National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Clinal variation in morphology in eastern populations of the stream-dwelling waterstrider, *Aquarius remigis* (Hemiptera: Gerridae)

Julie M. Brennan

A Thesis

in

The Department

of

Biology

Presented in Partial Fulfilment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

July, 1993.

© Julie M. Brennan, 1993.

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-87253-5

Canada

ABSTRACT

Clinal variation in morphology in eastern populations of the stream-dwelling waterstrider, *Aquarius remigis* (Hemiptera: Gerridae).

Julie M. Brennan

I examined morphological variation among populations of the stream-dwelling waterstrider, *Aquarius remigis*. A previous analysis of allozyme variation along a transect from southern Quebec to southern North Carolina revealed two distinct 'types' of *A. remigis*, north and south of a transition zone in southern New York and northern Pennsylvania. To assess the concordance of morphological and allozymic differentiation, adults were sampled from 15 populations along the original transect, as well as 11 additional populations in New York, Connecticut and New Jersey. Eight measurements of leg and body size were taken of preserved individuals. Univariate and multivariate analyses (multiple regression, PCA, discriminant analysis) reveal significant differentiation between populations from Quebec and New York and populations from Pennsylvania, Connecticut, New Jersey, Virginia and North Carolina. Strong clines in body size and shape within each region and a sharp transition between regions supports previous suggestions that there are two distinct 'types' of this species in eastern North America. Increased within-population variance in the transitional populations supports the earlier conclusion of a restriction of gene flow within the zone. The transition zone in morphology coincides with that revealed by the allozyme variation, but extends further south within Pennsylvania. A greater proportion of northern 'type' females than males in the transitional populations suggests more introgression of northern alleles in females

(consistent with Haldane's rule). The covariance of allozyme frequency and morphometric traits suggests that differentiation has occurred through vicariance and secondary contact, and that in this system, cladistic and adaptive differentiation are coupled.

ACKNOWLEDGEMENTS

Firstly, and most importantly, my thanks goes to Dr. Daphne Fairbairn, whose support, patience, and wisdom have helped make this thesis possible. Dr. D. A. Roff graciously provided the crucial digitizing equipment used for all measurements in the thesis. I greatly appreciated the assistance in sample collection from R. Preziosi, Dr. D. J. Fairbairn, Dr. D. A. Roff, Dr. W. U. Blanckenhorn, V. Simoneau, and I. Weigensberg. I would also like to thank R. Preziosi for his helpful advice with statistics. The comments of committee members Dr. P. Albert, Dr. J. Grant and Dr. J. McNeil have helped improve this thesis. I would also especially like to thank my family for their moral support and particularly Sean Basquill for his caring and encouragement.

Thanks also to the National Park Service, Blue Ridge Parkway who provided permits for sampling.

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES viii

INTRODUCTION 1

METHODS 6

RESULTS 17

Univariate analysis 17

Multivariate analysis 23

A posteriori analysis 39

DISCUSSION 50

REFERENCES 58

APPENDIX A: multivariate analyses 65

APPENDIX B: population means and standard deviations 69

LIST OF FIGURES

Figure 1. Map of eastern Canada and USA indicating the location of sites from which *A. remigis* were collected 7

Figure 2. Photograph of a ventral view of an apterous female *A. remigis* with the corresponding male genital segments, indicating where the eight different morphometric measurements were taken 10

Figure 3. Map indicating the location of *a posteriori* sampling sites from which *A. remigis* were collected 13

Figure 4. Morphological measurements (trait means for each site) of *A. remigis* as a function of latitude 18

Figure 5. Mean population scores on PC1 of a general PCA based on seven morphological measurements of *A. remigis*, plotted against latitude ... 25

Figure 6. Mean regional scores on PC1 of a general PCA based on seven morphometric measurements of *A. remigis*..... 27

Figure 7. Mean population scores on both PC1 and PC2 of a size-adjusted PCA based on seven morphological measurements of *A. remigis*, plotted against latitude 31

Figure 8. Mean regional scores on PC1 and PC2 of a size-adjusted PCA based on seven morphological measurements of *A. remigis* 33

Figure 9. Within-population variance of log-transformed data for each of the eight morphological measurements of *A. remigis*, plotted against latitude ... 40

Figure 10. Within-population variance for the scores on PC1 of a general and a size-adjusted PCA based on seven morphological measurements of *A. remigis*, plotted against latitude 42

Figure 11. Mean population scores on PC1 of a general and a size-adjusted PCA based on seven morphological measurements of *A. remigis*, plotted against latitude, for both the original transect and the *a posteriori* samples 46

LIST OF TABLES

Table 1. Results of a stepwise multiple regression for *A. remigis* males 20

Table 2. Results of a stepwise multiple regression for *A. remigis* females 21

Table 3. Principal components analysis component loadings for *A. remigis* 29

Table 4. Classification coefficients and constants of a discriminant analysis based on seven morphological measurements of *A. remigis* 36

Table 5. Discriminant analysis canonical loadings for *A. remigis* based on seven morphological measurements 37

Table 6. Results of a 1-tailed Mann-Whitney U test comparing the within-site variances in the transition zone with those outside the zone for the eight morphological measurements of *A. remigis* 44

Table 7. Discriminant analysis results indicating the proportion of *A. remigis* individuals characterized as northern 'types' within each of the *a posteriori* sampling sites 48

INTRODUCTION.

Geographic gradients in trait values, clines, have long been of interest to evolutionary biologists, particularly for the opportunity they provide to study evolution (Sumner, 1929; Haldane, 1948) and their possible importance in speciation (Endler, 1977). Clinal variation is a widespread phenomenon found in almost every group of organisms (Endler, 1977). It has been examined on every continent in many taxa including toads and frogs (e.g. Szymura and Farana, 1978; Gartside et al., 1979, Gollmann, 1984), birds (e.g. Selander, 1964; Yang and Selander, 1968), mammals (e.g. Smith, 1979, Searle, 1984), insects (e.g. Turner, 1971; Moran et al., 1980; Shaw, 1981; Butlin and Hewitt, 1985; Mallet, 1986; Hamrick and Hamrick, 1989) and plants (Millar, 1983) (for a more thorough review see Barton and Hewitt, 1985). Trait values may change gradually over long distances (smooth cline) or abruptly over a short distance with little change in the character outside the area of transition (stepped cline). A gradual change in a trait over a long distance may be interrupted by a sharp break over a short distance (saw-tooth cline). These patterns arise from the interaction of selection and gene flow along the cline (Mayr, 1970; Endler, 1977).

Debates about the origin of clinal variation remain unresolved (Barton and Hewitt, 1985; Harrison, 1990) because clinal variation may arise as a result of either primary or secondary intergradation (Mayr, 1942). Primary intergradation exists if variation develops within a continuous series of populations, producing intermediates in the zones of contact between phenotypically distinct populations (Mayr, 1942; Lincoln et al., 1982). Secondary intergradation results from contact between phenotypically distinct populations that differentiated in allopatry (*op cit.*). Primary intergradation thus implies that differentiation between populations and subsequent speciation are sympatric.

Secondary intergradation implies that differentiation and speciation are allopatric (vicariance event). Evidence supporting either primary or secondary intergradation as the origin of clinal variation thus provides insights into modes of speciation. These two forms of intergradation can produce identical patterns of variation, making interpretation of the observed patterns difficult or, according to Endler (1977), even futile.

The ambiguity regarding the origin of clinal variation is reduced when more than one set of characters is examined and their patterns compared. Clines produced by primary intergradation should reflect underlying environmental gradients. Selection associated with these underlying gradients would not be expected to influence all traits in a similar manner because some will influence fitness with respect to the changing environmental parameters, and some will not (Harrison, 1990; Thorpe, 1991). Therefore, congruence among diverse (uncorrelated) character sets such as morphology, life history, and allozyme frequencies, is unlikely if the cline has arisen through primary intergradation.

Covariation between trait values and environmental variables is often taken as evidence for primary intergradation, with selection *in situ* responsible for the observed clinal variation (e.g. Langvatn and Albon, 1986; Confalonieri and Colombo, 1989; James, 1991; Thorpe, 1991). If there is a sharp change in the environment, a sharp change in a particular character may result, producing a stepped or saw-tooth cline. Where clines occur in the absence of environmental gradients, secondary intergradation is often inferred (Smith, 1979; Butlin and Hewitt, 1985; Thorpe, 1987a; Tardif and Morisset, 1990). Generally, when secondary intergradation has taken place, covariance is expected among many characters due to the combined effects of genetic drift and

selection in the isolated populations (Thorpe, 1987b, 1991).

Studies of the patterns of variation in multiple characters also address the question of the role of selection in cladistic evolution (Futuyma, 1987, 1989; Grant, 1989; Brooks and McLennan, 1991). Clines and hybrid zones provide the opportunity, in this case, to examine the process of speciation in systems where reproductive isolation is not yet complete (Barton and Hewitt, 1985; Harrison, 1990). Patterns of covariation, both within and among populations, of adaptive and potentially neutral traits across a hybrid zone can be examined to directly address the hypothesis that cladogenesis is uncoupled from processes of adaptive differentiation (Brooks and McLennan, 1991).

In this study, I examine clinal variation in size and shape among populations of waterstriders, *Aquarius remigis* Say (Hemiptera: Gerridae), in the eastern U.S. and Canada. *Aquarius remigis* is a common semiaquatic bug found on streams throughout North America (Drake and Harris, 1934; Calabrese, 1977; Andersen, 1990). Because most populations are monomorphically wingless (Calabrese, 1979), migration among populations is limited. Both mark-and-recapture studies (Fairbairn 1985a,b, 1986; Blanckenhorn, 1991) and allozymic studies of genetic population structure (Zera, 1981; Preziosi and Fairbairn, 1992) have demonstrated a high level of differentiation and genetic isolation among populations on streams separated by as little as 1 km. *Aquarius remigis* is thus an excellent subject for studies of clinal variation because the limited gene flow facilitates differentiation of local populations through the effects of both genetic drift (Zera, 1981; Preziosi and Fairbairn, 1992) and selection (Blanckenhorn, 1991).

Previous descriptions of intraspecific morphological variation across North America lead to conflicting conclusions about the specific status of *A. remigis* (Andersen,

1990). Blanckenhorn (1991) and Blanckenhorn and Fairbairn (unpublished data) have found adaptive differentiation in body size and life-history over geographic scales from 1km to 750km. Size, colour, internal morphology, and wing form (macropterous or apterous forms) also vary on a broader scale, among populations throughout the United States (Drake and Harris, 1934; Michel, 1961; Calabrese, 1974). Subspecific (Calabrese, 1974) or even specific (Michel, 1961) status has been proposed for various regional 'populations', but these formal taxonomic revisions have not been generally accepted (Polhemus and Chapman, 1979).

Preziosi and Fairbairn (1992) studied allozyme variation in populations of *A. remigis* from across North America and found that those from North Carolina showed unexpectedly high biochemical genetic differentiation from the remaining populations. The degree of genetic differentiation was large enough to suggest specific status for these populations. A study by Zera (1981) of genetic variation in *A. remigis* showed a sharp break in allele frequencies between populations from Maine to New Jersey and those from Virginia. A clinal analysis of the allozyme variation along a transect extending from southern Quebec to southern North Carolina (Gallant et al., 1993) revealed two distinct 'types' of *A. remigis*, north and south of a geographically restricted transition zone in northern Pennsylvania and southern New York. The presence of heterozygote deficiencies in this zone of transition indicates restriction of gene flow between the northern and southern populations. The degree of differentiation was again sufficient to suggest speciation. Gallant et al. (1993) postulated that the differentiation occurred as a result of isolation in glacial refugia east and west of the Appalachian Mountains during the Pleistocene (Espenshade, 1964; Redfern 1983; Pielou, 1991). The geographic

location of the restricted zone of transition found in their study corresponded with the location of the southern limit of glaciation, rather than the location of any abrupt transition in the current environment (Espenshade, 1964; Cohen, 1973; Preziosi and Fairbairn, unpublished data). The zone of transition is thus hypothesized to be the result of secondary contact between the populations from these refugia.

Lack of clinal variation in allozyme frequencies within regions (Preziosi and Fairbairn, 1992; Gallant et al., 1993) suggests that the observed biochemical genetic variation is primarily neutral. In contrast, the pattern of morphological variation in size and shape is expected to be adaptive (Fairbairn, 1984; Blanckenhorn, 1991; Blanckenhorn and Fairbairn, unpublished data). In light of the hypothesis that vicariance events were responsible for the pattern of allozymic variation and that this variation is potentially neutral, the purpose of the present study is to examine the relationship between adaptive differentiation in morphometric traits and the possibly neutral biochemical differentiation. I will examine the variation in morphology in *A. remigis* to determine if these morphological characters display clinal variation and if this variation follows the latitudinal environmental cline, or rather reflects a sharp break in size and shape between the southern and northern regions, as defined in Gallant et al. (1993).

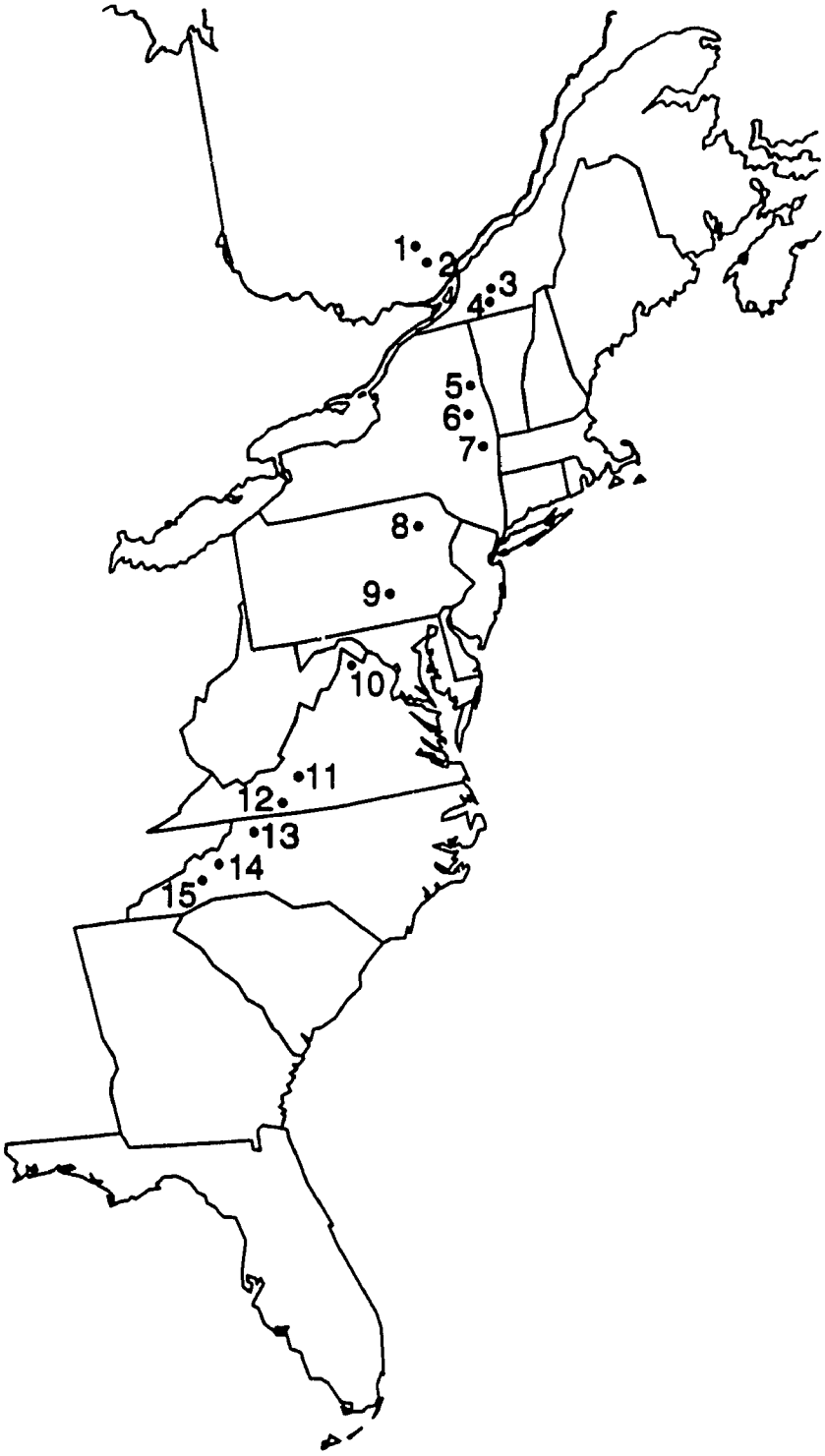
METHODS.

Adults from 15 populations of *Aquarius remigis* were sampled in September and October of 1990 and/or in May and June of 1991 along a transect at approximately 150 km intervals from southern Quebec to southern North Carolina (Fig. 1). Sites 1, 3, 5, 6, 7, 8, 9, 10, 11, and 13 were sampled in the fall only and sites 2 and 14 were sampled in the spring only. Sites 4, 12, 15 were sampled in both the fall and spring to assess possible seasonal effects. Multiple regression analysis indicated no significant effect of season on body size for any of the eight variables (p values ranged from 0.202 to 0.979). Samples 5 through 13 and sample 15 (fall samples only) were the same samples used for electrophoresis in Gallant et al. (1993). The samples were collected using pond nets from small, fast-flowing streams with little or no emergent vegetation and generally sheltered by overhanging trees and shrubs, the typical habitat of this species (Calabrese, 1977; Gailbraith and Fernando, 1977). Samples were returned alive to the lab and then either preserved in 70% ethanol or frozen at -60°C.

Measurements were made on preserved animals. Only apterous individuals were measured as the sample sizes for macropterous individuals captured were very low: 1 of 83 males (1.2%) and 1 of 223 females (0.4%) from site 3; 1 of 45 females (2.2%) from site 4 (spring only); 2 of 34 females (5.9%) from site 9; 4 of 76 males (5.3%) and 6 of 71 females (8.5%) from site 15 (fall only). Within each sample, 20 males and 20 females were subsampled at random for measurement. Samples from sites 7, 8, and 9 contained only 17, 15, and 10 females and 20, 17 and 13 males, respectively, and thus all individuals were included in the measurement sample.

Each individual was measured ventrally by placing it in a glass bottomed box and holding it in place with a block of foam. The box was placed glass side up under a

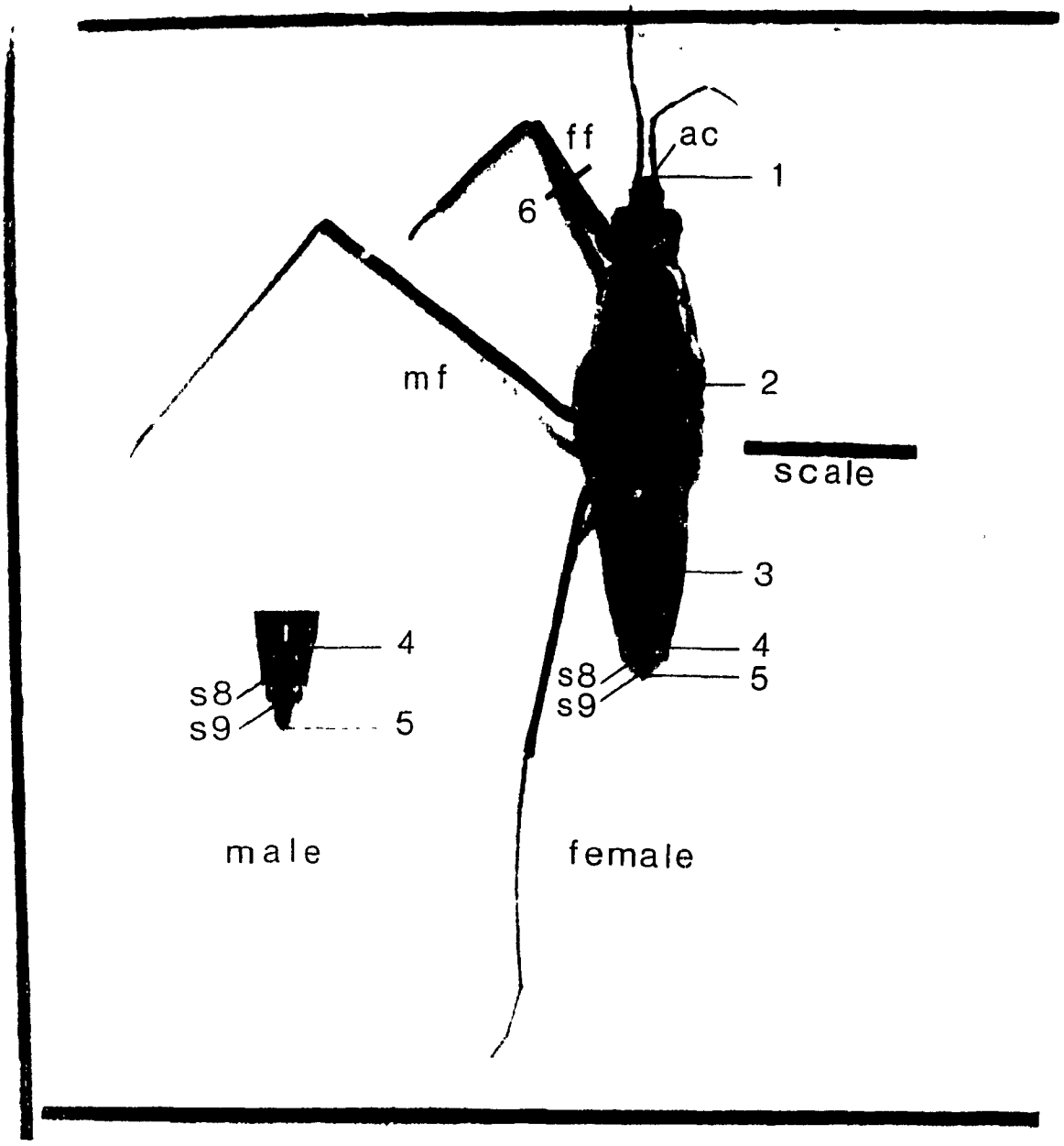
Figure 1. Map of eastern Canada and USA indicating location of sites from which *Aquarius remigis* were collected. In most cases, site names refer to the nearest town. Latitude and longitude are given in parentheses. 1. Mont Tremblant Park, Quebec (46°30'N, 74°33'W); 2. Morin Heights, Quebec (45°57'N, 74°10'W); 3. Acton Vale, Quebec (45°39'N, 72°33'W); 4. Bromont, Quebec (45°18'N, 72°36'W); 5. Underwood, New York (44°37'N, 73°43'W); 6. Weavertown, New York (43°38'N, 73°55'W); 7. Albany, New York (42°34'N, 73°55'W); 8. Scranton, Pennsylvania (41°27'N, 75°45'W); 9. Harrisburg, Pennsylvania (40°19'N, 76°45'W); 10. Winchester, Virginia (39°11'N, 78°07'W); 11. Roanoke, Virginia (37°27'N, 79°41'W); 12. Hurricane Knob, Blue Ridge Parkway, Virginia (36°42'N, 80°27'W); 13. Doughton Park, Blue Ridge Parkway, North Carolina (36°12'N, 81°11'W); 14. Julian Price Memorial Park, Blue Ridge Parkway, North Carolina (35°53'N, 81°55'W); 15. Buck Cr. Gap, Blue Ridge Parkway, North Carolina (35°39'N, 82°07'W).


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

microscope and the image projected onto a monitor where linear distances were measured using a digitizing system attached to a PC. Each image was calibrated using a reference line of known length affixed to the glass. Eight measurements were made of each individual (Fig. 2). Total length (ttl) was assessed as the distance along the ventral midline from the anterior edge of the anteclypeus to the distal end of the last genital segment. Head and thoracic length (hdth) was assessed as the ventral midline distance from the anterior edge of the anteclypeus to the suture between the thoracic and abdominal segments. Abdominal length (abdl) was assessed as the distance along the midline between this suture and the distal end of the seventh abdominal sternum. External genital length (gen) was assessed as the length of the eighth and ninth abdominal segments, which are modified to form the genital segments. Abdominal width was measured at the suture of the fifth and sixth segments (abdw). Three leg measurements were taken from the right side of each individual: front femoral width at the midlength (ffw) and length (ffl) and mid-femoral length (mf). Total length was not used in the multivariate analyses because it is the sum total of genital, abdominal and head and thoracic lengths. In multivariate analyses, patterns of correlation among variables are analyzed. If a variable is used that is a composite of other variables used, correlations become inflated (Tabachnick and Fidell, 1989).

These measurements were chosen because of their ecological relevance. General size is potentially under selection for a number of reasons. Larger animals may have increased overwinter survival (Gunnarsson, 1988) and foraging success (Rubenstein, 1984). Loading constraints on females associated with carrying mates may select for smaller males (Fairbairn, 1990). Smaller males may also be more agile (McLauchlan

Figure 2. Photograph of a ventral view of an apterous female *Aquarius remigis* with the corresponding male genital segments, indicating where the eight different morphological measurements, rounded off to the cm x 10^{-3} for calculations, were taken: total length (1-5); head and thoracic length (1-2); abdominal length (2-4); genital length (4-5); abdominal width, 3; front femoral width at the midlength, 6; front femoral length; mid-femoral length. Abbreviations: ac, anteclypeus; s8, abdominal segment 8 (first genital segment); s9, abdominal segment 9 (second genital segment); ff, front femur; mf, mid-femur. Scale: 5.006mm (remaining lines show frame of mounting box).

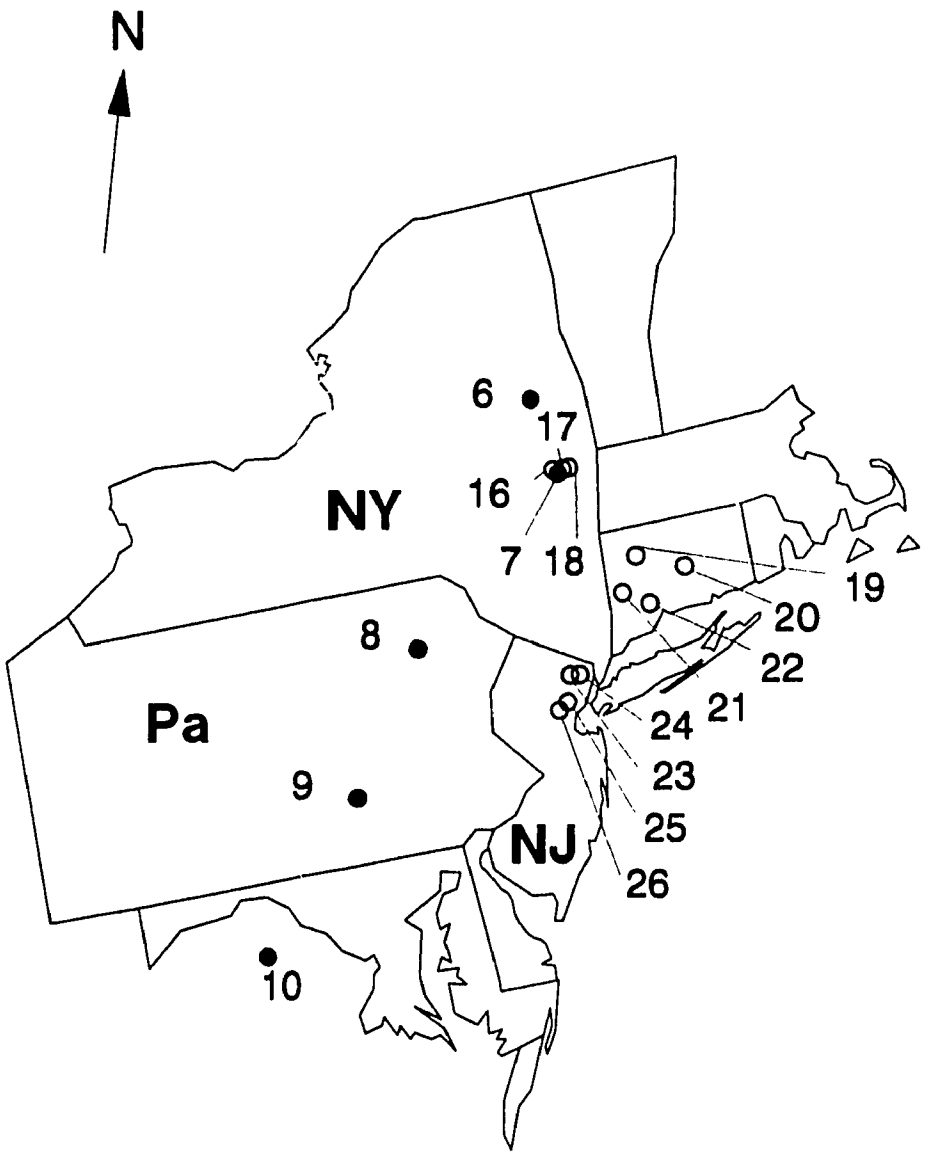

ff
ac
1
6
mf
2
scale
3
4
s8
s9
5
4
s8
5
s9
male
female

and Allen, 1987). Thoracic size has been shown to differ between apterous and macropterous individuals (Fairbairn, 1992) and is a measurement of general size. Larger abdomens (length and width) in females may be favoured to accommodate more or larger eggs (Fairbairn, 1988, 1990; Blanckenhorn, unpublished data). Wide front femora may be favoured in males because of their association with increased mating success (Rubenstein, 1984; McLean, 1990). Larger front femora may also improve foraging ability in males and females (Rubenstein, 1984). Mid-femoral length may be larger in response to intrasexual selection (Hayashi, 1985) in males or to loading in females (McMahon and Bonner, 1983). The length of the genitalia is an important trait as its size differs between males and females. Specific characteristics of the genitalia have been used to characterize species of gerrids (Michel, 1961; Polhemus and Chapman, 1979; Andersen, 1990) although they are not thoroughly examined here. Only the length of the external portion of the genital segments was measured.

To verify the findings of the above data set, eleven populations of adult *A. remigis* were sampled from sites in New York (sites 16 through 18), Connecticut (sites 19 through 22) and New Jersey (sites 23 through 26) (Fig. 3) comparable to those sampled along the Quebec to North Carolina transect. Sites 17, 18, 24 and 26 were collected in May, 1992. The remaining sites were sampled in June, 1993. Samples were returned alive to the lab and frozen at -60°C.

The same eight body size measurements were taken of each individual using the same method as described for the previous data set. Sample sizes ranged from 3 to 41 for males and from 4 to 33 for females. Individual sample sizes, means and standard deviations are given in Appendix B. All individuals collected were included in the

Figure 3. Map indicating the location of *a posteriori* sampling sites from which *Aquarius remigis* were collected (open circles). In most cases, site names refer to the nearest town. Latitude and longitude are given in parentheses. 16. Paint Mine Cr., Albany, New York (42°39'N, 74°05'W); 17. Five Rivers, Albany, New York (42°37'N, 73°53'W); 18. Albany Rural Cemetery, Albany, New York (42°42'N, 73°44'W); 19. John A. Minetto State Park, Connecticut (41°53'N, 73°13'W); 20. Stratton Brook State Park, Connecticut (41°52'N, 72°49'W); 21. Putnam Memorial State Park, Connecticut (41°20'N, 73°21'W); 22. Redding, Connecticut (41°19'N, 73°15'W); 23. Wanaque, New Jersey (41°05'N, 74°18'W); 24. Ramapo Valley, New Jersey (41°05'N, 74°13'W); 25. Hibernia, New Jersey (40°58'N, 74°28'W); 26. Lincoln, New Jersey (40°56'N, 74°21'W). Sites 6, 7, 8, 9 and 10 were included from the original transect as reference points (closed circles).


N
NY
Pa
NJ
6
17
16
7
18
19
20
22
21
24
23
25
26
8
9
10

measurement samples.

The precision of the measurements of both the transect and *a posteriori* individuals was assessed as in Fairbairn (1992) by repeated measures (N=5) of three males and three females. Precision was expressed as the 95% confidence interval for each individual. Precision did not differ between the sexes, and ranged from $\pm$0.01mm for female front femoral width to $\pm$0.16mm for male head and thoracic length and female mid-femoral length. As a percentage of the mean, precision varied from $\pm$0.13% for male mid-femoral length to $\pm$4.20% for male abdominal length. Most traits had a percentage of the mean below $\pm$1%.

All statistical analyses were conducted using SYSTAT (Wilkinson, 1989) and are described in detail in the results section. When the frequency distributions for each of the eight measurements were examined within each sex (for PCA and analysis of within population variance) or within each region for each sex (for discriminant analysis), these distributions were significantly deviant from normality in most traits. These deviations are not surprising, however, because the individuals are not taken from a single population. Though a normal distribution in the data is recommended, the analyses are robust to these deviations (Tabachnick and Fidell, 1989). All other assumptions for these analyses have been met and are described in Appendix A. While not completely eliminating the deviations from normality within the data set, $\log_{10}(X+1)$ transforming the data was effective in reducing these deviations (Tabachnick and Fidell, 1989). The data were thus log-transformed prior to principal components (PCA) and discriminant analyses. Log transformation of the data also removed the potential correlation between means and variances, making them independent. Comparisons of within population

variances are thus based on log-transformed data.

RESULTS.

Univariate analysis:

The mean values for each variable by site by sex are plotted against latitude in Figure 4 and the results of the statistical analyses are given in Tables 1 and 2 (means and standard deviations for the 8 measurements taken for all 15 populations are given in Appendix B). Preservation method had a slight effect (ranging from an increase of 0.0003% for female total length to an increase of 0.05% for male head and thoracic length from frozen to ethanol preserved individuals) on body size. Measurements of frozen individuals were thus corrected to ethanol preserved individuals. To examine latitudinal and regional trends, stepwise regressions were performed using population means for each sex with latitude as the covariate, region as the grouping variable, and latitude*region as an interaction term. Region was coded as 0 for 'southern populations' from Harrisburg (site 9) south to Buck Cr. Gap (site 15), and as 1 for 'northern populations' from Scranton (site 8) north to Mont Tremblant Park (site 1). This classification is based on the location of the transition in allozyme frequency identified in Gallant et al. (1993) and is consistent with the shifts in body size identified in Figure 4. In Gallant et al. (1993), Scranton was determined to be a transition site and therefore difficult to classify. This site has, however, been included with the northern populations because it is the southern limit for northern alleles for three of five loci and its location is north of the southern limit of Pleistocene glaciation. The choice of variables included at the first step of the stepwise regression analysis was based on which of the three variables (latitude, region, and latitude*region) correlated most strongly with the dependent variable (morphological measurement). The correlations are independent of the presence of the other variables. In all cases, latitude was most strongly correlated

Figure 4. Morphological measurements of *Aquarius remigis* as a function of latitude. Trait means in mm for each site plotted for total length (ttl), head and thoracic length (hdth), abdominal length (abdl), abdominal width (abdw), genital length (gen), mid-femoral length (mf), front femoral length (ffl) and width (ffw). (●) males; (▲) females

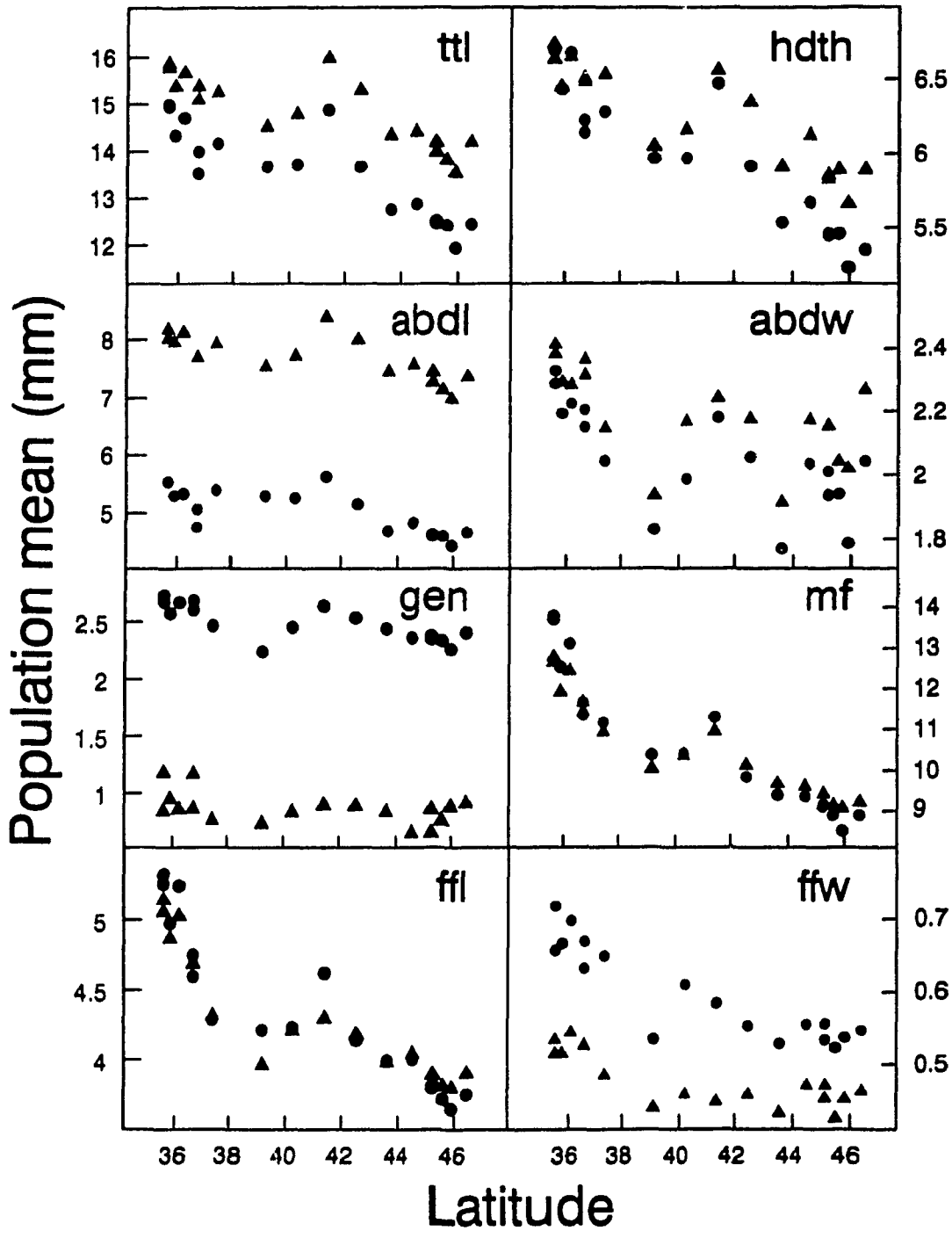

ttl
hdth
abdl
abdw
gen
mf
ffl
ffw
Population mean (mm)
Latitude
16
15
14
13
12
6.5
6
5.5
8
7
6
5
2.4
2.2
2
1.8
2.5
2
1.5
1
14
13
12
11
10
9
5
4.5
4
0.7
0.6
0.5
36
38
40
42
44
46

Table 1. Results of stepwise multiple regressions for *A. remigis* males, with latitude as the covariate, region as the grouping variable and a latitude*region interaction term.

Dependent Variable	Coefficients			Full model	
	Region[1]	Latitude	Interaction	R^2	F
ttl	11.574*	-0.233**	-0.250*	0.876	33.057***
hdth	0.623***	-0.178***	ns	0.913	78.308***
abdl	ns	-0.067***	ns	0.556	20.041***
abdw	ns	-0.063***	0.007**	0.615	11.989***
gen	0.349***	-0.066***	ns	0.770	25.108***
mf	1.613*	-0.559***	ns	0.909	74.712***
ffl	0.649**	-0.196***	ns	0.911	76.398***
ffw	ns	-0.014***	ns	0.808	67.382***
PC1[a]	1.014*	-0.314***	ns	0.901	68.016***
PC1[b]	-0.794*	0.297***	ns	0.917	82.452***
PC2[b]	ns	ns	ns	ns	ns

* $p \leq 0.05$ ** ≤ 0.01 *** ≤ 0.005

[1] northern region coded as 1; southern region coded as 0

[a] first principal component of a general principal components analysis (PCA)

[b] first and second principal components of a size-adjusted PCA

Table 2. Results of stepwise multiple regressions for *A. remigis* females, with latitude as the covariate, region as the grouping variable and a latitude*region interaction term.

Dependent Variable	Coefficients			Full model	
	Region[1]	Latitude	Interaction	R^2	F
ttl	1.580***	-0.330***	ns	0.852	43.233***
hdth	0.563***	-0.137***	ns	0.884	56.918***
abdl	6.626*	-0.094**	-0.141*	0.813	20.236***
abdw	ns	-0.020*	ns	0.341	8.294**
gen	ns	ns	ns	ns	ns
mf	1.346*	-0.438***	ns	0.918	84.096***
ffl	-5.885*	-0.226***	0.138**	0.921	54.282***
ffw	ns	-0.008***	ns	0.703	37.818***
PC1[a]	1.204*	-0.327***	ns	0.868	49.116***
PC1[b]	-0.772*	0.297***	ns	0.922	88.142***
PC2[b]	ns	ns	ns	ns	ns

* $p \leq 0.05$ ** ≤ 0.01 *** ≤ 0.005

[1] northern region coded as 1; southern region coded as 0

[a] first principal component of a general principal components analysis (PCA)

[b] first and second principal components of a size-adjusted PCA

with the morphological measurements and was therefore included at the first step in the stepwise regression (Tabachnick and Fidell, 1989). In order for region or the interaction term to be included at the next step, the variable must predict a significant portion of the variance in the morphological trait not already explained by latitude. One of these variables may not do so significantly despite the fact that it is almost as highly correlated with the morphological trait as was latitude. Therefore, the final subset of variables is the best predictor of the morphological measurement and does not include any variables that do not provide additional predictive power (*op cit.*). A variable considered for the regression, but not included in the final model, does not influence the correlation coefficients or significance levels (or any other statistic) of the variables already in the equation. All of the traits except female genital length increased significantly with decreasing latitude (Tables 1 and 2, Figure 4). When the probabilities for entry into the model were adjusted for multiple comparisons (sequential Bonferroni, Rice, 1989), latitude remained significant in predicting all of the variables, with the exceptions of male total length and female abdominal length and width. If the interaction term was not included in the model, latitude also remained significant in predicting male total length and female abdominal length. Region was significant for all the traits except front femoral width, abdominal width, male abdominal length, and female genital length. When the probabilities for entry into the model were adjusted, region remained significant in predicting head and thoracic length, male genital length, and female total length. The significant regional effect can be seen in Figure 4 as a sharp break (increase in size) between latitudes 40° and 42°, giving an overall saw-tooth cline for these traits. The latitudinal trend can be seen in Figure 4 as an overall decrease in body size with

increasing latitude within regions. Region*latitude was significant for female abdominal length and front femoral length and for male total length and abdominal width, indicating that the pattern of clinal variation within each region differs. The significant effects of region and/or of region*latitude for 6 of the 8 traits in males and 5 of the 8 traits in females indicate significant regional differentiation of *A. remigis* populations, over and above the strong effects of latitude. The models including region account for 77 to 92% of the among-population variance in morphological variables. With the exception of female abdominal width, these univariate models all remain significant in predicting variation in the morphological variables when corrected for multiple comparisons (*op cit.*).

Multivariate analysis:

Principal components analysis allows me to reduce the seven morphometric variables to one or two sets of component scores. The pattern of scores can be examined in the same way as the individual variables were examined in a univariate sense. Principal components analysis is a common method of analyzing geographic variation in several traits (e.g. Smith, 1979; Thorpe, 1987a,b).

To determine if the analyses were robust to the presence of outliers and deviations from normality, the principal components analysis was performed using both transformed and untransformed data, with both univariate and multivariate outliers, and transition sites (sites 7 and 8, Scranton and Albany) included and removed. The patterns of correlation of the traits with the principal components, the number of components extracted, and the amount of variation explained by each component, were similar in all cases. Therefore,

the complete data set was used for the analysis presented here and was log-transformed in order to remove the correlation between means and variances. Males and females differ in size and shape (Fairbairn, 1992), and when both sexes are included in a PCA, the variance due to sex (53.2%) obscures visualization of any other pattern within the data set. The sexes were therefore analyzed separately. For both sexes, the general PCA extracted a single component (PC1), which explained 66 to 74% of the among population variance in size and shape (Table 3). All seven variables are positively correlated with that component, indicating that overall size accounts for the greatest portion of the variance among sites. The first principal component generally extracts an index of size (Somers, 1989). A stepwise regression of principal component scores on latitude, region and latitude*region reveals significant regional and latitudinal effects for both sexes (Tables 1 and 2). These two variables predict 90.1% and 86.8% of the variation in overall size for males and females, respectively. To illustrate this pattern of overall variation in size with latitude, the mean component scores for PC1 for each site and each sex were plotted against latitude (Fig. 5). Size decreases with increasing latitude but increases between latitudes 40°N and 42°N, giving a saw-tooth cline pattern similar to that shown for the single traits (Fig. 4). Overall size, therefore, varies clinally within regions and differs between regions, with Harrisburg (site 9, 40°19'N) being intermediate between northern and southern forms. Mean principal component scores for each of the regions are plotted along PC1 and illustrate the regional differentiation in overall size: southern *A. remigis* are on average larger than their northern counterparts (Fig. 6).

When the isometric size effect in the covariance matrix is removed by doubly

Figure 5. Mean population scores on the first PC of a general principal components analysis based on seven morphometric measurements of *Aquarius remigis*, plotted against latitude. Principal components analysis was performed for each sex and the mean of the scores corresponding to each site calculated. Separate analyses was conducted for males (●) and females (▲), but the results are presented in the same figure for conciseness.

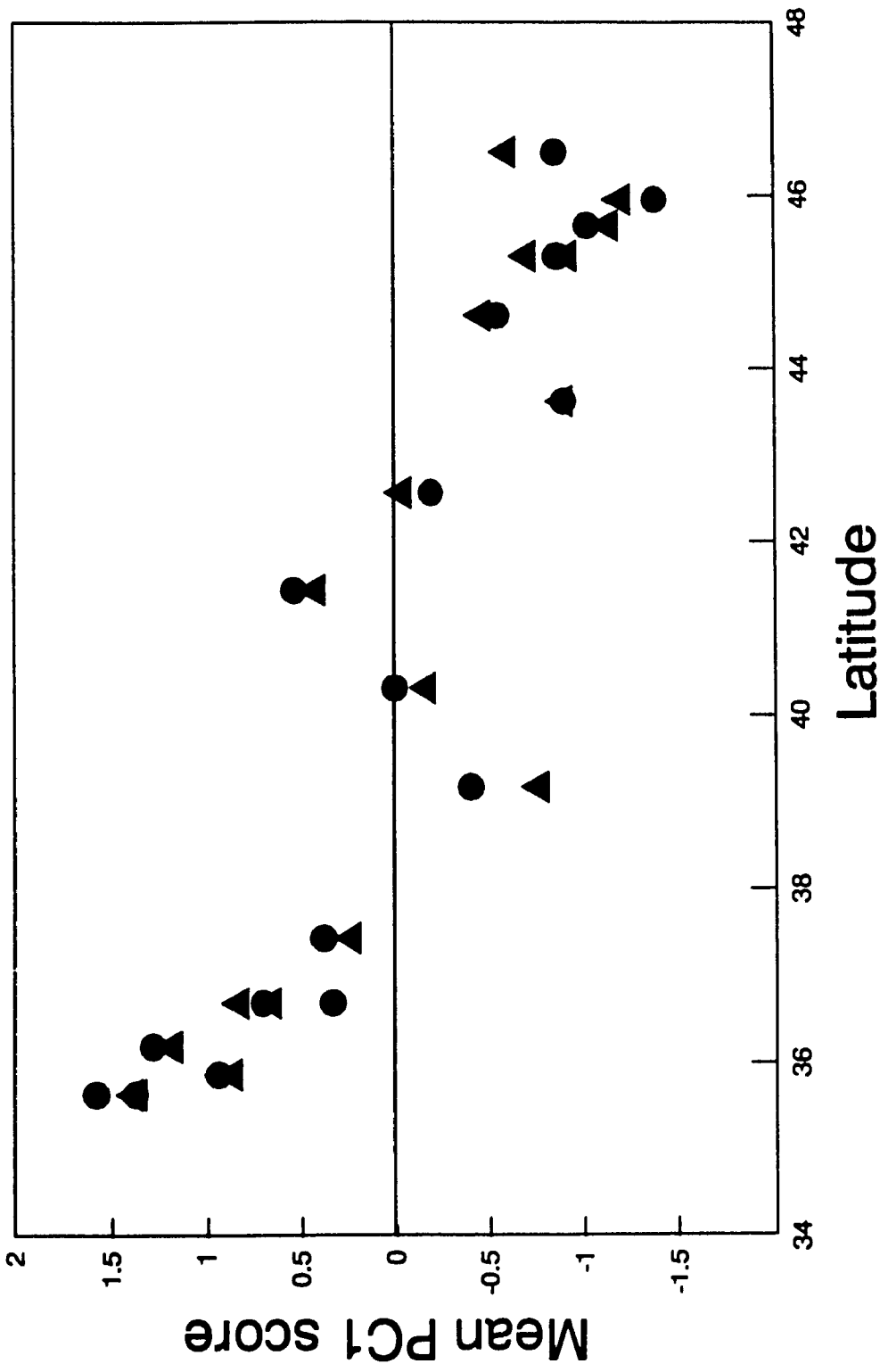

Mean PC1 score
2
1.5
1
0.5
0
-0.5
-1
-1.5
34
36
38
40
42
44
46
48
Latitude

Figure 6. Mean regional scores on the first PC of a general PCA based on seven morphometric measurements of *Aquarius remigis* for southern males (o), northern males (●), southern females (△) and northern females (▲). Northern individuals are those from sites 1 through 8 and southern individuals are those from sites 9 through 15. Component scores were calculated separately for each of the sexes, but plotted on the same figure for conciseness. Horizontal bars indicate $\pm$ 1 s.d.

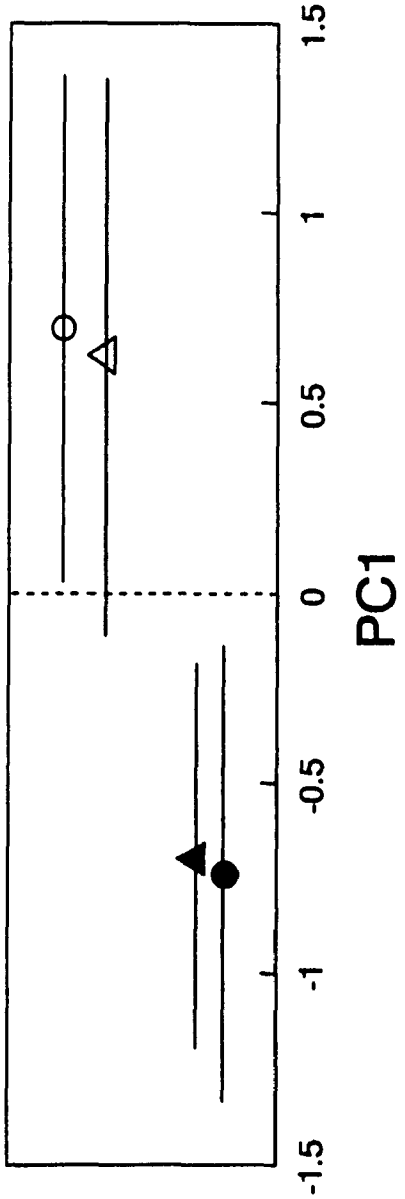

1.5
1
0.5
0
-0.5
-1
-1.5
PC1

Table 3. Principal components analysis component loadings for *Aquarius remigis*. Loadings represent the correlations between each of the seven morphometric traits measured and the corresponding principal component. Analyses were conducted separately for males and females.

Trait	General PCA		Size-adjusted PCA			
	PC1		PC1		PC2	
	Male	Female	Male	Female	Male	Female
abdl	0.788	0.811	0.272	0.388	-0.786	0.694
hdth	0.944	0.873	-0.112	0.160	-0.686	0.757
abdw	0.827	0.770	0.706	0.619	0.172	-0.030
ffl	0.950	0.949	-0.899	-0.891	0.199	0.192
ffw	0.878	0.807	0.770	0.702	-0.070	0.258
mf	0.962	0.945	-0.955	-0.945	0.094	0.172
gen	0.613	0.419	0.606	0.241	0.557	-0.911
Eigenvalue	5.173	4.633	3.266	2.798	1.481	2.019
%Variance explained	73.9	66.2	46.7	40.0	21.2	28.8

centering the data matrix (Somers, 1989), patterns of differentiation in shape are more evident. For both males and females, individuals with short, thick legs and wide abdomens score highly on the first component of the size-adjusted PCA (Table 3). Males with long genital segments, as well as short abdomens and head and thoracic segments, and females with short genitals and a long head, thorax and abdomen score highly on the second component (PC2; Table 3). To illustrate the relationship between shape and latitude, the mean scores for the first and second principal components for each population are plotted against latitude (Fig. 7). Mean PC1 scores increase with latitude and show a slight, but significant, shift between latitudes 40° and 42° (Tables 1 and 2, Figure 7). Scores for PC2 show neither an apparent linear relationship with latitude, nor any regional differentiation. Both latitude and region were significant in predicting the variation in PC1 scores, accounting for 92% of the variation for both males and females, but were not significant in predicting the variation in PC2 scores (Tables 1 and 2). Based on the pattern of correlation of each trait with the first component, northern individuals have short, thick legs and wide abdomens relative to their body size and southern individuals have long, thin legs and narrower abdomens relative to their overall size. Mean principal component scores for each of the regions are plotted along PC1 and illustrate that, on average, regions are differentiated in shape along this component (Fig. 8).

Discriminant analysis was used to determine if individuals could be accurately classified by regions and if so, which of the measurements were important in distinguishing between regions. To determine if the presence of outliers in the data set and deviations from normality reduced the reliability of the analysis, the discriminant

Figure 7. Mean population scores on PC1 and PC2 of a size-adjusted principal components analysis based on seven morphometric measurements of *Aquarius remigis*, plotted against latitude. Principal components analysis was performed for each sex and the mean of the scores corresponding to each site calculated for both principal components. Separate analyses were conducted for males (●) and females (▲), but the results are presented in the same figure for conciseness.

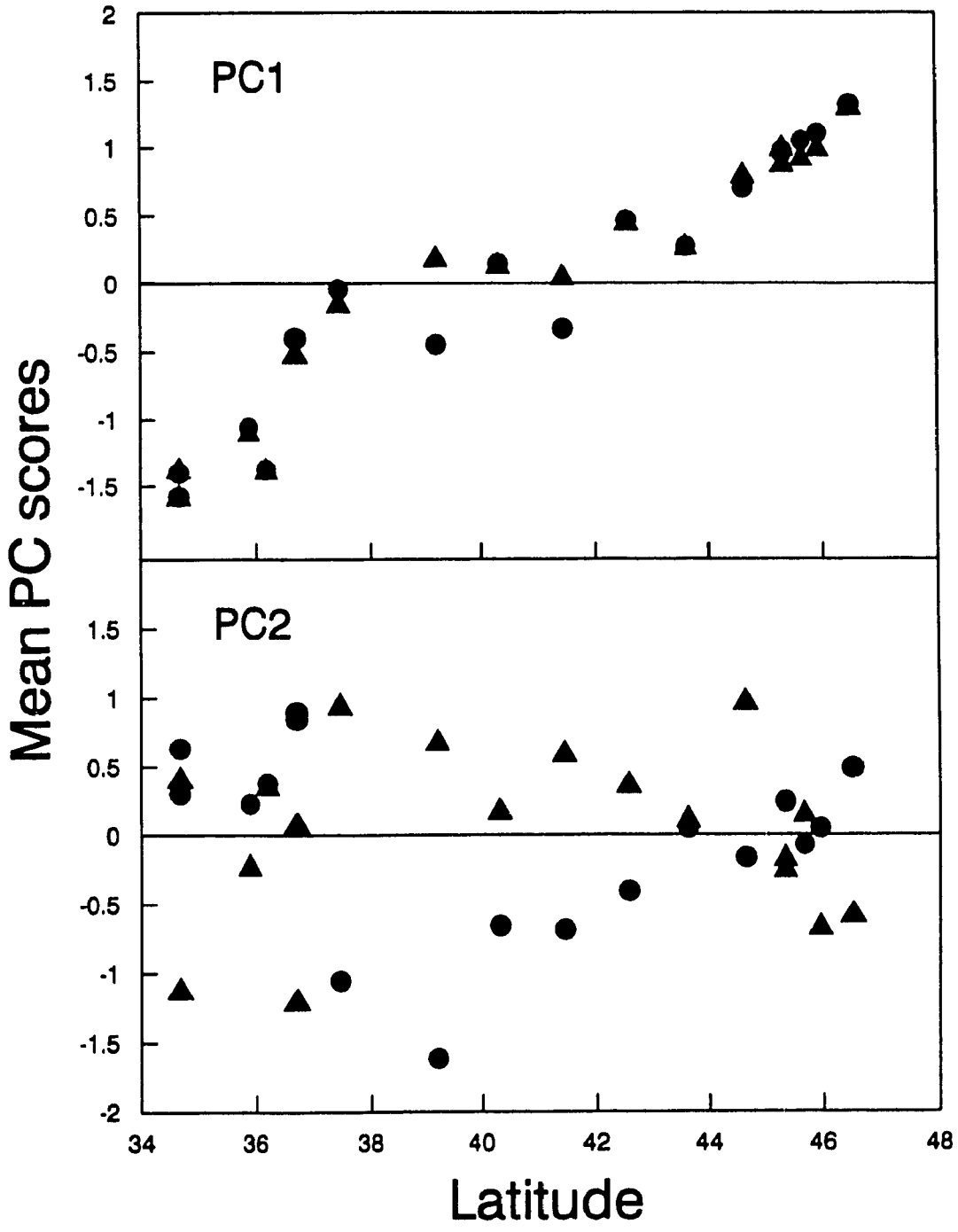
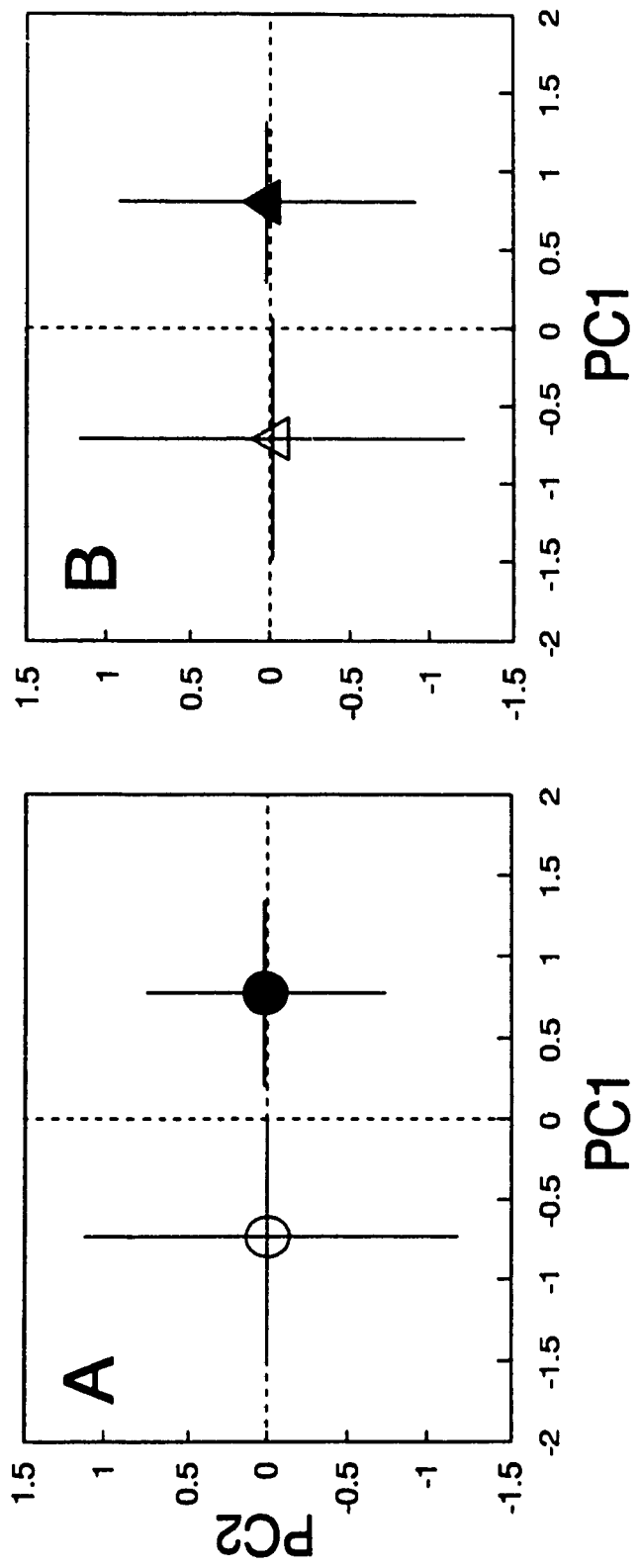

PC1
PC2
Mean PC scores
Latitude
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
34
36
38
40
42
44
46
48

Figure 8. Mean regional scores on the first and second PC of a size-adjusted PCA based on seven morphometric measurement of *Aquarius remigis* for A. southern males (o), northern males (●), and B. southern females (▵) and northern females (▲). Component scores were calculated separately for each of the sexes, but plotted on the same figures for conciseness. Vertical and horizontal bars indicate $\pm$ 1 s.d.


A
B
PC2
PC1
PC1
1.5
1
0.5
0
-0.5
-1
-1.5
-2
2

analysis was performed using both transformed and untransformed data, with both univariate and multivariate outliers and the *a priori* transition sites (sites 7 and 8, Scranton and Albany) included or removed. Percent correct classification, significance of the model, and patterns of canonical correlation of the traits, were similar in all cases. The complete, log-transformed data set was therefore used for the analysis presented here. Groups were defined as for the stepwise regressions. To assign individuals to one of the two groups, a classification equation was developed for each group (Tabachnick and Fidell, 1989). A classification score was produced for each individual for each of the two classification equations (north or south) by inserting the data into this equation. The raw score for each of the seven morphometric traits was multiplied by its function coefficient, summed over the seven predictors, and a constant was added, producing the classification score. The individual was assigned either as a northern or a southern 'type' based on the region for which it had the higher classification score. The group classification function coefficients and constants are given in Table 4. The discriminant function accounted for 67% and 63% of the between group variability for males and females respectively indicating that the regions can be reliably separated (Table 5). The models significantly predict group membership (Wilks' lambda 0.334, $p=0.000$ for males; 0.372, $p=0.000$ for females) with a high percent correct classification (93% for males, 89% for females). When outliers were removed, the percent correct classification was increased only slightly (to 97% for males, to 93% for females). Removing the transition zone sites from the calculations improved the model as well (to 95% for males, to 92% for females), indicating that individuals from these sites are difficult to accurately classify. When the model excluding the transition sites was used to predict group

Table 4. Classification coefficients and constants of a discriminant analysis based on seven morphometric measurements of *Aquarius remigis*, run for each sex separately on log-transformed data. Classification function coefficients and constants are used to produce classification scores (see explanation in text). Individuals are assigned to the region for which they have the highest score.

Trait	Group classification function coefficients			
	Males		Females	
	South	North	South	North
abdl	230.271	234.216	768.426	792.351
hdth	617.278	612.312	530.134	531.060
abdw	-61.080	-36.385	-66.004	-53.744
ffl	-244.005	-224.258	-392.588	-381.227
ffw	-283.308	-307.093	-69.516	-80.998
mf	134.127	86.241	186.982	131.519
gen	127.504	134.557	61.358	64.721
constant	-760.655	-699.223	-1234.311	-1183.272

Table 5. Discriminant analysis canonical loadings for *Aquarius remigis* from analyses based on seven morphological measurements run for each sex separately. Loadings represent the correlations between each of the predictors (the seven morphometric traits) and the discriminant function.

Trait	Canonical loadings	
	Males	Females
abdl	-0.494	-0.410
hdth	-0.712	-0.549
abdw	-0.352	-0.278
ffl	-0.759	-0.770
ffw	-0.637	-0.507
mf	-0.885	-0.897
gen	-0.227	-0.150
proportion of variance (squared canonical correlation)	.666	.627

membership for individuals from the transition sites, these sites appear mixed (16 of the 30 males (53%) and 12 of the 25 females (48%) were classifiesd as northern), indicating that the transition zone does not clearly belong to either the northern region or the southern region. This discriminant analysis indicates that sites outside the transition zone are easily classified (the regions are different) but group membership for some of the individuals from the suspected transition sites is ambiguous.

The potential generalization of the model obtained was tested by performing two types of cross validation. When the individuals from half the populations (sites 2, 4, 5, 9, 10, 13, and 15) were used to build the model and the individuals from the remaining half (1, 3, 6, 11, 12, and 14) classified using that model, 87% of the males and 88% of the females were correctly classified. The transition sites (7 and 8) were not included in the data set used in this cross validation procedure as conformity to the model had already been determined. If half of all individuals were used to build the model (half of each sex for each site were chosen at random, and the remaining half classified using that model, 95% of the males and 90% of the females were correctly classified. Thus, if a new individual or population is measured, this model can effectively predict whether it is a northern or a southern 'type'.

In a hybrid zone we expect an increase in the trait variance (Endler, 1977, p14). Scranton and Albany (sites 7 and 8) were defined in Gallant et al. (1993) as transition sites: in Scranton, two of the five polymorphic loci found are typical of northern populations, two others are typical of southern populations and one is intermediate; in Albany, one of the loci is typical of southern populations, while the other four are typical of northern populations. To determine if a transition or hybrid zone could be identified

based on an analysis of within population variances, within site variances of log-transformed data were plotted against latitude for each of the eight variables measured (Fig. 9). Variances were also calculated for each population for both the general principal components and the size-adjusted principal components analyses (Fig. 10).

The average variance in the transition sites (sites 7 and 8) is significantly higher than the average variance in the remaining sites for all traits except genital length and female front femoral width (Table 6, Figure 9). When the probabilities were adjusted for multiple comparisons (Bonferroni), none of the univariate traits was significant, but the general trend is evident. Analyses of the scores from the general and size-adjusted PCA's indicate that male and female size and female shape are more variable within the transition zone. The trend is therefore significant for the multivariate trait. This increase in variance indicates some form of restriction of gene flow.

A posteriori analysis:

Based on the preceding univariate and multivariate analyses, and the pattern of biochemical genetic variation described in Gallant et al. (1993), a transition zone in morphology (size and shape) and allozyme frequency was identified in Pennsylvania and southern New York. The 11 samples from New York, Connecticut and New Jersey were collected *a posteriori* to correspond latitudinally with this transition zone and used here to describe the area of transition in greater detail and to confirm the sharp shift in morphology found between latitudes 40° and 42°.

The classification function previously developed from the discriminant analysis of individuals from populations along the original Quebec-North Carolina transect,

Figure 9. Within-population variance of log-transformed data for each of the eight morphometric measurements of *Aquarius remigis* males (●) and females (▲), plotted against latitude. Ringed data points indicate *a priori* transition sites.

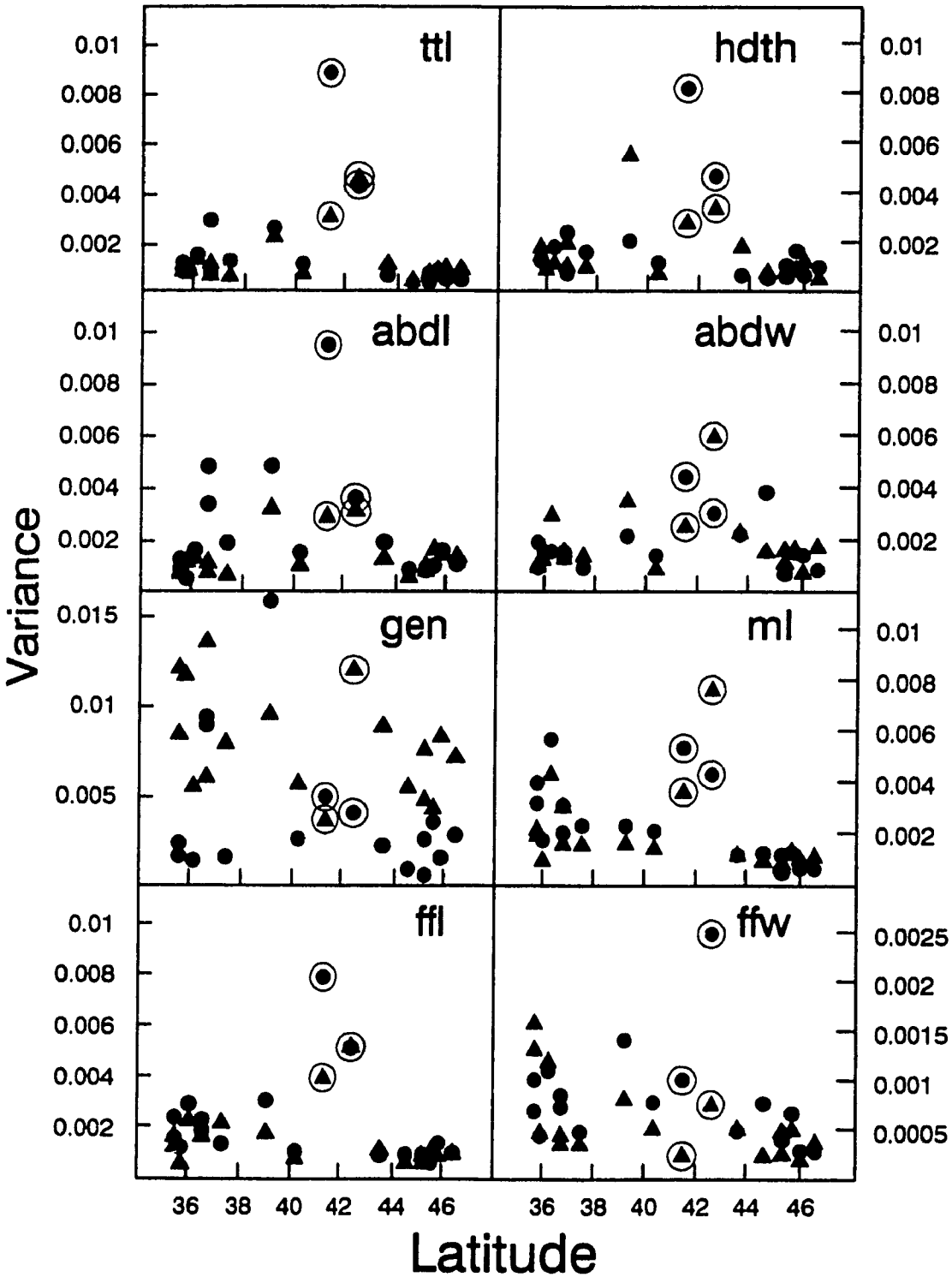

ttl
hdth
abdl
abdw
gen
ml
ffl
ffw
Variance
Latitude
0.01
0.008
0.006
0.004
0.002
0.015
0.005
0.0025
0.0015
0.001
0.0005
36
38
40
42
44
46

Figure 10. Within-population variance for the scores on the first principal components of a general principal components and a size-adjusted principal components analyses based on seven morphometric measurements of *Aquarius remigis* males (●) and females (▲), plotted against latitude. Ringed data points indicate the *a priori* transition sites. Component scores were calculated separately for each of the sexes, but plotted on the same figures for conciseness.

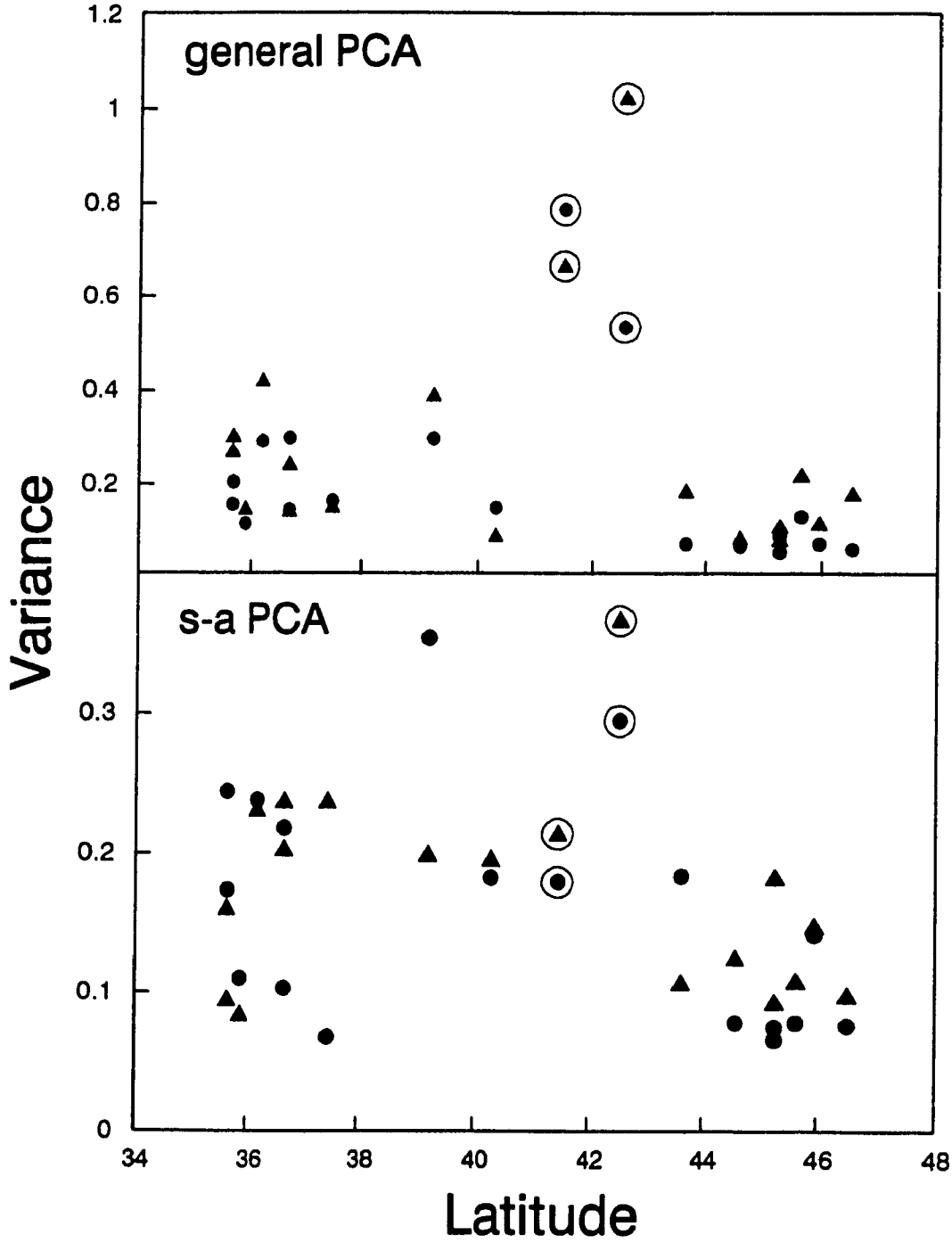

general PCA
s-a PCA
Variance
Latitude
1.2
1
0.8
0.6
0.4
0.2
0.3
0.2
0.1
0
34
36
38
40
42
44
46
48

Table 6. Results of a 1-tailed Mann-Whitney U test comparing the within-site variances in the transition zone (sites 7 and 8) with those outside the zone for the eight morphometric measurements of *Aquarius remigis*.

Trait	Mann-Whitney test statistic (p value)	
	Males	Females
ttl	32 (0.012)	32 (0.012)
abdl	30 (0.025)	30 (0.025)
hdth	32 (0.012)	30 (0.025)
abdw	32 (0.012)	30 (0.025)
ffl	30 (0.025)	30 (0.025)
ffw	29.5 (0.029)	18 (0.389)
mf	30 (0.025)	31 (0.018)
gen	24 (0.131)	18 (0.389)
general PCA	32 (0.0125)	32 (0.0125)
s-a PCA, PC1	25 (0.103)	29 (0.034)

excluding the *a priori* transition zone populations (Scranton and Albany), was used to classify individuals from the new populations as northern or southern 'types'. All samples except New York sites 16 and 17, which were composed entirely of northern individuals, and the males from New Jersey sites 24 and 25, which were entirely classified as southern individuals, contained both northern and southern 'types' in varying proportions (Table 7). The presence of both 'types' within the remaining *a posteriori* samples indicates that these populations are transitional and thus supports the previous conclusion of a zone of transition between northern and southern 'types'. It further suggests an eastern extension of this zone into Connecticut. There is a greater proportion of northern 'type' females than males in the transitional populations. In samples from sites 24 and 25, there are no northern 'type' males. This pattern suggests that introgression is occurring from north to south and preferentially in females. Such asymmetric introgression by sex is consistent with "Haldane's rule" (Haldane, 1922), which predicts that in cases of interspecific hybridization, the heterogametic sex (in this case males) is absent, rare or sterile (Lincoln et al. 1982).

To illustrate the pattern of variation in overall size and shape with latitude, the mean component scores for PC1 for each site and each sex of both a general principal components analysis and a size-adjusted principal components analysis based on both the original and *a posteriori* samples are plotted in Figure 11. For overall size, the *a posteriori* sites follow the saw-tooth cline pattern defined by the pattern of variation along the Quebec-North Carolina transect and confirm the location of the transition area between northern and southern 'types'. Based on this pattern of variation, *A. remigis* from New Jersey (latitudes 40°56' to 41°05') are similar in size to those from Harrisburg

Figure 11. Mean population scores on the first principal component of (A) a general principal components analysis and of (B) a size adjusted principal components analysis based on seven morphometric measurements of *Aquarius remigis*, plotted against latitude, using both the original transect (solid symbols) and the *a posteriori* samples (open symbols). Both the general and the size-adjusted principal components analyses were performed for each sex and the mean of the scores corresponding to each site calculated. A separate analyses was conducted for males (●, o) and females (▲, △), but the results are presented in the same figure for conciseness.

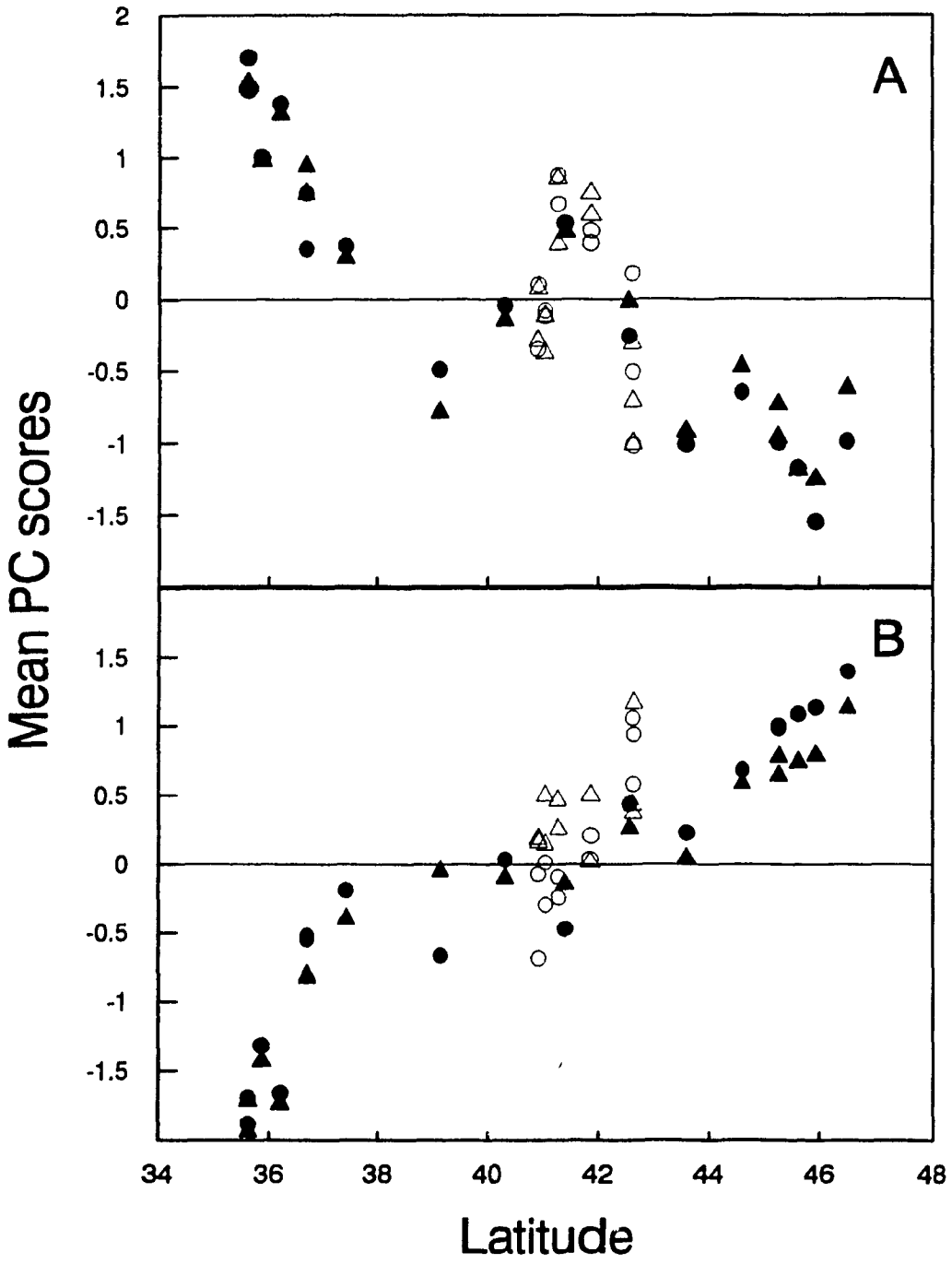

A
B
Mean PC scores
Latitude
2
1.5
1
0.5
0
-0.5
-1
-1.5
34
36
38
40
42
44
46
48

Table 7. Discriminant analysis results indicating the proportion of *Aquarius remigis* individuals characterized as northern 'types' within each of the *a posteriori* sampling sites. Sample sizes are given in parentheses. Analyses were performed separately for each sex.

sites	males	females
16 (NY)	1.000 (29)	1.000 (29)
17 (NY)	1.000 (7)	1.000 (7)
18 (NY)	0.600 (5)	0.875 (8)
19 (CT)	0.122 (41)	0.576 (33)
20 (CT)	0.083 (12)	0.168 (6)
21 (CT)	0.111 (9)	0.750 (4)
22 (CT)	0.039 (26)	0.333 (15)
23 (NJ)	0.367 (30)	0.731 (26)
24 (NJ)	0.000 (10)	0.333 (9)
25 (NJ)	0.000 (3)	0.250 (4)
26 (NJ)	0.100 (10)	0.500 (10)

(latitude 40°19'), and thus also appear to be transition sites for overall size. Sites from Connecticut (latitudes 41°19' to 41°53') are similar in overall size to Scranton (latitude 41°27'), and show the sharp increase in overall body size between latitudes 40° and 42° apparent in Figures 3 and 4. The *a posteriori* samples collected from Albany (42°39' to 42°42') are smaller than those from the original analysis but their size is not atypical of the other northern populations. Shape variation in the *a posteriori* samples does not closely follow the variation in shape defined by the pattern of variation in the earlier samples but it does confirm that a transition zone exists for shape between latitudes 40° and 42°. Northern populations have positive scores on PC1, southern scores have negative scores on PC1, and intermediate populations have mixed scores. Although *A. remigis* from some of the Albany samples appear to not follow the saw-tooth pattern of variation in size and shape, being smaller, with shorter, thicker legs relative to their body size, the individuals do not deviate in size and shape from those typical of the northern 'type'. This *a posteriori* analysis of the transition zone supports the earlier conclusion of a saw-tooth cline pattern of variation in morphology (size and shape) in eastern populations of *A. remigis*. The analysis also confirms the location of the transition between northern and southern 'types' at latitudes 40° to 42°, but extends this zone longitudinally, east to Connecticut and New Jersey.

DISCUSSION.

The preceding univariate, multivariate and *a posteriori* analyses indicate differentiation between waterstrider populations from Quebec and New York and populations from Pennsylvania, Connecticut, New Jersey, Virginia and North Carolina. The analyses reveal strong clines in body size within each of these regions, and a sharp transition in both size and shape between regions. Populations found within the transition zone between the northern and the southern 'types' of waterstrider can be identified based on an analysis of within population variance for morphological traits. This supports the hypothesis that Scranton and Albany are transition sites. However, the analysis of latitudinal variation in morphology suggests that Harrisburg may be transitional between the cline typical of southern sites and that typical of northern sites. The *a posteriori* analysis of latitudinal variation suggests that populations further east, at similar latitudes to Harrisburg and Scranton, are also transitional between northern and southern 'types'. The transition zone for morphological traits (size and shape) between the northern and southern 'types' extends further south and is therefore broader than that identified based on the analysis of allozyme variation by Gallant et al. (1993) of the same transect.

Heterozygote deficiencies found within the transition zone in Gallant et al. (1993) indicate partial reproductive isolation between northern and southern 'types'. An increase in variance in morphology within these same populations further supports the hypothesis that the gene flow between northern and southern 'types' is restricted. The partial restriction of gene flow within hybrid populations combined with limited gene flow among populations (Fairbairn, 1986; Preziosi and Fairbairn, 1992; Gallant et al., 1993) will favour the persistence of the steep cline between northern and southern 'types'. The fate of the hybrid zone and any persistence of the existing differentiation

of the northern and southern 'types' will depend on how introgression proceeds in the zone of contact. Evidence that interspecific hybridization is common in the Gerromorpha (Parshley, 1924; Calabrese, 1982; Spence and Maddison, 1986; Spence, 1990; Sperling and Spence, 1991), leads to the expectation of some gene flow, despite high levels of genetic differentiation. Post-reproductive isolating mechanisms may reduce the viability of these hybrids and influence the direction of the resulting introgression (*op cit.*), which may allow the hybrid zone to persist even in the presence of hybridization. The higher proportion of northern 'type' females than males in the *a posteriori* transition zone and the absence of northern 'type' males in two of the New Jersey populations suggests that the direction of this introgression is north to south. The absence of male northern 'types' suggests that any post-reproductive isolation may be following Haldane's rule, which states that in the offspring produced from interspecific hybridization, one sex (the heterogametic sex) is absent, rare or sterile (Haldane, 1922; Lincoln et al., 1982).

Studies in this and other gerrid species suggest some of the possible mechanisms responsible for the restriction of gene flow and the maintenance of this steep cline. The discriminant analysis classification of individuals from the transition sites in Pennsylvania, New Jersey, Connecticut, and southern New York suggest that both 'types' are present within each population. Therefore, some mechanism must be isolating, at least partially, these 'types' within the stream. Interspecific crosses between *Limnoporus dissortis* and *L. notabilis* (Spence, 1990) revealed that the F_1 hybrids had low survival to adulthood and were almost exclusively males, thus creating a partial post-mating barrier to gene flow due to asymmetrical introgression. The results of this study suggest, however, that crosses of northern and southern 'types' may result in female offspring

only. Another possible partial reproductive isolating mechanism may be behaviourial differences that somehow segregate southern 'types' from northern 'types'. This may range from passive avoidance due to habitat choice within the stream that simply reduces the probability of encountering the other 'type' of mate, to active, non-random, avoidance of the other 'type' through mechanisms such as mate choice. In *Aquarius remigis*, larger individuals within a stream have been shown to be better competitors for food as they can hold positions in the fastest flowing portions of the stream where prey capture rates are highest (Rubenstein, 1984). This may result in the segregation of larger individuals from smaller individuals within the habitat, producing a weak pattern of assortative mating. Assortative mating on the basis of total length has been shown in eastern populations of *A. remigis* (Fairbairn, 1988). Habitat segregation by size has been found in a neotropical waterstrider *Potamobates tridentatus* (Wheelwright and Wilkinson, 1985). Northern and southern 'types' differ in size and thus assortative mating by size could lead to restriction of gene flow.

Reproductive isolation may also result from phenological differences between individuals within the transition sites. Masaki (1972), and later Roff (1980), proposed a theoretical explanation for saw-tooth patterns of variation. As season length increases with decreasing latitude, so does the length of time that an animal has to develop. When the season is short, only one generation will be produced within the time constraint (univoltinism). As the season length gets progressively longer, there will be enough time for two generations to be produced that year (bivoltinism). Assuming that body size is correlated with development time, individuals will become progressively larger as season increases because they simply have more time to develop. When season length is long

enough for two generations to mature, body size decreases because development time is at its minimum. In the transition zone between univoltine and bivoltine populations, individuals from northern univoltine populations would be larger while those from southern bivoltine populations would be smaller. The resulting patterns of body size variation would be a saw-tooth cline. Based on an analysis of populations that include New Jersey sites 24 and 26 and New York sites 17 and 18, sites used in the present study, Blanckenhorn and Fairbairn (unpublished data) have shown a transition in voltinism corresponding loosely to the zone of transition. The potential mechanism of isolation may be that within the transition sites in spring, both univoltine and bivoltine individuals mate randomly. Later in the season, however, only the F_1 bivoltine southern 'types' are mating. This would lead to partial assortative mating by regional 'type', and a restriction of gene flow between regional 'types'. This explanation has its limits in explaining the partial reproductive isolation and body size variation. Firstly, we have no evidence that body size is correlated with development time within populations in this species, or in the Gerridae in general (Fairbairn, 1990). Secondly, if indeed season length dictates voltinism, then southern populations should all be bivoltine and northern populations should be univoltine. Populations as far north as Quebec have displayed partial bivoltinism (Fairbairn, 1986). This would mean that variation in body size within the transition sites may not be the result of differences in voltinism in accordance with the saw-tooth cline theory. At the present time, premating barriers due to mate choice or postmating barriers due to limited success of hybrid offspring seem more likely causes of the reduction of gene flow between these 'types'.

The southern and the northern populations are highly differentiated on the basis

of the morphometric traits measured in this study. One of the main distinguishing characters between regions is mid-femoral length, which is proportionately larger in the south. Multivariate analyses revealed that leg length was highly predictive of group membership. *Aquarius remigis* show strong hyperallometry for leg length, which explains the increase in relative size of mid-femoral length to total length. This hyperallometric relationship, both static and ontogenetic, is typical of the waterstrider genera *Gerris, Aquarius*, and *Limnoporus* (Matsuda, 1960; Andersen, 1990; Klingenberg and Zimmermann, 1992). A qualitatively similar and much stronger hyperallometry is found when comparisons are made among species of waterstriders (Klingenberg and Zimmermann, 1992). The pattern of differentiation between northern and southern 'types' of *Aquarius remigis* is therefore consistent with typical interspecific patterns of variation in this family.

Leg length may reflect adaptation to habitat type, and, more specifically, to the presence of vegetation (Spence, 1981). Individuals with short legs would be better able to manoeuvre along shorelines in the presence of emergent vegetation. Individuals with longer legs would be more suited to open water habitats where there is little or no emergent vegetation to impede movement, longer stride length and thus greater speed. Leg length may also be associated with fighting ability in intrasexual competition where the mid legs may be used as weapons (Hayashi, 1985) and may reflect possible intraspecific differences in reproductive behaviour. A study of sexual selection in Californian *A. remigis* indicates significant selection on mid-femoral length (Fairbairn and Preziosi, 1993, in press).

The concordance of morphometric variation found in the present study and the

allozyme variation found in Gallant et al. (1993), combined with evidence of partial reproductive isolation within the hybrid zone, supports the earlier suggestion of distinct northern and southern 'types' of *A. remigis* with restriction of gene flow between them. The degree of genetic differentiation between the southern and northern regions suggests re-evaluation of the taxonomic status of *A. remigis* in southeastern North America (Preziosi and Fairbairn, 1992; Gallant et al., 1993). The morphological differentiation in size and shape clearly supports these conclusions.

The pattern of morphological variation found in this study supports the hypothesis that the origin of the zone of transition and of the present genetic divergence of the northern and southern 'types' of *A. remigis* is due to a vicariance event, most likely Pleistocene isolation as suggested in Gallant et al. (1993). The location of the transition between the morphological 'types' does not correspond to any sharp change in the environment (habitat or climate: Espenshade, 1964; Cohen, 1973; Preziosi and Fairbairn, unpublished data) but rather to the location of the southern limit of Pleistocene continental glaciation (Espenshade, 1964; Cohen, 1973; Gallant et al., 1993).

If a vicariance event has taken place and secondary intergradation of the isolated populations has occurred, character sets may be expected to show concordant patterns (Barton and Hewitt, 1985; Harrison, 1990; Thorpe, 1991). Studies of *Peromyscus californicus* in California (Smith, 1979), of grasshoppers in the Pyrenees (Butlin and Hewitt, 1985), and of grass snakes in Europe (Thorpe, 1985a,b; 1987a,b) suggest that concordant variation of geographic patterns at several loci and of morphological traits is the result of a history of past isolation due to vicariance events and subsequent secondary contact. Furthermore, concordant variation in morphological variation and biochemical

genetics in *Bombina bombina* and *Bombina variegata* (fire-bellied toads) in lower Austria has been interpreted as reflecting adaptation to different environmental conditions in the Pleistocene refuges (Gollmann, 1984). Pleistocene glaciation has been invoked to explain differentiation between populations of *Aeropedellus clavatus* (tundra grasshopper) in Colorado (Hamrick and Hamrick, 1989), between populations of zooplankton in eastern lakes (Roff et al., 1981), and between populations of fishes (Seeb et al., 1987; Hinch et al., 1991). The loci observed in Gallant et al. (1993) showed no linkage disequilibria, suggesting that they were independent characters. The pattern of morphological variation is concordant with this observed variation in allozyme frequency, suggesting a similar history of past isolation due to vicariance events and subsequent secondary contact. Vicariant speciation has been suggested to be the predominant mode of speciation (Mayr, 1963; Futuyma and Mayer, 1980; Brooks and McLennan, 1991). Only 6% of 66 cases of vertebrate speciation documented in an analysis by Lynch (1989) were the result of sympatric speciation events, while 71% were the result of vicariance.

One of the major issues in speciation research is the question of whether the roles of adaptation and speciation are coupled (Brooks and McLennan, 1991, p124). Reproductive isolation is not yet complete in clines and hybrid zones of adaptive and potentially neutral traits, providing the opportunity to examine the relationship between adaptation and speciation. Though the roles of speciation and adaptation do not need to be coupled for cladistic evolution to take place, any adaptive changes would be more permanent as recombinations would be limited by the lack of gene flow (Futuyma, 1989), and thus adaptive differentiation may typically accompany the speciation event. The covariation of morphological and biochemical genetical traits in this system indicates that

the two processes, cladistic differentiation (speciation) and adaptive differentiation, are not uncoupled and it is likely that both selection and drift are responsible for the patterns of differentiation seen in the allozymes and morphological traits.

In summary, the patterns of clinal morphological differentiation indicate the presence of adaptive differentiation in *Aquarius remigis* and support previous suggestions that there are two distinct 'types' of this species in eastern North America, north and south of a geographically restricted zone of transition in Pennsylvania and southern New York. We also provide evidence that this transition extends zone eastward to include Connecticut and New Jersey. The results of this study also support earlier suggestions of restriction of gene flow between these two 'types' within the zone of transition. The pattern of covariance of allozyme frequency and morphometric traits is consistent with patterns of variation in other species hypothesized to be the result of vicariance and secondary contact. This pattern of covariation also suggests that in this system, cladistic and adaptive differentiation are coupled, which is relevant to the broader study of the origins of clines and the process of speciation.

REFERENCES.

Andersen, N. M. 1990. Phylogeny and taxonomy of water striders, genus *Aquarius* Schellenburg (Insecta, Hemiptera, Gerridae), with a new species from Australia. Steenstrupia 16:37-81.

Barton, N. H. and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology and Systematics 16:113-148.

Blanckenhorn, W. U. 1991. Life-history differences in adjacent water strider populations: phenotypic plasticity or heritable responses to stream temperature? Evolution 45:1520-1525.

Brooks, D. R. and D. A. McLennan. 1991. Phylogeny, Ecology, and Behavior. A Research Program in Comparative Biology. The University of Chicago Press, Chicago and London.

Butlin, R. K. and G. M. Hewitt. 1985. A hybrid zone between *Chorthippus parallelus* and *C. p. erythropus* (Orthoptera: Acrididae): morphological and electrophoretic characters. Biological Journal of the Linnean Society 26:269-285.

Calabrese, D. 1974. Population and subspecific variation in *Gerris remigis* Say. Entomological News 85:27-28.

————————. 1977. The habits of *Gerris* F. (Hemiptera: Gerridae) in Connecticut. Annals of the Entomological Society of America 70:977-983.

————————. 1979. Pterygomorphism in 10 nearctic species of *Gerris*. American Midland Naturalist 101:61-68.

————————. 1982. Hybridization between *Gerris alacris* Hussey and *Gerris comatus* Drake and Hottes (Hemiptera: Heteroptera: Gerridae) in nature. Proceedings of the Entomological Society of Washington 84:209-210.

Cohen, S. B. 1973. (ed) Oxford World Atlas, Oxford University Press, New York.

Comrey, A. L. 1973. A First Course in Factor Analysis. Academic Press, New York.

Confalonieri, V. A. and P. C. Colombo. 1989. Inversion polymorphism in *Trimerotropis pallidipennis* (Orthopters): clinal variation along an altitudinal gradient. Heredity 62:113-116.

Drake, C. J. and Harris, H. M. 1934. The Gerrinae of the western hemisphere. Annals of the Carnegie Museum. 23:179-240.

Endler, J. E. 1977. Geographic Variation, Speciation and Clines. Princeton University Press, Princeton.

Espenshade, E. B. 1964. (ed) Good's World Atlas, 12th Edition. Rand McNally, Chicago.

Fairbairn, D. J. 1984. Microgeographic variation in body size and development time in the waterstrider *Limnoporus notabilis*. Oecologia 61:126-133.

————————. 1985a. Comparative ecology of *Gerris remigis* (Hemiptera: Gerridae) in two habitats: a paradox of habitat choice. Canadian Journal of Zoology 63:2594-2603.

————————. 1985b. A test of the hypothesis of compensatory upstream dispersal using a stream-dwelling waterstrider, *Gerris remigis* Say. Oecologia 66:147-153.

————————. 1986. Does alary polymorphism imply dispersal polymorphism in the waterstrider, *Gerris remigis*? Ecological Entomology 11:355-368.

————————. 1988. Sexual selection for homogamy in the Gerridae: an extension of Ridley's comparative approach. Evolution 42:1212-1222.

————————. 1990. Factors influencing sexual size dimorphism in temperate waterstriders. The American Naturalist 136:61-86.

————————. 1992. The origins of allometry: size and shape polymorphism in the common waterstrider, *Gerris remigis* Say (Heteroptera, Gerridae). Biological Journal of the Linnean Society 45:167-186.

Fairbairn, D. J. and R. F. Preziosi. 1993. Sexual selection and the evolution of allometry for sexual size dimorphism in the waterstrider, *Aquarius remigis*. The American Naturalist, in press.

Futuyma, D. J. 1987. On the role of species in anagenesis. The American Naturalist 130:465-473.

————————. 1989. Speciation trends and the role of species in macroevolution. The American Naturalist 134:318-321.

Futuyma, D. J. and G. C. Mayer. 1980. Non-allopatric speciation in animals. Systematic Zoology 29:254-271.

Gailbraith, D. F. and C. H. Fernando. 1977. The life history of *Gerris remigis* (Heteroptera: Gerridae) in a small stream in southern Ontario. Canadian Entomologist 109:221-228.

Gallant, S. L., R. F. Preziosi, and D. J. Fairbairn. 1993. Clinal variation in eastern populations of the waterstrider, *Aquarius remigis*: Gradual intergradation or

discontinuity? Evolution, in press.

Gartside, D. F., M. J. Littlejohn, and G. F. Whatson. 1979. Structure and dynamics of a narrow hybrid zone between *Geocrinia laevis* and *G. victrianna* (Anura: Leptodactylidae) in southern Australia. Heredity 43:165-77.

Gollmann, G. 1984. Allozymic and morphological variation in the hybrid zone between *Bombina bombina* and *B. variegata* (Anura: Discoglossidae) in northeastern Austria. Zeitschrift fuer Zoologische Systematik und Evolutionsforschung 22:51-64.

Grant, V. 1989. The theory of speciation trends. The American Naturalist 133:604-612.

Gunnarsson, B. 1988. Body size and survival: implications for an overwintering spider. Oikos 52:274-282.

Haldane, J. B. S. 1922. Sex ratio and unisexual sterility in hybrid animals. Journal of Genetics 12:101-109.

————————. 1948. The theory of a cline. Journal of Genetics 48:277-284.

Hamrick, K. J. and J. L. Hamrick. 1989. Genetic variation within and among populations of an alpine grasshopper, *Aeropedellus clavatus*. Journal of Heredity 80:186-192.

Harrison, R. C. 1990. Hybrid zones: windows on evolutionary process. Oxford Surveys in Evolutionary Ecology 7:69-128.

Hayashi, K. 1985. Alternative mating strategies in the water strider *Gerris elongatus* (Heteroptera, Gerridae). Behavioral Ecology and Sociobiology 16:301-306.

Hinch, S. G., N. C. Collins and H. H. Harvey. 1991. Relative abundance of littoral zone fishes: biotic interactions, abiotic factors, and postglacial colonization. Ecology 72:1314-1324.

James, F. C. 1991. Complementary descriptive and experimental studies of clinal variation in birds. American Zoologist 31:694-706.

Klingenberg, C. P. and M. Zimmermann. 1992. Static, ontogenetic, and evolutionary allometry: a multivariate comparison in nine species of waterstriders. The American Naturalist 140:601-620.

Langvatn, R. and S. D. Albon. 1986. Geographic clines in body weight of Norwegian red deer: a novel explanation of Bergmann's rule? Holarctic Ecology 9.285-293.

Lincoln, R. J., G. A. Boxshall, and P. F. Clark. 1982. A Dictionary of Ecology, Evolution and Systematics. Cambridge University Press, Cambridge.

Lynch, J. D. 1989. The gauge of speciation: On the frequencies of modes of speciation. Pp 527-553 in Speciation and its Consequences, (ed) D. Otte and J. A. Endler. Sinauer Associates, Sunderland.

Mallet, J. 1986. Hybrid zones of *Heliconius* butterflies in Panama and the evolution of warning colour. Heredity 56:191-202.

Masaki, S. 1972. Climatic adaption and photoperiodic response in the band-legged ground cricket. Evolution 26:587-600.

Matsuda, R. 1960. Morphology, evolution and classification in the Gerridae (Hemiptera, Heteroptera). The University of Kansas Science Bulletin. 41:25-632.

Mayr, E. 1942. Systematics and the Origin of Species. Columbia University Press, New York.

———————. 1963. Animal Species and Evolution. Harvard University Press, Cambridge.

———————. 1970. Populations, Species, and Evolution. Harvard University Press, Cambridge.

McLauchlan, A. J. and D. F. Allen. 1987. Male mating success in Diptera: advantages of small size. Oikos 48:11-14.

McLean, E. B. 1990. Sexual dimorphism and predacious feeding habits of the waterstrider *Gerris remigis* Say (Heteroptera; Gerridae). Canadian Journal of Zoology 68:2688-2691.

McMahon, T. A. and J. T. Bonner. 1983. On Size and Life. Scientific American Books, New York.

Michel, F. A. 1961. The taxonomic value of the male genitalia of the genus *Gerris Fabricius* (Hemiptera: Heteroptera: Gerridae). Unpublished M. S. thesis, Oregon State University, Corvallis.

Millar, C. I. 1983. A steep cline in *Pinus muricata*. Evolution 37:311-319.

Moran, C., P. Wilkinson, and D. D. Shaw. 1980. Allozyme variation across a narrow hybrid zone in the grasshopper *Caledia captiva*. Heredity 44:69-81.

Parshley, H. M. 1924. An instance of miscegenation in the Gerridae (Hemiptera).

Annales de la Société Entomologique de Belgique 61:42-44.

Pielou, E. C. 1991. After the Ice Age. The Return of Life to Glaciated North America. The University of Chicago Press, Chicago.

Polhemus, J. T. and Chapman, H. C.. 1979. Family Gerridae - Water Striders, Pond Skaters, Wherrymen. Pp 58-69 in The Semiaquatic and Aquatic Hemiptera of California, (ed) A. S. Menke. Bulletin of the California Insect Survey 21.

Preziosi, R. F. and D. J. Fairbairn. 1992. Genetic population structure and levels of gene flow in the stream-dwelling waterstrider, *Aquarius (=Gerris) remigis* (Hemiptera: Gerridae). Evolution 46:430-444.

Redfern, R. 1983. The Making of a Continent. Time Books, New York.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223-225.

Roff, D. A. 1980. Optimizing development time in a seasonal environment: the 'ups and downs' of clinal variation. Oecologia 45:202-208.

Roff, J. C., W. G. Sprules, J. C. H. Carter, and M. J. Dadswell 1981. The structure of crustacean zooplankton communities in glaciated North America. Canadian Journal of Fisheries and Aquatic Sciences 38:1428-1437.

Rubenstein, D. J. 1984. Resource acquisition and alternative mating strategies in water striders. American Zoologist 24:345-353.

Searle, J. B. 1984. Three new karyotypic races of the common shrew *Sorex araneus* (Mammalia: Insectivora) and a phylogeny. Systematic Zoology 33:185-194.

Seeb, J. E., L. W. Seeb, D. W. Oates, and F. M. Utter. 1987. Genetic variation and postglacial dispersal of populations of northern pike (*Esox lucius*) in North America. Canadian Journal of Fisheries and Aquatic Sciences 44:556-561.

Selander, R. K. 1964. Speciation in wrens of the genus *Camylorhynchus*. University of California Publications in Zoology 74:1-305.

Shaw, D. D. 1981. Chromosomal hybrid zones in Orthopteroid insects. Pp 109-145 in Evolution and Speciation: Essays in Honour of M. J. D. White, (ed) W. R. Atchley, D. S. Woodruff. Cambridge University Press, Cambridge.

Smith, M. F. 1979. Geographic variation in genic and morphological characters in *Peromyscus californicus*. Journal of Mammalogy 60:705-722.

Somers, K. M. 1989. Allometry, isometry and shape in principal components analysis.

Systematic Zoology 38:169-173.

Spence, J. R. 1981. Experimental analysis of habitat selection in waterstriders (Heteroptera: Gerridae). Ecology 63:1505-1514.

————. 1990. Introgressive hybridization in Heteroptera: the example of *Linmoporus* Stål (Gerridae) species in western Canada. Canadian Journal of Zoology 68:1770-1782.

Spence, J. R. and D. R. Maddison. 1986. Chromosomes of two hybridizing species of *Limnoporus* (Heteroptera: Gerridae). Proceedings of the Entomological Society of Washington 88:502-508.

Sperling, F. H. and J. R. Spence. 1991. Structure of an asymmetric hybrid zone between two water strider species (Hemiptera: Gerridae: *Limnoporus*). Evolution 45:1370-1383.

Sumner, F. B. 1929. The analysis of a concrete case of intergradation between two subspecies. Proceedings of the National Academy of Sciences of Washington 15:110.

Szymura, J. M. and I. Farana. 1978. Inheritance and linkage analysis of five enzyme loci in interspecific hybrids of toadlets, genus *Bombina*. Biochemical Genetics 16:307-319.

Tabachnick, B. G. and L. S. Fidell. 1989. Using Multivariate Statistics, 2nd Edition. Harper Row, New York.

Tardif, B. and P. Morisset. 1990. Clinal morphological variation of *Allium schoenophrasum* in eastern North America. Taxon 39:417-429.

Thorpe, R. S. 1985a. Character number and the multivariate analysis of simple patterns of geographic variation: categorical or 'stepped-clinal' variation. Systematic Zoology 34:127-139.

————. 1985b. Clines: character number and the multivariate analysis of simple patterns of variation. Biological Journal of the Linnean Society 26:201-214.

————. 1987a. Geographic variation: a synthesis of cause, data, patterns and congruence in relation to subspecies, multivariate analyses and phylogenesis. Bulletin of Zoology 53:3-11.

————. 1987b. Congruence between independent character systems across a hybrid zone: patterns in geographic space. Zeitschrift fuer Zoologische Systematik und Evolutionsforschung 25:161-169.

———————. 1991. Clines and cause: microgeographic variation in the tenerife gecko (*Tarentola delelandii*). Systematic Zoology 40:172-187.

Turner, J. R. G. 1971. Two thousand generations of hybridization in a *Heliconius* butterfly. Evolution 25:471-482.

Wheelwright, N. T. and G. S. Wilkinson. 1985. Space use by a neotropical water strider (Hemiptera: Gerridae): sex and age-class differences. Biotropica 17:165-169.

Wilkinson, L. 1989. Systat. The system for Statistics for the PC., 2nd Edition, Systat, Inc., Evanston, Il.

Yang, S. Y. and R. K. Selander. 1968. Hybridization in the grackle *Quiscalus quiscula* in Louisiana. Systematic Zoology 17:107-143.

Zera, A. J. 1981. Genetic structure of two species of waterstriders (Gerridae: Hemiptera) with differing degrees of winglessness. Evolution 35:218-225.

APPENDIX A
Multivariate analyses

The following discussion is based on information provided in Tabachnick and Fidell (1989) on data treatment prior to principal components and discriminant analyses.

Principal components analysis:

The main objective of PCA is to reduce a large number of predictor variables to a smaller number of principal components and describe, if possible, the relationship between these variables. Certain considerations of the data set must be made to insure the power and reliability of the results, especially regarding the size of the correlations.

Large sample sizes are important as correlation coefficients become more reliable with increasing sample sizes. In this study, sample sizes were 349 for males and 340 for females, which are consid· red good (Comrey, 1973; Tabachnick and Fidell, 1989). The distributions of individual trait values for each sex deviated significantly from normality based on Lilliefors goodness-of-fit to a normal distribution tests. All but male abdominal width and female abdominal width and length and total length were significantly nonnormal. These deviations from normality are not surprising as the data were not sampled from a single population. When PCA is being used to summarize the relationship in a large set of variables, the analysis does not assume normality. The more normal the distributions are, however, the more the results of the analysis are reliable. Measures designed to make the data set normal (log transformation and removal of outliers) did not change the overall results of this analysis, suggesting that the deviations from normality do not detract from the inherent patterns within the data set revealed through the PCA. Outliers tend to have strong influences on the principal

components, distorting the underlying patterns within the data set. Univariate outliers are individual points that are 3 standard deviations above or below the mean. Multivariate outliers were determined by examining Mahalanobis distances and considering as outliers any individual point 3 standard deviations from the group centroid. In this study, the removal of the 19 male and 21 female univariate and multivariate outliers from the analysis did not change the overall results suggesting that they have little effect.

In order to get reliable indices of the relationship between the variables, certain aspects of the data set should be examined. Multivariate normality within the data set implies that relationships among variables are linear. Although the data were not normal, linearity among pairs of variables is important in order to accurately determine the relationship between these variables. Correlation measures linear relationships and does not reflect nonlinear relationships. The analysis is therefore degraded when linearity fails. By the inspection of scatterplots of pairs of variables, the relationship between each pair should appear to be linear. Out of the 21 possible combinations of the 7 morphometric variables in this data set, none show any obvious non-linearity.

Multicollinearity and singularity occur within the correlation matrix when predictor variables are highly redundant and can impede the analysis by creating an identity matrix, which cannot be reliably inverted. In PCA, however, the matrix is not inverted during calculations and therefore any multicollinearity or singularity within the data set does not affect the analysis. The data matrix should, however, include several sizable correlations in order to be factorable. At least some of the correlations should be in excess of .30. Only 1 of the 21 correlations for males (genital length with

abdominal length, 0.26) and 2 of the 21 for females (genital length with head and thoracic length, 0.212; genital length with abdominal length, 0.137) are low. All others are above 0.3 and most are above 0.7.

Based on the above information, the principal components analysis performed in this study should reliably extract the underlying relationships among the 7 morphometric variables and accurately summarize the variation in the variables.

Discriminant analysis:

Discriminant analysis is used to predict group membership from a set of predictors. This analysis is very robust to violations of the assumption of normality when deviation are due to skewness rather than outliers. Removal of the outliers from the analysis improved the classification rate only slightly. If classification rates are high, as they are in the present study, the shape of the distribution is obviously not significantly hindering the analysis, especially in this case where sample sizes are large and equal (Tabachnick and Fidell, 1989). Therefore, even though the assumption of normality has been violated, discriminant analysis can onducted with reliable results.

Some considerations should be given to the data matrices and the patterns of correlation among the variables. The variances from each group should be equal, otherwise cases tend to be overclassified into groups with greater dispersion. Homogeneity of variance-covariance matrices can be achieved by log-transformation of predictors, which equalizes the variances of the groups. Within each of the two groups, the relationship between all pairs of predictor variables should be linear. As indicated for the principal components analysis, these relationships were found to be linear. In

discriminant analysis, though, a lack of linearity is less serious because violation leads to reduced power rather than increased type I error. Multicollinearity (when variables are too highly correlated, i.e. correlation $>.90$) and singularity (perfect correlations, i.e. one variable is a combination of the others) within the data matrix occur when predictor variables are highly redundant, making matrix inversion unreliable. Singular matrices produce determinants of zero and thus cannot be used as devisors. Multicollinearity in a matrix produces determinants of zero to several decimal places and thus produces high, unstable values in the resulting inverted matrix. This condition was tested by examining tolerance levels for each of the predictors. Any predictors that have insufficient tolerances (<0.04, Tabachnick and Fidell, 1989) should be excluded. Tolerances were tested by generating a variable with a random normal distribution with a mean of 0 and unit variance and regressing all of the 7 variables against this new variable. This multiple regression analysis using the independent variables abdl, hdth, abdw, gen, ffl, ffw, mf to predict the random variable, produced high tolerance values for all the predictors. None of the tolerance levels were less than .07 and most were above 0.3 so there was no need to exclude any of the variables. All of the 7 morphometric variables could therefore be tested to determine their ability to predict group membership.

As with the principal components analysis, based on the above information, potential problems with the data set were shown to not affect the outcome or change the conclusions reached. Classification rates are high and the models significantly predict group membership, indicating the reliability if this analysis in depicting patterns within the data set.

APPENDIX B
Population means and standard deviations

Mean values (± 1 standard deviation) for each of the eight morphometric traits measured (abbreviations are listed in the text), for each population, for each sex of *Aquarius remigis*. If the population was sampled in both spring and fall, both values are given. Site numbers correspond to those given in Figure 1; s=spring-caught samples; f=fall-caught samples.

Northern males:

Site number	Trait			
	ttl	hdth	abdl	gen
1	12.446 (0.303)	5.358 (0.202)	4.666 (0.181)	2.423 (0.179)
2	11.948 (0.301)	5.238 (0.166)	4.437 (0.208)	2.275 (0.130)
3	12.416 (0.407)	5.465 (0.267)	4.598 (0.174)	2.353 (0.192)
4f	12.519 (0.372)	5.464 (0.214)	4.633 (0.165)	2.369 (0.083)
4s	12.457 (0.269)	5.456 (0.165)	4.607 (0.159)	2.394 (0.173)
5	12.875 (0.282)	5.675 (0.164)	4.932 (0.519)	2.379 (0.104)
6	12.762 (0.359)	5.539 (0.173)	4.691 (0.245)	2.458 (0.166)
7	13.668 (1.028)	5.919 (0.503)	5.152 (0.378)	2.551 (0.233)
8	14.866 (1.508)	6.480 (0.683)	5.614 (0.648)	2.653 (0.254)

Northern males, continued:

Site number	Trait			
	abdw	ffl	ffw	mf
1	2.045 (0.087)	3.763 (0.152)	0.548 (0.027)	8.902 (0.251)
2	1.787 (0.105)	3.659 (0.177)	0.538 (0.027)	8.523 (0.249)
3	1.942 (0.113)	3.729 (0.119)	0.524 (0.040)	8.904 (0.353)
4f	2.012 (0.077)	3.831 (0.149)	0.535 (0.031)	9.114 (0.342)
4s	1.938 (0.088)	3.810 (0.142)	0.557 (0.034)	9.164 (0.229)
5	2.037 (0.060)	4.019 (0.157)	0.556 (0.044)	9.359 (0.370)
6	1.771 (0.127)	4.008 (0.151)	0.530 (0.035)	9.400 (0.358)
7	2.057 (0.178)	4.158 (0.399)	0.554 (0.077)	9.844 (0.774)
8	2.182 (0.208)	4.544 (0.398)	0.580 (0.050)	11.317 (0.898)

Southern males:

Site number	Trait			
	ttl	hdth	abdl	gen
9	13.702 (0.502)	5.971 (0.242)	5.253 (0.241)	2.467 (0.172)
10	13.667 (0.769)	5.970 (0.326)	5.288 (0.457)	2.256 (0.367)
11	14.150 (0.544)	6.287 (0.296)	5.382 (0.275)	2.481 (0.137)
12f	13.517 (0.401)	6.147 (0.199)	4.752 (0.369)	2.618 (0.402)
12s	13.988 (0.811)	6.227 (0.358)	5.061 (0.357)	2.701 (0.335)
13	14.702 (0.616)	6.690 (0.333)	5.328 (0.252)	2.683 (0.139)
14	14.329 (0.445)	6.445 (0.232)	5.294 (0.136)	2.591 (0.315)
15f	14.931 (0.464)	6.711 (0.284)	5.530 (0.233)	2.690 (0.150)
15s	14.984 (0.547)	6.706 (0.272)	5.528 (0.192)	2.750 (0.184)

Southern males, continued:

Site number	Trait			
	abdw	ffl	ffw	mf
9	1.988 (0.111)	4.249 (0.171)	0.610 (0.045)	10.410 (0.524)
10	1.831 (0.132)	4.231 (0.299)	0.536 (0.058)	10.392 (0.558)
11	2.045 (0.090)	4.307 (0.198)	0.650 (0.037)	11.182 (0.579)
12f	2.151 (0.122)	4.611 (0.263)	0.633 (0.044)	11.365 (0.557)
12s	2.206 (0.115)	4.767 (0.252)	0.670 (0.049)	11.680 (0.714)
13	2.226 (0.126)	5.259 (0.342)	0.698 (0.057)	13.120 (1.073)
14	2.195 (0.122)	4.991 (0.211)	0.667 (0.036)	12.537 (0.569)
15f	2.289 (0.143)	5.273 (0.309)	0.658 (0.053)	13.702 (0.832)
15s	2.329 (0.100)	5.340 (0.252)	0.718 (0.046)	13.796 (0.922)

Northern females:

Site number	Trait			
	ttl	hdth	abdl	gen
1	14.182 (0.465)	5.895 (0.165)	7.355 (0.304)	0.932 (0.175)
2	13.535 (0.458)	5.666 (0.236)	6.970 (0.296)	0.899 (0.181)
3	13.797 (0.437)	5.894 (0.210)	7.122 (0.321)	0.780 (0.116)
4f	13.971 (0.416)	5.836 (0.204)	7.291 (0.272)	0.823 (0.125)
4s	14.179 (0.386)	5.855 (0.208)	7.419 (0.219)	0.876 (0.168)
5	14.405 (0.328)	6.124 (0.200)	7.563 (0.197)	0.672 (0.124)
6	14.328 (0.513)	5.915 (0.298)	8.453 (0.329)	0.856 (0.170)
7	15.288 (1.158)	6.351 (0.439)	7.994 (0.528)	0.906 (0.213)
8	15.964 (0.934)	6.568 (0.401)	8.372 (0.501)	0.914 (0.114)

Northern females, continued:

Site number	Trait			
	abdw	ffl	ffw	mf
1	2.270 (0.135)	3.917 (0.153)	0.464 (0.029)	9.221 (0.342)
2	2.021 (0.080)	3.810 (0.145)	0.453 (0.021)	9.065 (0.325)
3	2.044 (0.124)	3.822 (0.155)	0.426 (0.032)	9.144 (0.368)
4f	2.164 (0.098)	3.914 (0.119)	0.428 (0.034)	9.346 (0.306)
4s	2.148 (0.129)	3.903 (0.149)	0.450 (0.026)	9.429 (0.279)
5	2.174 (0.126)	4.060 (0.127)	0.472 (0.023)	9.603 (0.326)
6	1.916 (0.140)	3.999 (0.166)	0.434 (0.033)	9.680 (0.364)
7	2.177 (0.255)	4.197 (0.379)	0.456 (0.041)	10.132 (1.001)
8	2.245 (0.160)	4.268 (0.327)	0.481 (0.025)	10.708 (0.725)

Southern females:

Site number	Trait			
	ttl	hdth	abdl	gen
9	14.766 (0.435)	6.165 (0.199)	7.709 (0.269)	0.851 (0.140)
10	14.508 (0.738)	6.050 (0.490)	7.530 (0.488)	0.755 (0.172)
11	15.246 (0.408)	6.543 (0.237)	7.921 (0.219)	0.782 (0.156)
12f	15.654 (0.610)	6.516 (0.245)	7.687 (0.279)	0.877 (0.136)
12s	15.369 (0.552)	6.495 (0.336)	7.684 (0.226)	1.190 (0.257)
13	15.654 (0.610)	6.663 (0.265)	8.114 (0.321)	0.877 (0.136)
14	15.375 (0.493)	6.459 (0.229)	7.951 (0.229)	0.965 (0.218)
15f	15.782 (0.566)	6.748 (0.331)	8.174 (0.252)	0.860 (0.173)
15s	15.862 (0.493)	6.645 (0.290)	8.016 (0.234)	1.201 (0.243)

Southern females, continued:

Site number	Trait			
	abdw	ffl	ffw	mf
9	2.169 (0.092)	4.225 (0.149)	0.459 (0.034)	10.373 (0.434)
10	1.936 (0.170)	3.980 (0.213)	0.441 (0.042)	10.043 (0.447)
11	2.147 (0.116)	4.334 (0.243)	0.486 (0.029)	10.946 (0.465)
12f	2.315 (0.132)	4.700 (0.222)	0.528 (0.033)	11.261 (0.479)
12s	2.363 (0.116)	4.709 (0.227)	0.527 (0.029)	11.452 (0.676)
13	2.285 (0.176)	5.041 (0.278)	0.545 (0.054)	12.446 (0.891)
14	2.294 (0.114)	4.883 (0.144)	0.516 (0.033)	11.937 (0.399)
15f	2.381 (0.108)	5.159 (0.251)	0.535 (0.535)	12.789 (0.603)
15s	2.411 (0.104)	5.072 (0.215)	0.535 (0.061)	12.660 (0.633)

Mean values (±1 standard deviation) for each of the eight morphometric traits measured for each of the 11 *a posteriori* populations, for each sex, of *Aquarius remigis*. Site numbers (sample size) correspond to those given in Figure 3.

Males:

Site number	Trait			
	ttl	hdth	abdl	gen
16(29)	12.654 (0.486)	5.514 (0.257)	4.683 (0.210)	1.977 (0.090)
17(7)	13.610 (0.629)	5.842 (0.360)	5.226 (0.339)	2.542 (0.129)
18(5)	14.014 (0.527)	6.043 (0.329)	5.283 (0.356)	2.687 (0.141)
19(41)	14.683 (0.512)	6.377 (0.270)	5.565 (0.271)	2.741 (0.138)
20(12)	14.443 (0.293)	6.339 (0.151)	5.445 (0.204)	2.659 (0.137)
21(9)	14.665 (0.523)	6.397 (0.360)	5.698 (0.197)	2.570 (0.122)
22(26)	15.101 (0.539)	6.670 (0.332)	5.759 (0.228)	2.672 (0.160)
23(30)	13.710 (0.819)	6.066 (0.371)	5.170 (0.392)	2.474 (0.141)
24(10)	14.036 (0.355)	6.150 (0.266)	5.534 (0.289)	2.352 (0.107)
25(3)	14.179 (0.551)	6.302 (0.394)	5.492 (0.165)	2.385 (0.028)
26(10)	13.648 (0.454)	6.108 (0.175)	5.205 (0.309)	2.335 (0.221)

Males, continued:

Site number	Trait			
	abdw	ffl	ffw	mf
16(29)	1.977 (0.090)	3.879 (0.156)	0.507 (0.033)	8.940 (0.427)
17(7)	1.983 (0.099)	4.055 (0.230)	0.518 (0.047)	9.371 (0.330)
18(5)	2.465 (0.807)	4.119 (0.204)	0.552 (0.029)	9.982 (0.480)
19(41)	2.147 (0.131)	4.373 (0.233)	0.595 (0.035)	10.540 (0.519)
20(12)	2.154 (0.078)	4.397 (0.185)	0.584 (0.026)	10.607 (0.341)
21(9)	2.194 (0.105)	4.557 (0.221)	0.619 (0.049)	11.015 (0.604)
22(26)	2.226 (0.108)	4.544 (0.211)	0.631 (0.044)	10.998 (0.488)
23(30)	2.016 (0.100)	4.249 (0.320)	0.577 (0.054)	10.294 (0.824)
24(10)	1.960 (0.066)	4.281 (0.203)	0.566 (0.045)	10.354 (0.366)
25(3)	1.913 (0.037)	4.376 (0.139)	0.588 (0.039)	10.915 (0.304)
26(10)	1.923 (0.114)	4.129 (0.189)	0.562 (0.043)	10.041 (0.439)

Females:

Site number	Trait			
	ttl	hdth	abdl	gen
16(29)	14.072 (0.536)	5.873 (0.206)	7.332 (0.335)	0.867 (0.109)
17(7)	14.358 (0.231)	6.088 (0.197)	7.479 (0.163)	0.791 (0.070)
18(8)	14.718 (0.503)	6.239 (0.260)	7.633 (0.293)	0.846 (0.064)
19(33)	15.775 (0.424)	6.562 (0.184)	8.298 (0.283)	0.915 (0.106)
20(6)	15.928 (0.232)	6.647 (0.090)	8.390 (0.203)	0.890 (0.071)
21(4)	15.451 (0.972)	6.376 (0.272)	8.228 (0.672)	0.847 (0.117)
22(15)	16.055 (0.469)	6.773 (0.246)	8.370 (0.303)	0.911 (0.136)
23(26)	14.687 (0.718)	6.185 (0.318)	7.660 (0.440)	0.842 (0.101)
24(9)	15.045 (0.459)	6.366 (0.293)	7.742 (0.272)	0.937 (0.235)
25(4)	15.333 (0.565)	6.316 (0.269)	8.033 (0.224)	0.984 (0.111)
26(11)	15.080 (0.554)	6.363 (0.270)	7.863 (0.362)	0.855 (0.209)

Females, continued:

Site number	Trait			
	abdw	ffl	ffw	mf
16(29)	2.177 (0.103)	3.789 (0.162)	0.422 (0.035)	9.027 (0.336)
17(7)	2.048 (0.071)	4.021 (0.118)	0.441 (0.034)	9.547 (0.425)
18(8)	2.251 (0.122)	4.097 (0.160)	0.433 (0.043)	9.870 (0.425)
19(33)	2.367 (0.110)	4.282 (0.166)	0.492 (0.029)	10.422 (0.490)
20(6)	2.307 (0.078)	4.414 (0.106)	0.495 (0.022)	10.828 (0.342)
21(4)	2.285 (0.143)	4.295 (0.234)	0.492 (0.059)	10.445 (0.468)
22(15)	2.435 (0.126)	4.330 (0.200)	0.499 (0.039)	10.668 (0.352)
23(26)	2.169 (0.119)	4.083 (0.276)	0.452 (0.046)	9.730 (0.775)
24(9)	2.133 (0.148)	4.122 (0.138)	0.456 (0.032)	10.249 (0.287)
25(4)	2.140 (0.062)	4.249 (0.205)	0.467 (0.047)	10.132 (0.418)
26(11)	1.998 (0.128)	4.069 (0.141)	0.459 (0.021)	10.033 (0.312)