of incipient speciation in the ringed snake *Natrix natrix* (L.). Experientia 31:180–182.

Thorpe, R.S. 1976. Biometric analysis of geographic variation and racial affinities. Biol. Rev. 51:407–452.

Thorpe, R.S. 1980. A comparative study of ordination techniques in numerical taxonomy in relation to racial variation in the ringed snake *Natrix natrix* (L.). Biol. J. Linn. Soc. 13:7–40.

Thorpe, R.S. 1983a. A review of the numerical methods for recognising and analyzing racial differentiation, pp. 404–423. **In**: J. Felsenstein (ed.), Numerical Taxonomy. NATO ASII series, series G, Ecological Sciences, No. 1., Springer-Verlag, Berlin.

Thorpe, R.S. 1983b. A biometric study of the effects of growth on the analysis of geographical variation: Tooth number in green geckos (Reptilia: *Phelsuma*). J. Zool. 201:13–26.

Thorpe, R.S. 1985a. Character number and the multivariate analysis of simple patterns of geographic variation: Categorical or "stepped clinal" variation. Syst. Zool. 34:127–139.

Thorpe, R.S. 1985b. Clines: Character number and the multivariate analysis of simple patterns of geographic variation. Biol. J. Linn. Soc. 26:201–214.

Thorpe, R.S. 1985c The effect of insignificant characters on the multivariate analysis of simple patterns of geographic variation. Biol. J. Linn. Soc. 26:215–223.

Thorpe, R.S. and L. Leamy. 1983. Morphometric studies in inbred and hybrid house mice (*Mus* sp.): Multivariate analysis of size and shape. J. Zool. 199:421–432.

Timm, R.M. and R.D. Price. 1980. The taxonomy of *Geomydoecus* (Mallophaga: Trichodectidae) from the *Geomys bursarius* complex (Rodentia: Geomydidae). J. Med. Entom. 17:126–145.

Tipper, H.W. 1971. Glacial geomorphology and Pleistocene history of central British Columbia. Bull. Geol. Surv. Can. 196:1–89.

Tsuyuki, H., J.F. Uthe, E. Roberts and L.W. Clarke. 1966. Comparative electropherograms of *Coregonus clupeaformis*, *Salvelinus namaycush*, *S. alpinus*, *S. malma*, and *S. fontinalis* from the family Salmonidae. J. Fish. Res. Bd. Can. 23:1599–1606.

Tukey, J.W. 1962. The future of data analysis. Ann. Math. Stat. 33:1–67.

Ueda, T. and Y. Ojima. 1984. Karyological characteristics of the brown trout, the Japanese char and their hybrids. Proc. Jap. Acad. B. Phys. Biol. Sci. 60(7):249–252.

Utter, F.M. 1981. Biological criteria for definition of species and distinct intraspecific populations of anadromous salmonids under the U.S. Endangered Species Act of 1973. Can. J. Fish. Aquat. Sci. 38:1626–1635.

Uyeno, T. and R.R. Miller. 1963. Summary of late Cenozoic freshwater fish records from North America. Occ. Pap. Mus. Zool. Univ. Michigan 631:1–34.

van den Wollenberg, A.L. 1977. Redundancy analysis: An alternative for canonical correlation analysis. Psychometrika 42:207–219.

Van Valen, L. 1978. The statistics of variation. Evol. Theory 4:33–43.

Vasilyev, V.P. 1975. Karyotypes of some forms of Arctic char from Kamchatka. J. Ichthyol. 15:374–386.

Veitch, L.G. 1965. The description of Australian pressure fileds by principal components. Quart. J. Royal Meteorol. Soc. 91:184–195.

Viktorovsky, R.M. 1975a. Karyotypes of the endemic chars from Kronotski Lake. Tsitologiya 17:464–466. [In Russian].

Viktorovsky, R.M. 1975b. Karyotypes of *Salvelinus leucomaenis* (Pallas) and *S.malma* (Walbaum) (Pisces, Salmoniformes, Salmonidae). Zool. Zh. 54:787–789. [In Russian].

Viktorovsky, R.M. 1978. **Mechanism of speciation in chars from Lake Kronotskoe.** Moskva, Izd. Nauk. [In Russian].

Viktorovsky, R.M. and M.K. Glubokovsky. 1977. Mechanisms and rate of speciation in the charr genus *Salvelinus* (Salmonidae, Pisces). Dokl. Akad. Nauk. SSR 235:946-949. [In Russian].

Vladykov, V.D. 1954. Taxonomic characters of the eastern America charrs (*Salvelinus* and *Cristivomer*). J. Fish. Res. Bd. Can. 11:904–932.

Vladykov, V.D. 1964. A review of salmonid genera and their broad geographical distribution. Trans. Roy. Soc. Can. (4)1, sec.3:459-504.

von Baer, K.E. 1828. **Entwicklungsgeschichte der Thiere: Beobachtung und Reflexion.** Bornträger, Königsberg.

von Bertalanffy, L. 1960. Principles and theory of growth, pp. 137–260. **In:** W.W. Nowinski (ed.), Fundamental aspects of normal and malignant growth. Elsevier, Publ. Co., Berlin.

Waddington, C.H. 1957. **The strategy of the genes.** Allen and Unwin, London.

Waddington, C.H. 1962. **New patterns in genetics and development.** Columbia Univ. Press, New York, NY.

Wagel, B. 1968. Multivariate beta distribution and a test for multivariate normality. J. Roy. Stat. Soc. 30B:511–516.

Wake, D.B. 1966. Comparative osteology and evolution of the lungless salamanders, family Plethodontidae. Mem. So. Calif. Acad. Sci. 4:1–111.

Wanntorp, H.-E. 1983. Historical constraints in adaptation theory: Traits and non-traits. Oikos 41:157–160.

Warner, B.C., R.W. Mathewes and J.J. Clague. 1982. Ice-free conditions on the Queen Charlotte Islands, British Columbia, at the height of Late Wisconsin glaciation. Science 218:675–677.

Wartenberg, D.E. 1985a. Canonical trend surface analysis: A method for describing geographic patterns. Syst. Zool. 34:259–279.

Wartenberg, D.E. 1985b. Multivariate spatial autocorrelation: A method for exploratory geographical analysis. Geogr. Anal. 17:263–283.

Wayne, R.K. 1986. Cranial morphology of domestic and wild canids: The influence of development on morphological change. Evolution 40:243–261.

Weiner, J.M. and O.J. Dunn. 1966. Elimination of variates in linear discriminant problems. Biometrics 22:268–275.

Weismann, A. 1904. **The evolution theory: the biogenetic law.** Volume 2. Edward Arnold, London.

Wheeler, H.E. and E.F. Cook. 1954. Structural and stratigraphic significance of the Snake River capture, Idaho–Oregon. J. Geol. 62:525–536.

White, J.F. and S.J. Gould. 1965. Interpretation of the coefficient in the allometric equation. Amer. Nat. 99:5–18.

Wiley, E.O. 1978. The evolutionary species concept reconsidered. Syst. Zool. 27:17-26.

Wiley, E.O. 1981. **Phylogenetics: The theory and practise of phylogenetic systematics.** John Wiley and Sons, New York, NY.

Wiley, E.O. and R.L. Mayden. 1985. Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. Ann. Missouri Bot. Garden. 72:596–635.

Wilk, M.B. and R. Gnanadisikan. 1968. Probability plotting methods for the analysis of data. Biometrika 55:1–17.

Wilk, M.B., R. Gnanadesikan and M.J. Huyett. 1962. Probability plots for the Gamma distribution. Technometrics 4:1–20.

Wilk, M.B. and S.S. Shapiro. 1968. The joint assessment of normality of several independent samples. Technometrics 10:825–839.

Wilk, S.J., W.G. Smith, D.E. Ralph and J. Sibunka. 1980. Population structure of summer flounder between New York and Florida based on linear discriminant analysis. Trans. Amer. Fish. Soc. 109:205–271.

Wilkie, J.S. 1967. Preformation and epigenesis: A new historical treatment. Hist. Sci. 6:138–150.

Willig, M.R., R.D. Owen and R.L. Colbert. 1986. Assessment of morphometric variation in natural populations: The inadequacy of the univariate approach. Syst. Zool. 35:195–203.

Willig, M.R. and R.D. Owen. 1987. Univariate analyses of morphometric variation do not emulate the results of multivariate analyses. Syst. Zool. 36:398–400.

Wilson, J.W. III. 1974. Analytical zoogeography of North American mammals. Evolution 28:124–140.

Wilson, M.V.H. 1977. Middle Eocene freshwater fishes from British Columbia. Royal Ontario Mus. Life Sci. Contrib. 113.

Winans, G.A. 1984. Multivariate morphometric variability in Pacific salmon: Technical demonstration. Can. J. Fish. Aquat. Sci. 41:1150–1159.

Winans, G.A. and R.S. Nishioka. 1987. A multivariate description of change in body shape of coho salmon (*Oncorhynchus kisutch*) during smoltification. Aquaculture 66:235–245.

Workman, W.B. 1978. Prehistory of the Aishihik-Kluane area, southwest Yukon Territory. Nat. Mus. Man. Arch. Surv. Can. Pap. 74.

Yang, S.H. and R.K. Selander. 1968. Hybridization in the grackle *Quiscalus quiscalus* in Louisiana. Syst. Zool. 17:107–143.

Yoshiyasu, K. 1973. Starch-gel electrophoresis of hemoglobins of freshwater salmonid fishes in northeast Japan. Bull. Jap. Soc. Sci. Fish 39:449-459.

Zakharova, L.A., G.G. Novikov and K.A. Savvaitova. 1971. Relationships between species of the genus *Salvelinus* based on precipitation and immunoelectrophoresis in agar gel. Zool. Zh. 50:537-546. [In Russian].

Zanardi, P., G. Dell'Acqua, C. Menini and I. Barrai. 1977. Population genetics in the province of Ferara. I. Genetic distances and geographic distances. Amer. J. Hum. Gen. 29:169–177.

Zar, J.H. 1968. Calculation and miscalculation of the allometric equation as a model in biological data. Bioscience 18:1118–1120.

Zar, J.H. 1984. **Biostatistical analysis.** 2nd ed. Prentice-Hall, Englewood Cliffs, NJ.

Appendix A. Morphology and Meristics Used.

Morphology For Part II

Body Morphology

1. body depth: as in Hubbs and Lagler 1958.
2. body width: width at point body depth is measured.
3. peduncle length: as in Hubbs and Lagler 1958.
4. peduncle width: width at point peduncle depth is measured.
5. snout length: as in Hubbs and Lagler 1958.
6. predorsal length: as in Hubbs and Lagler 1958.
7. dorsal fin-adipose fin length: insertion of dorsal fin to origin of adipose fin.
8. adipose fin-caudal fin length: insertion of adipose fin to base of most dorsal caudal fin ray.
9. pectoral fin-pelvic fin length: insertion of pectoral fin to origin of pelvic fin.
10. pelvic fin-anal fin length: insertion of pelvic fin to origin of anal fin.
11. anterior gular plate length: distance across gular plate at its anterior point.
12. lateral line-dorsal depth: lateral line to dorsal body surface at origin of dorsal fin.
13. lateral line-ventral depth: lateral line to ventral body surface below origin of dorsal fin.

Fin Morphology

14. adipose fin base length: origin to insertion of adipose fin.
15. adipose fin length: base at centre of adipose fin to its furthest tip.
16. adipose fin depth: distance from ventral to dorsal surface of adipose fin at its centre.
17. dorsal fin base: as in Hubbs and Lagler 1958.
18. posterior dorsal fin height: distance from insertion of dorsal fin to furthest tip of posterior dorsal fin ray.

19. anal fin base: as in Hubbs and Lagler 1958.

20. body-caudal fin fork length: posterior part of hypural plate to anterior part of caudal fin fork.

21. inner pectoral fin length: insertion of pectoral fin to posterior tip of innermost pectoral fin ray.

22. pectoral fin base: origin to insertion of pectoral fin.

23. pelvic fin base: origin to insertion of pelvic fin.

24. pelvic axillary process length: anterior to posterior point of pelvic axillary process.

Head Morphology

25. preopercle length: posterior edge of orbit to posterior edge of preopercle.

26. opercle length: posterior edge of preopercle to posterior edge of opercle.

27. nostril size: longest distance across nostril from one skin edge to another.

28. orbit depth: distance from ventral to dorsal edge of orbit.

29. orbit skin flap: anterior edge of orbit to posterior edge of orbit skin flap at its centre.

30. eye length: anterior edge of eyeball to posterior edge of eyeball.

31. pupil length: anterior edge of eye pupil to posterior edge of eye pupil.

32. orbit-dorsal depth: distance from centre of orbit to dorsal body surface directly above it.

33. orbit-ventral depth: distance from centre of orbit to ventral body surface directly below it.

34. orbit-maxillary depth: distance from centre of orbit to dorsal edge of maxillary directly below it.

35. orbit-nostril length: centre of orbit to posterior edge of nostril.

36. snout width: inside edge of left nostril to inside edge of right nostril.

37. head width: dorsal edge at centre of left orbit to dorsal edge at centre of right orbit.

38. **premaxillary length**: anterior tip of premaxillary to anterior tip of maxillary.

39. anterior mandible depth: distance from ventral to dorsal edges of mandible at its most anterior point.

40. anterior mandible distance: ventral inside edge of left mandible to ventral inside edge of right mandible at their most anterior point.

Truss Morphology

41. measurement 1–2 (see fig. 1).

42. measurement 1–3 (see fig. 1).

43. measurement 1–4 (see fig. 1).

44. measurement 2–4 (see fig. 1).

45. measurement 4–6 (see fig. 1).

46. measurement 6–8 (see fig. 1).

47. measurement 7–9 (see fig. 1).

48. measurement 9–12 (see fig. 1).

49. measurement 10–11 (see fig. 1).

50. measurement 11–12 (see fig. 1).

51. **standard length**: as in Hubbs and Lagler 1958.

Meristics For Part II

1. dorsal fin rays: as in Hubbs and Lagler 1958.

2. **anal fin rays**: as in Hubbs and Lagler 1958.

3. caudal fin rays: count of all caudal fin rays.

4. pectoral fin rays: as in Hubbs and Lagler 1958.

5. pelvic fin rays: as in Hubbs and Lagler 1958.

6. **branchiostegal rays**: total branchiostegal fin ray count (Hubbs and Lagler 1958). Counts were made separately on the right and left halves.

7. mandibular pores: as in Hubbs and Lagler 1958. They were exposed by drying with a towel and then marked with a felt pen run across the mandible. The pore at the end of each mandible was not counted (Cavender 1978).
8. lateral line: as in Hubbs and Lagler 1958.
9. gill-rakers: all rakers, including rudimentary ones, were counted on the removed first right gill-raker arch. Counts were made separately on the upper and lower arches.
10. pyloric caeca: were counted by actually removing them.

Other Morphology and Meristics

Body Morphology

52. total length: as in Hubbs and Lagler 1958.
53. fork length: as in Hubbs and Lagler 1958.
54. peduncle depth: as in Hubbs and Lagler 1958.
55. head length: as in Hubbs and Lagler 1958.
56. dorsal fin-caudal fin length: insertion of dorsal fin to base of most dorsal caudal fin ray.
57. pre-pectoral length: anterior snout edge to origin of pectoral fin.
58. anal vent length: anterior to posterior edge of anal vent.
59. anal vent width: largest distance from left to right inside edges of anal vent.
60. gular branch width: left anterior edge of gular branch to right anterior edge of gular branch.
61. branchiostegal length: anterior edge to posterior edge of branchiostegal region.
62. anterior branchiostegal width: left anterior edge to right anterior edge of visible branchiostegal region.
63. posterior branchiostegal width: left posterior edge to right posterior edge of visible branchiostegal region.
64. branchiostegal distance: left outside edge to right outside edge of branchiostegal region at point where edges meet pectoral fins.

65. gape length: anterior snout tip to posterior mouth edge.

66. gape width: left side to right side of mouth at its most posterior point.

67. largest spot size: length of qualitatively largest spot.

68. smallest spot size: length of qualitatively smallest spot.

69. maximum parr mark height: distance from dorsal to ventral edge of largest parr mark if present.

70. minimum parr mark height: distance from dorsal to ventral edge of smallest parr mark if present.

71. maximum parr mark width: distance from anterior to posterior edge of widest parr mark if present.

72. minimum parr mark width: distance from anterior to posterior edge of narrowest parr mark if present.

73. parr mark distance: maximum distance between the edges of the widest parr mark and a neighbouring parr mark if present taken at the lateral line.

Fin Morphology

74. adipose fin width: width of adipose fin at its centre.

75. anterior dorsal fin height: distance from origin of dorsal fin to furthest tip of anterior dorsal fin ray.

75. dorsal fin width: width of dorsal fin at the centre of its most anterior fin ray.

76. anterior anal fin height: distance from origin of anal fin to furthest tip of anterior anal fin ray.

77. posterior anal fin height: distance from insertion of anal fin to furthest tip of posterior anal fin ray.

78. anal fin width: width of anal fin at the centre of its most anterior fin ray.

79. caudal fin width: width of most dorsal caudal fin ray at its centre.

80. outer pectoral fin length: origin of pectoral fin to posterior tip of outermost pectoral fin ray.

81. pectoral fin width: width of most anterior pectoral fin ray at its centre.

82. inner pelvic fin length: insertion of pelvic fin to posterior tip of innermost pelvic fin ray.

83. outer pelvic fin length: origin of pelvic fin to posterior tip of outermost pelvic fin ray.

84. pelvic fin width: width of most anterior pelvic fin ray at its centre.

85. pelvic axillary process height: ventral to dorsal surface of pelvic axillary process at its centre.

Head Morphology

86. preopercle height: ventral to dorsal edge of preopercle.

87. opercle height: ventral to dorsal edge of opercle.

88. opercle width: width of opercle at its most posterior edge.

89. orbit length: anterior to posterior edge edge of orbit including anterior skin flap.

90. eye depth: distance from dorsal to ventral edge of eyeball.

91. eye-opercle length: posterior edge of orbit to posterior edge of opercle.

92. snout depth: dorsal to ventral surface of snout at nostrils.

93. head depth: dorsal to ventral surface of head at occiput.

94. **maxillary length**: anterior to posterior edge of maxillary.

95. anterior maxillary depth: dorsal to ventral edge of maxillary at its most anterior point.

96. posterior maxillary depth: dorsal to ventral edge of maxillary at its largest posterior point.

97. maxillary width: width at centre of most posterior edge of maxillary.

98. mandible length: anterior edge to posterior edge of mandible.

99. posterior mandible depth: dorsal to ventral edge of mandible at its most posterior point.

100. anterior mandible width: inside to outside edge of mandible at its most anterior point.

101. posterior mandible width: inside to outside edge of mandible at its most posterior point.

102. posterior mandible distance: ventral inside edge of left mandible to ventral inside edge of right mandible at their most posterior point.

Truss Morphology

103. measurement 2–3 (see fig. 1).

104. measurement 3–4 (see fig. 1).

105. measurement 3–5 (see fig. 1).

106. measurement 3–6 (see fig. 1).

107. measurement 4–5 (see fig. 1).

108. measurement 5–6 (see fig. 1).

109. measurement 5–7 (see fig. 1).

110. measurement 5–8 (see fig. 1).

111. measurement 6–7 (see fig. 1).

112. measurement 7–8 (see fig. 1).

113. measurement 7–10 (see fig. 1).

114. measurement 8–9 (see fig. 1).

115. measurement 8–10 (see fig. 1).

116. measurement 9–10 (see fig. 1).

117. measurement 9–11 (see fig. 1).

118. measurement 10–12 (see fig. 1).

Meristics

11. spot number: number of spots above lateral line on one side.

12. spot number below lateral line: number of spots below lateral line on one side.

13. spot number above lateral line: number of spots above lateral line on one side.

14. parr marks: number of parr marks on one body side if present.

15. scales above lateral line: as in Hubbs and Lagler 1958.

16. scales below lateral line: as in Hubbs and Lagler 1958.

17. presence/absence of basibranchial teeth: as in Cavender 1978 and McPhail 1961.

18. presence/absence of mandibular symphysis: as in Cavender 1978.

19. presence/absence of vermiculations

20. sex: male or female determined by gonad inspection.