Gene Flow and Dispersal Among Populations of the Stonefly *Peltoperla tarteri* (Plecoptera: Peltoperlidae) in the Southern Appalachians

Alicia S. Schultheis

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment for the requirements for the degree of

Doctor of Philosophy in Biology

Albert C. Hendricks, Chair
Ernest F. Benfield
Bruce J. Turner
Reese Voshell
Lee Weigt
David West

November 14, 2000
Blacksburg, Virginia

Key words: Plecoptera, dispersal, control region, population genetics, nested clade analysis, phylogeography

Copyright 2000, Alicia S. Schultheis
Gene Flow and Dispersal Among Populations of the Stonefly *Peltoperla tarteri* (Plecoptera: Peltoperlidae) in the Southern Appalachians

**Alicia S. Schultheis**

**(ABSTRACT)**

Despite a number of recent studies focused on the issue, patterns of stream insect dispersal in temperate streams of North America remain poorly understood. Movement of benthic invertebrates is notoriously difficult to measure directly using traditional means; however, genetic markers provide an ideal method for estimating stream insect dispersal. In this study, the control region of mitochondrial DNA was used to study gene flow and dispersal among populations of the stonefly *Peltoperla tarteri* in the Southern Appalachians. The control region of *P. tarteri* is approximately 1270 base pairs (bp) in length, 81% AT-rich, and contains variable numbers of a 74 bp tandem repeat containing the sequence motif ‘5’-GGGGGC-3’.

Many stream insects have long life cycles that result in the simultaneous existence of multiple cohorts throughout the larval development period. If larval development is fixed, genetic isolation among cohorts may confound genetic estimates of dispersal. Although larval head width data indicated that *P. tarteri* is semivoltine in southwestern Virginia, low levels of genetic differentiation among cohorts suggested that larval development of *P. tarteri* is not fixed and that some individuals complete development in one or three years rather than two. This ‘cohort splitting’ would result in individuals from distinct cohorts maturing at the same time and mating with one another. Such developmental plasticity may allow some individuals to avoid adverse environmental conditions.

The extent and likely mechanism of dispersal was determined by comparing levels of population genetic differentiation within drainages to that among drainages. While there was no consistent relationship between genetic and geographic distance, genetic differentiation was generally high within and among drainages. Nested clade analysis indicates that historical fragmentation and range expansion coupled with contemporary gene flow explain the present day pattern of genetic variation in *P. tarteri*. In order for these historical patterns to have such a
strong influence on present day genetic structure, both larval and adult dispersal must be restricted. However, the genetic data suggest that larval dispersal is more frequent than adult dispersal.
GRANT INFORMATION

This work was supported by the Sigma Xi Grants-in-aid-of-research Program, the Virginia Academy of Science, and the Graduate Research Development Project at Virginia Tech. In addition, the Department of Biology at Virginia Tech supplied matching funds for all grants up to $500.
ACKNOWLEDGEMENTS

Support and encouragement from many people made this work possible. Foremost, I thank my committee chair, Dr. Albert Hendricks, for guidance and allowing me to work independently. In no particular order, I thank Dr. Fred Benfield for advice on a full range of topics and for keeping an extra chair in his office. Dr. Bruce Turner allowed me access to his laboratory and endured numerous discussions during the ‘troubleshooting’ phase of my project. Dr. Reese Voshell assisted with the life history analysis and allowed members of his laboratory to aid in species identification. Dr. David West is the ultimate editor, and I appreciate his careful review of my work. Last, but certainly not least, I thank Lee Weigt whose expertise ‘jump started’ (pun intended) my laboratory work and ultimately brought this project to fruition.

I have worked closely with several excellent teachers during my tenure at Virginia Tech. I could not have asked for better mentors than Drs. Art Buikema, Jack Webster, and George Simmons. I strive to reflect their excellent examples in my own teaching.

I have been fortunate to receive unwavering support from the administration and staff of the Biology Department. I thank the Department Head, Dr. Joe Cowles, for financial support and a plethora of teaching opportunities. I also thank Dee Lee, Sue Rasmussen, Michelle Woodell, and Melanie Huffman for their assistance.

Graduate school would not have been the same without Dr. Jason Bond. I thank him for sharing his experience and knowledge with me and for being a good friend. Other graduate students who have provided assistance include Barbara Bennett, Matt McTammany, Michael Fisher, Stacey Smith, and the Virginia Tech Stream Team.

A wide circle of family and friends provided me with the experience, strength, and hope needed to get to the starting block. In many ways, this degree belongs to all of them. Kimberly Richards-Thomas, Jill Alcorn, Lynn Coyle and K.C. Cardos have been the best support system I could have asked for. I look forward to many more years of friendship with them. My family has provided support and assistance in every way imaginable. I thank my parents and my sisters, Michele Slater and Sheri Gibson, for all they have done along the way.

Finally, I thank my soul mate and partner, John Schultheis, for everything.
# TABLE OF CONTENTS

Abstract ........................................................................................................................................... ii  
Grant Information ........................................................................................................................... iv  
Acknowledgments .......................................................................................................................... v  
Table of Contents ......................................................................................................................... vi  
List of Tables .................................................................................................................................. vii  
List of Figures .............................................................................................................................. viii  
Chapter 1: Introduction .................................................................................................................1  
    Introduction ................................................................................................................................. 12  
    Methods ..................................................................................................................................... 13  
    Results and Discussion ........................................................................................................... 16  
Chapter 2: Arrangement, structural conservation, and population genetic utility...  
    Introduction ................................................................................................................................. 12  
    Methods ..................................................................................................................................... 13  
    Results and Discussion ........................................................................................................... 16  
Chapter 3: Genetic evidence for ‘leaky’ cohorts in the semivoltine stonefly...  
    Introduction ................................................................................................................................. 24  
    Methods ..................................................................................................................................... 27  
    Results ....................................................................................................................................... 29  
    Discussion ................................................................................................................................. 36  
Chapter 4: Gene flow, dispersal, and nested clade analysis among populations...  
    Introduction ................................................................................................................................. 40  
    Methods ..................................................................................................................................... 43  
    Results ....................................................................................................................................... 46  
    Discussion ................................................................................................................................. 57  
Chapter 5: Conclusions ..................................................................................................................63  
Literature Cited ............................................................................................................................ 64  
Vita ................................................................................................................................................ 73
LIST OF TABLES

Table 2.1: Primers used for PCR and sequencing of control region.................................15
Table 2.2: Summary comparison of *P. tarteri* and *P. arcuata* nucleotide .........................19
Table 3.1: Variable sites for 13 haplotypes of the control region from............................33
Table 3.2: Haplotype frequencies by cohort .....................................................................34
Table 3.3: Partitioning of genetic variance according to AMOVA ...................................35
Table 3.4: Above diagonal: Pairwise non-differentiation exact P-values..............................36
Table 4.1: Variable sites for 21 haplotypes of the control region ....................................47
Table 4.2: Haplotype frequencies by site .........................................................................48
Table 4.3: Partitioning of genetic variance according to AMOVA ...................................49
Table 4.4: Above diagonal: Pairwise F_{ST} values .............................................................50
Table 4.5: Haplotype frequencies used in the nested clade analysis ...............................54
Table 4.6: Results of GeoDis analysis showing clade (D_c), nested (D_n) .......................57
LIST OF FIGURES

Figure 2.1: Schematic of primers used to amplify and sequence .............................................15
Figure 2.2: Schematic of control regions of P. tarteri (A) and P. arcuata (B) .........................16
Figure 2.3: Aligned sequences (5' to 3') for complete Peltoperla tarteri ..............................18
Figure 2.4: Alignment of ‘Conserved Block 1’ sequences in peltoperlid stoneflies ...............20
Figure 2.5: Alignment of ‘Conserved Block 2’ sequences in peltoperlid stoneflies .............20
Figure 2.6: Possible conserved secondary structure in peltoperlid stoneflies ......................23
Figure 3.1: Study sites in Giles, County, Virginia ..............................................................26
Figure 3.2: Mean monthly temperature in Sartain Branch ....................................................30
Figure 3.3: Growth phenology of Peltoperla tarteri at Hunters Branch .............................31
Figure 3.4: Phenology of adult emergence in Hunters Branch, Sartain Branch, ..........32
Figure 4.1: Study sites in Giles, County, Virginia .............................................................43
Figure 4.2: Neighbor-joining tree indicating genetic differentiation ....................................51
Figure 4.3: Pairwise Fst’s vs. aerial and aquatic distances (km) .............................................52
Figure 4.4: Unrooted phylogram indicating strict consensus of 654 .....................................53
Figure 4.5: Haplotype network showing 95% most probable connections .......................55
Figure 4.6: Nested clade diagram of haplotype network ..................................................56
Chapter 1: Introduction

Background

Animal movement is one of the most difficult problems to study in natural populations. Population-level studies often treat local populations as closed systems and ignore the effects of the gain or loss of individuals due to emigration or dispersal. However, the movement of individuals between populations can have significant effects (e.g., on population genetic structure) and, with the current rate of loss of biodiversity worldwide, understanding the dispersal dynamics of natural populations is more important than ever. Of particular concern is increasing habitat fragmentation, which in the absence of dispersal can lead to small effective population sizes and consequent loss of genetic diversity. In some cases, dispersal patterns may determine the persistence of populations and, ultimately, the survival of a species.

Most animal species studied to date move far less than is physically possible (Slatkin 1985). That the extent of dispersal depends upon more than an animal’s ability to move (vagility) is not surprising due to the considerable risks involved with dispersing. These risks include being unable to locate food and/or avoid predators en route, or not finding suitable habitat and/or a mate upon reaching a final destination. Dispersal behaviors are therefore likely to be influenced by strong selection pressures and may reflect species-specific reproductive strategies. However, numerous other factors, both biotic (e.g., vagility and life span) and abiotic (e.g., topography, height and density of vegetation, moisture, wind and availability of suitable habitat), undoubtedly affect animal dispersal as well.

Most stream insects alternate between a short-lived winged adult stage and a longer aquatic immature stage (Merritt and Cummins 1996). There are two primary ways stream insects can disperse: 1) larval dispersal, which is aquatic and includes swimming, crawling, and downstream displacement by the stream current (drift) and, 2) aerial dispersal by winged adults.

Numerous studies have shown that large numbers of stream insect larvae are carried downstream by drift (Allan 1987, Armitage 1977, Berner 1951, Townsend and Hildrew 1976, and Waters 1965). Though there has been some discussion about whether insects enter the drift actively or passively, it is widely held that larval drift is the primary dispersal mechanism for stream insects (Palmer et al. 1996). In part, this is due to the paucity of data concerning adult dispersal.
The first studies of adult movement used conventional techniques (e.g., nets and traps) to focus on Müller’s ‘colonization cycle’. After observing the upstream flight of adult stream insects, Müller (1982) proposed that gravid females fly upstream to oviposit, and thereby compensate for the downstream displacement of larvae by the stream current. In a few cases, larval drift rate seemed to affect the frequency of upstream flight and to support the idea of a colonization cycle for stream insects (Flecker and Allan 1988, Neves 1979). However, several other studies found that relatively few species exhibit classic ‘compensatory upstream flight’ (Bird and Hynes 1981, Madsen and Butz 1976, Svensson 1972, and Svensson 1974), suggesting that the pattern of adult movement is more complicated than described by Müller.

The movement of adults away from the stream channel (inland dispersal) has proved similarly difficult to characterize using conventional methods. For example, although overall abundance and biomass of 27 Trichopteran taxa decreased with increasing distance from the stream (Jackson and Resh 1989) and dispersal distance increased with increasing body size for several taxa (Kovats et al. 1996), adults of some species were always found close to the stream while others were captured up to 8 km away (Sode and Wiberg-Larsen 1993). The overall results are inconclusive with regard to movement patterns of adult stream insects, and, consequently, the relative importance of larval and adult dispersal in maintenance of benthic populations has been a source of continuing debate (Williams and Williams 1993, and Wilzbach and Cummins 1989).

Until recently, resolving this debate has been difficult because of several factors that confound direct measurement of stream insect dispersal. Stream insects tend to have small body sizes, reclusive behavior, and a brief adult stage--all of which make direct observation or measurement of their behaviors difficult. In addition, there are problems associated with conventional methods used to make direct estimates of dispersal.

Direct methods of estimating dispersal include the use of attracting (e.g. light) traps, non-attracting traps (e.g. drift nets, Malaise traps), and/or mark-recapture methods. One of the major problems associated with using traps is trap size, which limits the number of individuals that are captured. Furthermore, good swimmers (e.g. some Ephemeroptera) and flyers (e.g., Trichoptera and Odonata) can easily avoid small traps (Bird and Hynes 1981). On the other hand, some species may not be caught in traps for reasons unrelated to dispersal behavior. For example, mating swarms of *Brachyptera risi* have been observed over 50 m from the stream and often at
considerable heights (Madsen and Butz 1976); therefore, even large traps spanning the stream would be ineffective for sampling *B. risi*. Other species are known to fly close to the stream surface; thus, traps must stretch all the way down to the stream surface in order to capture them (Bird and Hynes 1981).

Since the force and direction of wind impacts flight direction, the position of traps relative to local air movements is also important. Improperly placed traps can create local eddies of air or lead to avoidance reactions (Bird and Hynes 1981). Trap location can also influence flight direction. Flecker and Allan (1988) observed that the mayfly, *R. hageni*, consistently flew towards clearings and/or downstream. Therefore, conclusions about flight directionality and dispersal in this species may depend on where traps are placed relative to clearings. Finally, light intensities can vary on different sides of traps (Bird and Hynes 1981) and this may affect results because swarming often occurs in sunlight (Madsen and Butz 1976).

Perhaps the most serious problem associated with conventional ‘capture’ methods is that they do not actually measure dispersal. Because the geographic origin of individuals captured in traps or nets is unknown, these methods give an estimate of the number of individuals moving past a fixed point rather than a true estimate of dispersal. Mark-recapture methods give a better estimate of dispersal; yet, they are limited to short-range movements and often fail to detect low-frequency long distance dispersal (Stenseth and Lidicker 1992). Therefore, the relative importance of adult and larval dispersal in the maintenance of benthic populations cannot be reliably established with either capture or mark-recapture techniques.

Fortunately, advances in molecular biology in the past quarter century have made it possible to use genetic information to address questions concerning dispersal that cannot be fully evaluated by conventional means. These ‘indirect’ measures of dispersal use the relationship between dispersal and population genetic structure to infer patterns of movement between populations (Slatkin 1985). If stream insects disperse widely, little genetic differentiation among populations would be expected. Conversely, if they do not disperse far from their natal stream, the populations may become genetically distinct due to the effects of genetic drift (random fixation or loss of alleles) and/or natural selection.

In the case of stream insects, the genetic approach is particularly useful because it can be used to determine both the *extent* and likely *mechanism* of dispersal. Because the surrounding terrestrial environment presents a formidable barrier to the dispersal of aquatic larvae, adult
flight is the primary mechanism for dispersal between streams located in separate watersheds. Thus, while the extent of dispersal between any two populations can be inferred from the level of genetic differentiation between them, the likely mechanism can be determined by using a hierarchical approach that compares levels of genetic differentiation between populations in the same drainage (where both larval and adult dispersal is possible) to those in separate drainages (where dispersal by larvae is extremely unlikely).

Genetic data provide several other advantages as well. Genetic markers can be used to screen large numbers of individuals and over very large spatial scales (Avise 1994). In addition, because very little gene flow (>1 migrant every other generation) is required to counteract the effects of genetic drift, genetic data can detect low-frequency long distance dispersal when appropriate sample sizes are used (Slatkin 1985). However, the genetic approach is not a panacea and there are several problems associated with genetic methods. Chief among them are the many assumptions involved in calculating population genetic parameters such as Wright’s $F$ statistics (Rousset and Raymond 1997), the fact that only dispersal that results in successful mating will be detected, and that factors other than vagility and dispersal can affect population genetic structure.

For instance, behaviors like habitat choice and philopatry (site faithfulness) can affect gene flow in ways that are not related to the ability of adults to disperse. Anadromous salmon and birds are two good examples of animals with high dispersal potential that often return to natal sites to mate (Avise 1994). Natal site philopatry in aquatic insects would lead to highly structured populations, perhaps even within streams.

Environmental factors such as temperature may ultimately lead to the genetic structuring of populations. Insect development is strongly affected by temperature. Hence, the timing of adult emergence and flight period are also influenced by temperature, and populations located in areas that receive different amounts of thermal input may emerge at different times. If the differences in emergence times are large enough, the flight periods of two populations may overlap only slightly or not at all, as is the case for some mayflies in the Colorado Rockies (Peckarsky et al. 2000). In such cases, the decrease in or elimination of gene flow between these populations would be a result of time serving as a genetic isolating mechanism (rather than geographic distance and/or dispersal ability). Because population genetic structure can be affected by factors other than dispersal, it is important to use information from genetic markers in conjunction with ecological and life history data.
Early studies of the population genetic structure of North American stream insects highlighted the importance of life history information in interpreting genetic data. For example, variation in life history helped explain the contrasting population genetic structures of two waterstrider species (*Gerris remigis* and *Limnoporus canaliculatus*, Hemiptera: Gerridae) with differing degrees of winglessness. The two species differed markedly, both in degree of population differentiation at polymorphic loci and in amounts of variability within populations (Zera 1981). Average heterozygosity values for *G. remigis*, the wingless species, were lower than is typical for other insects or invertebrates and were similar to values for species inhabiting islands, suggesting that gene flow was severely reduced in this species.

In contrast, allelic frequencies were spatially homogeneous for the winged species, *Limnoporus canaliculatus* (Zera 1981). Dispersal is common in this species, and some adults have been found in small pools on top of 3-story buildings (Calabrese 1974, as cited by Zera 1981). The importance of dispersal via adult flight was illustrated by highly divergent allelic frequencies in one population of *L. canaliculatus* that was composed of 94% wingless individuals.

While differences in life history were good predictors of population genetic structure in waterstriders, this was not the case for two species of blackflies (Diptera). *Prosimulium fuscum* may complete its first ovarian cycle without a blood meal, while *P. mixtum* requires a blood meal to complete first ovarian cycle. Therefore, it was predicted that populations of *P. fuscum* would be isolated relative to populations of *P. mixtum* since *P. mixtum* would have to search for the required blood meal (Snyder and Linton 1984). The results were exactly the opposite of what had been predicted: populations of *P. fuscum* were panmictic, while populations of *P. mixtum* were structured. Snyder and Linton (1984) suggested that the apparent population structure they found in *P. mixtum* might reflect the inclusion of more than one species in their *P. mixtum*.

Such taxonomic issues are a major concern in population genetic studies of stream insects. In most cases, adult males are required for positive species identification. It is therefore easy to include more than one species when several closely related species are present and larvae are used as source material.

Also, there is mounting evidence of significant genetic differentiation in the absence of morphological change in many invertebrates (Bond *in press*, Duan et al. 2000). Several studies of stream invertebrates have identified genetic groups that may constitute cryptic species.
Genetic distances among presumptive populations of *Hyallela azteca* were higher than genetic distances among *H. azteca* and *H. montezuma*, a recognized congener (Duan et al. 2000). Likewise, the presence of four genetically distinct groups among populations of the caddisfly *Helicopsyche borealis* in three eastern North American and three north-eastern Californian streams led Jackson and Resh (1992) to suggest that the groups were reproductively isolated species rather than genetic variants of a single species.

In summary, though molecular techniques circumvent some of the problems associated with conventional methods of directly estimating dispersal, they must be used in conjunction with accurate taxonomic information and life history data. Yet, even with this information it can be difficult to determine what causes the genetic structuring of populations.

In general, studies of North American stream insects have suggested that adult flight is not an effective mechanism for dispersal, because populations located in the same drainage tend to have lower levels of differentiation than those located in separate drainages. For example, levels of genetic differentiation in two mayfly species from different drainages were intermediate to high with respect to other insects (average $F_{ST} = 0.04-0.15$) and an order of magnitude higher than those of very mobile insects characterized by high levels of gene flow (Sweeney et al. 1987).

Restricted gene flow at distances of 26-2300 km was attributed to the short life span of adults (1 - 72h). However, large population sizes led Sweeney et al. (1987) to doubt that enough inbreeding was occurring to counteract the effects of gene flow, and they suggested that small effective population sizes coupled with patchily distributed populations were the cause of the genetic structure.

Strikingly different levels of genetic differentiation among populations of two other closely related mayfly species suggested that dispersal and gene flow were high within drainages in one case, (i.e., *Eurylophella verisimilis* (Ephemeroptera: Ephemerellidae), but not in the other (*Ephemerella subvaria*) (Sweeney et al. 1986). Life history data would have been very useful in understanding the different population structures of the two species.

Unfortunately, the design of many of the early genetic studies made it impossible to conclude much about larval dispersal. Those studies did not include any sites that were connected by water (Zera 1981), or combined all sites along a stream (Snyder and Linton 1984), or included only one or two distant sites on each stream (Funk et al. 1988, Jackson and Resh...
1992, Sweeney et al. 1987). Nevertheless, when a hierarchical approach is used, genetic information can make important contributions to the debate about the relative importance of adult and larval dispersal in the maintenance of stream insect populations.

For example, in the waterstrider, *Gerris remigis* (Hemiptera: Gerridae), very high levels of genetic differentiation (average $F_{ST} = 0.54$) at spatial scales greater than sites within streams suggested that in-stream movement was an effective dispersal mechanism and that there was little between-stream movement, which is not surprising for a species in which most adults are wingless (Preziosi and Fairburn 1992).

Interestingly, Hughes et al. (1999) found a similar pattern of genetic structure in the stonefly *Yoraperla brevis* (Plecoptera: Peltoperlidae) despite the presence of a winged adult stage and a relatively synchronous extended emergence period. The high levels of genetic differentiation among streams (average $F_{ST} = 0.156$), low levels within streams (average $F_{ST} \approx 0$) and consistent levels of heterozygosity upstream and downstream suggest that although dispersal is by flying adults, it is confined to their natal stream (Hughes et al. 1999). Clearly, further research is needed to determine the dominant pattern of stream insect dispersal in streams of North America.

By comparison, much more is known about the dispersal and population genetic structure of stream invertebrates in tropical streams of Australia. The evidence from Australian streams suggests that adult flight is an important dispersal mechanism for stream insects while larval movement in streams is restricted (Bunn and Hughes 1997), exactly the opposite of what was predicted by results from North American streams.

In addition, population genetic studies of Australian stream insects have revealed an unusual pattern of fine-scale genetic variation. That is, among-stream differentiation was at least as great as that between groups of small streams (subcatchments) or entire drainages, and levels of differentiation were even higher at the site or ‘reach’ scale (Hughes et al. 1998, Gallagher 1995, and Schmidt et al. 1995). At the reach scale, a number of samples were not in Hardy-Weinberg equilibrium, and this was not consistent across sites or loci for any of the species. Thus, the larval population in a reach was composed of ‘patches’ or groups of individuals that were as genetically different from each other as they were from patches in other streams and or drainages.
Based on these observations, Bunn and Hughes (1997) proposed that this pattern of genetic variation was a result of widespread adult dispersal coupled with a small number of females ovipositing (oviposition = egg laying) and repopulating a given reach. According to this ‘patchy recruitment’ hypothesis (recruitment = establishment of new cohort through oviposition and hatching), adult dispersal accounts for lower levels of genetic differentiation at the catchment or drainage level, while patchy oviposition explains high levels of genetic differentiation at the subcatchment and reach scales as well as the lack of Hardy-Weinberg equilibrium. A further assumption of this hypothesis is that in order for these small-scale genetic differences to be maintained, larval movement within streams must be limited. Though this idea contradicts the widely held view that larval dispersal is widespread and adult flight is not an effective means for dispersal, there are several features unique to tropical streams of eastern Australia that may give rise to the strong recruitment effects and limited larval movement suggested by the patchy recruitment hypothesis (Bunn and Hughes 1997).

First, tropical climates tend to be less predictable than those in temperate regions. As a result, life cycles in tropical streams tend to be asynchronous and adult emergence is ‘aseasonal’ or occurs throughout the year (Lake et al. 1986). Because only a few females oviposit at a given time when adults emerge asynchronously, larval density tends to be low in tropical streams. Consequently, distinct genetic patches are likely to be more pronounced in tropical streams than in temperate streams where many females oviposit at the same time and larvae are abundant. Second, larval drift rates in streams of southeast Queensland are much lower than those in temperate streams (Kerby et al. 1995) and the physical features of these streams (such as waterfalls, high gradients, and low flow for much of the year) are likely to impede larval movement (Bunn and Hughes 1997).

Though more work needs to be done to confirm the predictions of the patchy recruitment hypothesis, these studies have identified the importance of recruitment dynamics in determining population genetic structure of Australian stream insects. Of course, differences in climate and topography would lead to different predictions about recruitment effects in North American streams. However, before we can establish the underlying causes of population genetic structure in North American stream insects, we must first determine the patterns of genetic structure.
Focus of current work

In the Appalachian Mountains of eastern North America, headwater streams are being lost at an alarming rate as a result of human activities such as mountaintop removal mining, agriculture, and residential development (Meyer and Wallace in press). Because the diversity of many cold-water organisms (e.g., stoneflies) in Appalachian streams is among the highest in North America (Stewart and Stark 1993), the number of species lost due to the harmful effects of human activities on these streams could be significant (Riccardi and Rasmussen 1999). In those species that are not lost entirely, habitat fragmentation can lead to small effective population sizes and consequent loss of genetic diversity. Thus, it is now more important than ever to understand the dispersal dynamics of headwater stream organisms.

Peltoperla tarteri (Plecoptera: Peltoperlidae) is a stonefly that can be abundant in Appalachian headwater streams (Stark and Kondratieff 1987). Its diet consists chiefly of coarse particulate organic matter, such as leaves, which have fallen into the stream from the surrounding area (Ruggles and Tarter 1991). *P. tarteri* 'shreds' leaves into finer particles that are consumed by many other stream invertebrates. In heavily shaded streams where primary production is limited (as is the case for most Appalachian headwater streams), the shredding action of *P. tarteri* and other invertebrates initiates energy flow in the ecosystem. This role in energy transformation makes *P. tarteri* essential to ecosystem function (Cummins et al. 1989).

The life history and geographic distribution of *P. tarteri* make it an ideal organism for studying dispersal of stream insects. *P. tarteri* is hemimetabolous, lacking a pupal stage, and except for wings on adults, larvae and adults are similar in appearance. Despite appearances, adults and larvae occupy very different habitats; *P. tarteri* larvae are aquatic and restricted to cool springs, seeps and streams at higher elevations, while adults are aerial (Merritt and Cummins 1996). Thus, larval dispersal is limited both within and between drainages because it requires movement through unsuitable habitats (warmer, high flow valley streams or terrestrial environments). Therefore, the natural distribution of *P. tarteri* allows one to choose populations where adult flight is the most likely mechanism for dispersal among them.

The geography of southwestern Virginia, in particular, provides a unique opportunity to study dispersal of *P. tarteri*. Because the Eastern Continental Divide is located in this area, several distinct drainages occur together in a relatively small geographic area (i.e., within 15 km of each other): the James and Roanoke drainages, which ultimately flow to the Atlantic Ocean,
and the drainages of the New River, which ultimately flow to the Gulf of Mexico. Because there are significant physical barriers to larval dispersal between these drainages, gene flow between these populations can be attributed to adult dispersal.

My primary research goal was to determine the extent and likely mechanism for dispersal in the ecologically important stonefly species *Peltoperla tarteri*. My specific objectives were: 1) to use traditional ecological methods to study the life history of several populations of *P. tarteri* located in separate drainages in southwestern Virginia; 2) to ascertain the population genetic structure of these populations using mitochondrial DNA sequence data; and 3) to use life history and population genetic data to infer dispersal patterns of *P. tarteri*.

**Analytical approach**

All previously mentioned studies of stream insect dispersal have used protein electrophoresis or ‘allozymes’ to determine the genetic structure of stream insect populations. The allozyme technique takes advantage of the fact that non-denatured proteins with different net charges and structural conformations travel at different rates through supporting media (e.g., starch) in the presence of an electric field. Histochemical stains are used to single out products of particular genes encoded by nuclear DNA (nDNA), and the resulting banding pattern can usually be interpreted in simple genetic terms (Avise 1994). Protein electrophoresis has been widely used because it is inexpensive and requires little technical expertise (Hillis et al. 1996).

Allozymes are functional products of nuclear genes that are biparentally inherited. As such, they undergo recombination, can be subject to selection, and tend to be less variable than non-coding regions of DNA (Mitton 1994, Niegel 1997, and Parker et al. 1998). Thus, allozymes often do not provide the level of genetic variability that is required for intraspecific applications (e.g., studies of fine-scale dispersal). In contrast, animal mitochondrial DNA (mtDNA) is haploid and, in most cases, uniparentally inherited (usually maternally). This difference in ploidy and mode of inheritance results in a four-fold reduction of the ‘effective population size’ (N_e, or the genetic size of a population) of mtDNA relative to nDNA. The smaller N_e makes mtDNA more subject to the effects of genetic drift and ultimately more variable than many nuclear markers such as allozymes (Avise 1994). In many situations, mtDNA markers have revealed significant variation in the absence of allozyme variation, demonstrating the limited resolving power of protein electrophoresis (DeSalle et al. 1987).
One of the most variable regions in both vertebrates and invertebrates is the ‘control’ region of mtDNA (‘d-loop’ in vertebrates) (Simon 1991). Because the d-loop is ideal for assessing levels of genetic differentiation at the population level, it has been widely used as a population-level marker in vertebrates. However, because length variation and high A-T content often complicate amplification and sequencing of this region, the control region has been characterized in only a few insect groups (Caccone et al. 1996, Clary and Wolstenholme 1987, Crozier and Crozier 1993, Rand and Harrison 1989, and Zhang et al. 1995) despite its potential usefulness for studying gene flow and dispersal.

As part of this work, I refined a technique that allowed me to amplify and sequence the control region in *Peltoperla tarteri*. I have used the information from this region to infer patterns of gene flow and dispersal in several populations of this stream insect in the Appalachian Mountains of southwestern Virginia (Giles Co.). The four chapters that follow describe the results of this work and my conclusions about the life history, gene flow, and dispersal of this Appalachian stonefly species.  

**Chapter summary**

Chapter Two entitled, “Arrangement, structural conservation, and population genetic utility of the mitochondrial control regions of peltoperlid stoneflies”, characterizes the arrangement of the control region in two peltoperlid stoneflies and compares it to that of other insect groups. The utility of this region for insect population genetic studies is also discussed.

Chapter Three, “Genetic evidence for ‘leaky’ cohorts in the semivoltine stonefly *Peltoperla tarteri*” presents life history and genetic data, including growth and development, emergence phenology, and examines the genetic relationship between successive cohorts.

Chapter Four, “Gene flow, dispersal, and nested clade analysis among populations of the stonefly *Peltoperla tarteri* (Plecoptera: Peltoperlidae) in the Southern Appalachians” presents the results of the molecular study of gene flow and dispersal and discusses the extent and likely mechanism for dispersal in this species. The relative importance of current and historical events in forming present-day patterns of genetic structure is also discussed.

Chapters 2-4 will be submitted individually to peer-reviewed journals. Thus, each contains a separate methods section. The overall conclusions are presented in Chapter Five, which is followed by a collection of references used in writing the dissertation.
Chapter 2: Arrangement, structural conservation, and population genetic utility of the mitochondrial control regions of peltoperlid stoneflies

Introduction

Insect mitochondrial DNA (mtDNA) contains 13 protein coding genes, 22 tRNA genes, 2 rRNA genes and a single non-coding region that is thought to be the origin of mtDNA replication (Hoy 1994). The origin of replication or ‘control region’ (‘d-loop’ in vertebrates) is one of the most variable regions in both vertebrates and invertebrates and has been widely used as a population-level marker in vertebrates (Simon 1991). However, high A-T content (84-96%) and extensive size variation in the insect control region make sequencing this region difficult (Zhang and Hewitt 1994). Therefore, it has been characterized in only a few insect groups (Caccone et al. 1996, Clary and Wolstenholme 1987, Crozier and Crozier 1993, Rand and Harrison 1989, and Zhang et al. 1995) despite its potential usefulness for studying gene flow and dispersal.

Typically, sequence variation within the control region falls into three categories: 1) variable numbers of nucleotides in polynucleotide runs, 2) nucleotide substitutions, and 3) insertion/deletions of taxa-specific tandem repeats ranging in size from 150-750 bp (Zhang and Hewitt 1997). Across the entire control region, the pattern of sequence conservation varies among species. The control region of Drosophila spp. comprises two domains: a conserved sequence block near the tRNA^{ile} gene followed by a more variable domain (Clary and Wolstenholme 1987; Monnerot et al. 1990). In contrast, conserved sequence blocks are dispersed throughout the entire control region of grasshoppers, mosquitoes and butterflies (Zhang et al. 1995, Caccone et al. 1996, Taylor et al. 1993).

In general, primary sequence similarity in the control region is low, even among closely related species. The percent similarity between taxa is often less than that of computer-generated random sequences with the same nucleotide compositions (Clary and Wolstenholme 1987; Zhang and Hewitt 1997). However, the control regions of recently radiated groups tend to have high overall sequence similarities (e.g., the A. gambiae species complex; Caccone et al. 1996).

Low levels of primary sequence similarity across taxa have led to the suggestion of conserved structural elements in insect mitochondrial control regions (Zhang and Hewitt 1997). These ‘motifs’ are considered structural elements because their relative locations, rather than
their primary sequences, are conserved. Conserved structural elements have been identified in both hemimetabolous and holometabolous insects which may reflect the functional importance of these motifs.

Five structural elements have been identified through comparative analysis of dipteran and orthopteran control regions (Zhang and Hewitt 1997; Zhang et al. 1995). They include: 1) A poly-thymidine (poly-T) stretch at the 5' end of the control region (near the tRNA\textsuperscript{ile} gene) which may be involved in transcription control and/or the initiation of replication, 2) A [TA(A)]\textsubscript{n} stretch between the poly-T stretch (see above) and the conserved secondary structure (see below), 3) A highly conserved stem and loop structure that is potentially associated with the second strand replication origin, 4) Conservation of the sequences flanking the stem and loop structure, with a 5' consensus of ‘TATA’ and 3' consensus of ‘G(A)\textsubscript{n},’ and 5) A G+A-rich sequence block downstream of the highly conserved secondary structure (Zhang and Hewitt 1997).

The aim of this study was to characterize the control regions of two peltoperlid stoneflies, *Peltoperla tarteri* and *P. arcuata* (Plecoptera: Peltoperlidae), and determine whether the previously described motifs are found in these taxa. A further aim was to consider the implications of the structural elements of the control region with respect to its use in evolutionary studies.

**Methods**

*Tissue preparation*

DNA was isolated using a Gentra Systems Puregene® Genomic DNA Isolation Kit and the ‘mouse tail tissue protocol’ described in the kit instruction manual. Generally speaking, the procedure involves tissue homogenization, digestion of proteins with proteinase k, digestion of RNA with RNase, isopropanol precipitation of DNA, followed by ethanol washes and resuspension of the DNA in sterile ddH\textsubscript{2}O (50µL and 100µL, respectively, for small and large DNA pellets).

*PCR amplification and sequencing of the control region*

Initially, two conserved primers, 12sar (Kocher et al. 1989) and arf-met (Brasher and Ovenden 1992) were used to amplify the entire mitochondrial control region plus portions of the tRNA\textsuperscript{met} and 12S genes (Table 2.1). Each 25µL PCR reaction contained 22µL ddH\textsubscript{2}O, 1µL of
each primer (10µM), 1µL undiluted template, and a single Amersham Pharmacia Biotech ‘Ready To Go’ PCR bead. PCR reactions were cycled a total of 30 times in an MJ Research Model PTC-100 thermal cycler. The thermal cycling program was as follows: 3 min. @ 94ºC (initial denaturation), five cycles of 30 sec @ 94ºC (denature), 30 sec @ 44ºC (anneal), and 1:15 min. @ 72ºC (extension), followed by 25 cycles of 30 sec @ 94ºC, 30 sec @ 46ºC, and 1:15 min. @ 72ºC. A final extension step was carried out at 72ºC for 4 minutes.

PCR products were visualized by running 3µL on a 0.8% agarose gel stained with ethidium bromide. The initial amplification with 12sar and arf-met resulted in multiple PCR products. From these, a (1.5 kb) band was selected (based on the expected size of the control region in *P. tarteri*) and excised from the gel. DNA in the gel slice was extracted and purified using Qiagen’s Gel Extraction kit.

The purified PCR product was sequenced using ABI’s Big Dye Terminator kit. Each 15µL sequencing reaction contained 3µL dye terminator reaction mix, 1µL primer (10µM), 8µL ddH2O, and 3µL template. Cycle sequencing was performed in a Hybaid PCR Express thermal cycler and consisted of 30 cycles of 15 sec @ 94ºC, 15 sec. @ 50ºC, and 4 min. @ 60ºC. Sequencing reactions were cleaned using Millipore MultiScreen HV Plates and sequences were run on an ABI 377 automated sequencer.

A BLAST search revealed significant matches of the amplified sequence with tRNA^met and 12S in a variety of insects as well as a partial match with the control region of *Bactrocera* spp. (Diptera). These results were taken as confirmation that a portion of the control region had been successfully amplified. Several stonefly-specific internal primers designed from the original sequence were used to sequence the remainder of the control region (Table 2.1 and Figure 2.1).

Subsequent PCR and sequencing was performed as described above. All PCR products were sequenced in both directions. However, problems sequencing the 3’ to 5’ strand near the tRNA^met gene resulted in a small portion of the control region in *P. tarteri* being sequenced in only one direction and in an incomplete sequence for *P. arcuata.*
Figure 2.1: Schematic of primers used to amplify and sequence the control regions of *P. tarteri* and *P. arcuata*. Arrows indicate direction of primers. All primers except arf-met (Brasher and Ovenden 1992) and 12sar (Kocher et al. 1989) are first reported in this study.

Table 2.1: Primers used for PCR and sequencing of control region. The oligonucleotide sequences are in the 5' to 3' direction. R = A/G, Y = C/T. Primer locations are shown in Figure 2.1.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>arf-met</td>
<td>GGG GTA TGA GCC CAG TAG CTT AAT T</td>
<td>Brasher and Ovenden 1992</td>
</tr>
<tr>
<td>par-met</td>
<td>CTT TTT ATC AGG CAT CTC ATT TTT</td>
<td>This paper</td>
</tr>
<tr>
<td>par-dl24/2r</td>
<td>GAT RCT TTC CAT TTT AGT YCT AT</td>
<td>This paper</td>
</tr>
<tr>
<td>par-dl24/2</td>
<td>ATA GRA CTA AAA TGG AAA GYA TC</td>
<td>This paper</td>
</tr>
<tr>
<td>par-dl24</td>
<td>TGA CCC TCA AAC AAT ACA ATA CTT</td>
<td>This paper</td>
</tr>
<tr>
<td>par-12s2</td>
<td>GAC CCT CTA ATT AAC CAC ACC AA</td>
<td>This paper</td>
</tr>
<tr>
<td>12sar</td>
<td>ATA GTG GGG TAT CTA ATC CCA GTT T</td>
<td>Kocher et al. 1989</td>
</tr>
</tbody>
</table>

Sequence editing and alignment

Sequences were viewed and edited with ABI’s EditView version 1.0.1 and contigs (the consensus sequence from the region of overlap between 5' and 3' sequences) were assembled and edited with DNASTar SeqMan software version 4.03. Sequences were aligned using ClustalX (Thompson et al. 1997) and by eye.

Sequence comparison and secondary structure determination

Nucleotide composition, patterns of nucleotide substitution, and genetic distances were computed using the DAMBE software package (Xia 2000). Secondary structures and folding energies were determined using the Mfold server (http://mfold.wustl.edu/~mfold/rna/form1.cgi)
and the algorithms of Zuker et al. (1999) and Mathews et al. (1999). Secondary structures were drawn with RNAviz version 1.0 (DeRijk and DeWachter 1997).

**Results and Discussion**

*Sequence analysis*

The entire 1270 bp control region was sequenced in *Peltoperla tarteri* and a partial sequence (1072 bp) was obtained from *P. arcuata* (Figures 2.2 and 2.3). All sequence comparisons were made from the region of overlap between the two sequences.

**A. *P. tarteri***

![Diagram of Peltoperla tarteri control region](image)

**B. *P. arcuata***

![Diagram of Peltoperla arcuata control region](image)

**Figure 2.2:** Schematic of control regions of *P. tarteri* (A) and *P. arcuata* (B). The large shadowed boxes denote the control region sequences before and after the repeat region (BRP and ARP, respectively). The small shadowed boxes indicate the inverted repeats (rp1-rp3). Numbers below the boxes indicate nucleotide position.
\textit{P. tarteri} (tRNA) ATTTACATGATTTATCCCTATCACGATAATCCTTTTTATCAGGGTTCATT \[50\]
\textit{P. arcuata} (tRNA) --------------------------------------------- \[0\]
\textit{P. tarteri} TTTTTTTTATTATTTACGGGAACTTGATAGAGTATGGATTGTTATATAGA \[100\]
\textit{P. arcuata} --------------------------------------------- \[0\]
\textit{P. tarteri} TTTTATTATTTTTTTACCTTTAATTGAAGGAGCTGAAATAATTTCCAAA \[150\]
\textit{P. arcuata} -----------------.TT............T........T........ \[33\]
\textit{P. tarteri} TTTATTATTATTTTTTTACCTTAAATTGAAAGGAGCTGAAAATTTCCAAA \[200\]
\textit{P. arcuata} .................................................A \[83\]
\textit{P. tarteri} GGTTTAACCCATCCCCCAATTCTAAAGTTGTCCACGAAAACTAAAACCT \[250\]
\textit{P. arcuata} ..........T..........................A............ \[133\]
\textit{P. tarteri} ATCAGTTTTTTTTTTTGTATTTAATAATTACGTGTATATAAGAATTTA \[300\]
\textit{P. arcuata} .................A..A..A......................T........ \[183\]
\textit{P. tarteri} ATTTATATACATTTATACCAATTTATACCTATTAGCTTATATAAAAAAGAAA \[350\]
\textit{P. tarteri} AAATAAAAGATATTTAAATTTAATTATTTATATTTATTTTATA- \[399\]
\textit{P. arcuata} ........T-..................................G....C...A..T..T \[282\]
\textit{P. tarteri} ATATAATAAATACATATATATATATATATATATTATACATTATT \[449\]
Figure 2.3: Aligned sequences (5' to 3') for complete *Peltoperla tarteri* and partial *P. arcuata* control region. Dots indicate similarity to *P. tarteri*. Conserved sequence blocks similar to those reported in Orthoptera and Diptera (Zhang et al. 1995) and inverted repeats are top-marked with a line.

The nucleotide sequence of the control regions of both peltoperlid stoneflies were characterized by high A-T content (Table 2.2) and a 74 bp tandem repeat which begins with the sequence ‘GGGGGC’ (Figures 2.2 and 2.3). *Peltoperla tarteri* had three repeats; the first two were exact copies of one another; however, the third repeat was somewhat divergent (Figure 2.3). There were only two copies of the repeat in the *Peltoperla arcuata*, and they were very similar to each other (Figure 2.3). The difference in copy number of the tandem repeat accounted for most of the size variation between control regions of the two species (Table 2.2).
Across the entire region of overlap, the peltoperlid sequences were 89.4% similar (Table 2.2). The high percent similarity and low genetic distance (Table 2.2) in the control regions of *P. tarteri* and *P. arcuata* suggest that these species diverged more recently than *D. yakuba* and *D. virilis* which have 78% homology in their control regions (Clary and Wolstenholme 1987), but not as recently as *Jalmenus* spp. (Lepidoptera) which have 94% homology (Taylor et al. 1993).

In the peltoperlid stoneflies, sequence divergence was primarily due to transitions, which outnumbered transversions 1.4 to 1, and to a lesser extent, insertions and deletions (Table 2.2). This pattern of nucleotide substitution is somewhat unusual, since transversions usually outnumber transitions in the control regions of insects (e.g., Lessinger and Azeredo-Espin 2000). However, the strong directional mutation bias and high percentage of A→T transversions (65%) observed in *P. tarteri* and *P. arcuata* (Figure 2.3) strongly resemble the pattern described in other insects (e.g., Lessinger and Azeredo-Espin 2000).

Table 2.2: Summary comparison of *P. tarteri* and *P. arcuata* nucleotide sequences from the mitochondrial control region. Percent similarity, A-T content, number of substitutions, transition/transversion ratios, number of indels, Tamura Nei distances (D), and lengths are shown for each region. BRP = before repeat, ARP = after repeat. The lowest A-T content and highest percent similarity were found in the repeat region. *The insertion/deletion of rp3 in *P. tarteri* was considered a single event and scored as one indel.

<table>
<thead>
<tr>
<th>Region</th>
<th>% Sim.</th>
<th>% A-T</th>
<th>Subst.</th>
<th>Ts:Tv</th>
<th>Indels*</th>
<th>D</th>
<th>Length (bp) <em>P. tarteri</em></th>
<th>Length (bp) <em>P. arcuata</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>BRP</td>
<td>88.7</td>
<td>84.0</td>
<td>59</td>
<td>1.03:1</td>
<td>21</td>
<td>0.09</td>
<td>705</td>
<td>693</td>
</tr>
<tr>
<td>Repeat</td>
<td>91.8</td>
<td>74.9</td>
<td>12</td>
<td>5:1</td>
<td>6</td>
<td>0.09</td>
<td>220</td>
<td>151</td>
</tr>
<tr>
<td>ARP</td>
<td>89.5</td>
<td>81.1</td>
<td>24</td>
<td>1.67:1</td>
<td></td>
<td>0.12</td>
<td>228</td>
<td>228</td>
</tr>
<tr>
<td>Entire</td>
<td>89.4</td>
<td>81.7</td>
<td>95</td>
<td>1.4:1</td>
<td>27</td>
<td>0.10</td>
<td>1153</td>
<td>1072</td>
</tr>
</tbody>
</table>

Visual inspection of the aligned sequences of *P. tarteri* and *P. arcuata* revealed fewer nucleotide substitutions in the repeat region than in the remainder of the control region (Figure 2.3). Thus, the number and pattern of nucleotide substitutions in the repeats was compared to that of the regions before (BRP) and after (ARP) the repeat (Table 2.2). Indeed, the repeat region showed the highest percent similarity, lowest A-T content, and the highest transition to transversion ratio (Table 2.2) suggesting a different evolutionary pattern for the repeats. High percent similarity among tandem repeats (97.5%) has also been reported in Orthoptera (Zhang et al. 1995).
Conservation of structural elements

Conserved sequence blocks were distributed throughout the control regions of the peltoperlid stoneflies (Figure 2.3). Thus, the pattern of sequence conservation was more similar to that reported for grasshoppers, mosquitoes, and butterflies than to that of *Drosophila* spp.

Several motifs that correspond to structural elements previously described in Orthoptera and Diptera by Zhang et al. (1995) were found in *P.tarteri* and *P. arcuata*. The poly-thymidine stretch downstream of the tRNAile gene was present in the peltoperlid stoneflies (Conserved Block 1; Figure 2.3). The conservation of this motif in Plecoptera, Orthoptera, and Diptera (Figure 2.4) supports the idea that the poly-T region may be important in the initiation of mtDNA replication or have some other functional role.

![Figure 2.4: Alignment of ‘Conserved Block 1’ sequences in peltoperlid stoneflies and Orthoptera and Diptera. The poly-T region is shown in bold in *P. tarteri*.](image)

Another conserved sequence block was identified downstream of the poly-T region in *P. tarteri* and *P. arcuata* (Conserved Block 2; Figures 2.3 and 2.5). This motif, which may correspond to the ‘[TA(A)]ₙ’ sequence described by Zhang et al. (1995), contains the relatively conserved core sequence ‘5’-A…TA……ATTTA……TT…ATA……ACATTT-3’ (Figure 2.5). The core sequence resembles the template stop signals for D-loop synthesis in human and mouse mtDNA (Clayton 1982 as cited by Zhang et al. 1995) and the presence of this sequence in similar locations in Orthoptera, Diptera, and Plecoptera suggests its possible role in transcription or replication control.

![Figure 2.5: Alignment of ‘Conserved Block 2’ sequences in peltoperlid stoneflies and Orthoptera and Diptera. Consensus bases are shown in bold.](image)
Conservation of secondary structure

A conserved stem and loop (hairpin) structure was identified within the repeat regions of *Peltoperla tarteri* and *P. arcuata* (Figures 2.6A-2.6C). This stem and loop structure is formed by an inverted repeat that is centered around a ‘5’-GGGGGC-3’’ motif (Figures 2.3 and 2.6). Although there are some substitutions within copies 1 and 2 of the repeat, the structural integrity of the hairpin is maintained in both taxa (Figures 2.6A-2.6C). In contrast, the copy 3 of *P. tarteri* is less conserved (Figure 2.3) and did not form an energetically stable structure.

The stem and loop structures of the peltoperlid stoneflies are structurally and energetically similar to that of several other insect taxa albeit with somewhat smaller loops (Figures 2.6A-2.6G). Although repeats within the control region of the cricket, *Gryllus firmus*, contain the dyad symmetric sequence ‘5’-GGGGGCATGCCCCC-3’’ which is capable of forming a cruciform structure (Rand and Harrison 1989), the secondary structure proposed by Zhang et al. (1995) does not include this region (Figure 2.6E). The importance of the very similar ‘5’-GGGGGC-3’’ motif in the formation of a hairpin structure in the peltoperlid stoneflies suggests that, in *G. firmus*, the ‘true’ hairpin structure may be different from that shown in Figure 2.6E.

Interestingly, there is little primary sequence similarity among the stem and loop structures across insect taxa (Figure 2.6) while the flanking regions tend to be highly conserved. The 5' flanking sequences are A-T rich and, with the exception of *G. firmus* (which may form a secondary structure different from that shown in Figure 2.6E, see above), all contain a ‘TTAT’ motif (Figure 2.6). The 3' flanking sequences are G-A rich and have the consensus sequence ‘(GA)_n’. The similarity of the 5' and 3' flanking regions in *P. tarteri*, *P. arcuata*, and other insect taxa further supports Zhang and Hewitt’s (1997) proposal that structural elements rather than primary sequences are conserved within the control region.

Hairpin structures are important because the replication of circular DNA molecules has been shown to initiate within or close to them (Zannis-Hadjopoulos et al. 1988). However, a functional role cannot be assigned because of inferred secondary structures or conservation of primary sequence data. Thus, experimental data are required to confirm any of the proposed functional roles of structural elements in the control region.
Utility of the control region in evolutionary studies

The results of this study strongly support idea that structural elements are conserved throughout the control region of insect mitochondrial DNA. These distinct structural characteristics have important implications for the use of the control region in evolutionary studies. Of particular concern is the widespread occurrence of tandem repeats.

The high frequency of tandem duplication in the insect control region suggests that sequence variation between lineages may not result from descendence but instead from variation of tandem repetition (Zhang and Hewitt 1997). In addition, length variation may be due to recurrent gain or loss of mutations and represent convergent evolution. This complicates any analyses concerning genetic diversity and population history. Thus, the control region should only be used to study genetic polymorphisms and genetic structure where it does not contain tandem repetition.
Figure 2.6: Possible conserved secondary structure in peltoperlid stoneflies (A-C), Orthoptera (D and E), Hymenoptera (F), and Diptera (G). Underscoring indicates conserved flanking regions. $\Delta G =$ folding energy. Figures D-F are after Zhang et al. (1995).
Chapter 3: Genetic evidence for ‘leaky’ cohorts in the semivoltine stonefly, *Peltoperla tarteri* Stark and Kondratieff (Plecoptera: Peltoperlidae)

Introduction

Genetic techniques have been increasingly used to address important questions about dispersal and gene flow of freshwater invertebrates (Duan et al. 2000, Hughes et al. 1999, Hughes et al. 1998, Schmidt et al. 1995). Such ‘indirect’ studies of dispersal use the relationship between dispersal and population genetic structure to infer patterns of movement between populations (Bunn and Hughes 1997, Meyers et al. 2000, Slatkin 1985). A chief assumption of these studies is that dispersal (and resultant gene flow) is the primary determinant of population genetic structure.

However, numerous animal examples have shown that population genetic structure can be affected by factors other than dispersal (e.g., natal site philopatry, Avise 1994; mating systems, Meyers et al. 2000; effective population size, Duvernell and Turner 1998; historical patterns of gene flow, Culver et al. 2000; and life history, Preziosi and Fairburn 1992, Zera 1981). Therefore, when possible, behavior and life history should be taken into account when inferring patterns of stream insect dispersal from genetic data.

Many stream insects have long life cycles that result in the simultaneous existence of multiple cohorts throughout the larval development period (Merritt and Cummins 1996). If larval development is fixed, successive cohorts may be reproductively isolated and, as a result, genetically distinct. In such cases, significant levels of genetic differentiation between cohorts could confound estimates of dispersal based on population genetic structure.

Most genetic studies of stream insect dispersal have avoided this problem altogether by focusing on univoltine organisms (Hughes et al. 1998, Jackson and Resh 1992, Schmidt et al. 1995, Sweeney et al. 1986, Sweeney et al. 1987, Funk et al. 1988), or by assuming that cohorts were ‘leaky’ (Hughes et al. 1999). In the long-lived stonefly *Pteronarcys proteus*, genetic evidence suggested high levels of gene flow between cohorts (White 1989). However, it is questionable whether the individuals used were actually from different cohorts, since body length was used to identify cohorts (White 1989). Body length is an unreliable for this purpose in stoneflies because of the amount of variation within size classes (Latto 1992, Townsend and
Pritchard 1998) and the difficulty of assigning instars in hemimetabolous insects (Fink 1984, Stewart and Stark 1993).

Further, White (1989) used protein electrophoresis (allozymes) to determine levels of genetic differentiation between cohorts. Allozymes are functional products of nuclear genes that are biparentally inherited. They undergo recombination, can be subject to selection, and, as a result, tend to be less variable than non-coding regions of DNA (Mitton 1994, Niegel 1997, and Parker et al. 1998). Thus, allozymes are often not variable enough to distinguish populations (see DeSalle et al. 1987) and may not be able to detect variation between cohorts of Pteronarcys proteus. Thus, it is clear that further work is needed to determine the contribution of inter-cohort variation to the genetic structure of stream insect populations.

One of the most variable molecular markers in both vertebrates and invertebrates is the non-coding origin of replication of mitochondrial DNA (called the ‘control’ or ‘Control’ region in invertebrates, and the ‘d-loop’ in vertebrates) (Simon 1991). Because the d-loop is ideal for assessing levels of genetic differentiation at the population level, it has been widely used as a population-level marker in vertebrates. However, because length variation and high A-T content often complicate amplification and sequencing of this region (Zhang and Hewitt 1994), the control region has been characterized in only a few insect groups (Caccone et al. 1996, Clary and Wolstenholme 1987, Crozier and Crozier 1993, Rand and Harrison 1989, and Zhang et al. 1995), despite its potential usefulness.

As part of the current study, I refined a technique that allowed me to amplify and sequence the mitochondrial control region in the stonefly Peltoperla tarteri Stark and Kondratieff (Plecoptera: Peltoperlidae). I have used the information from this region to infer patterns of gene flow between cohorts in two P. tarteri populations in the Appalachian Mountains of southwestern Virginia.

Study Organism

The stonefly P. tarteri is abundant in many Appalachian headwater streams (Stark and Kondratieff 1987). Its diet consists chiefly of coarse particulate organic matter, such as leaves that have fallen into the stream (Ruggles and Tarter 1991). A P. tarteri population in West Virginia located 820m above sea level (a.s.l.) was reported to be univoltine (Ruggles and Tarter 1991). However, many insects switch to semivoltinism in colder areas (Danks 1992), and the
congener *P. arcuata* is semivoltine at elevations similar to that of the present study sites (Yokum et al. 1995). Therefore, I predicted that the study populations of *P. tarteri* would be semivoltine.

Multiple developmental pathways and ‘cohort splitting’ have been well documented in stoneflies (Brittain and Lillehammer 1987, Moreira and Peckarsky 1994, Marten and Zwick 1989, Wolf and Zwick 1989). Thus, I predicted that developmental plasticity would prevent successive cohorts of *P. tarteri* from becoming reproductively isolated, and as a result, I expected low levels of inter-cohort genetic differentiation.

*Site Description*

Hunters Branch and Sartain Branch (N 37º 22’ W 80º 31’) are headwater streams located in Giles County, Virginia, U.S.A. (Figure 3.1). The streams straddle the Eastern Continental Divide, with Hunters Branch flowing into the New River of the Mississippi Drainage and Sartain Branch flowing into the James River of the Atlantic Drainage. The streams are located in northern hardwood forest at ~1300m above sea level. A unique feature of this area is nearby Mountain Lake, the only natural lake in upland Virginia, and the large area of protected forest contained in the Mountain Lake Wilderness of Jefferson National Forest and the privately held Wilderness Conservancy at Mountain Lake.

![Figure 3.1: Study sites in Giles, County, Virginia. The dashed line represents the Eastern Continental Divide, and circles indicate sampling locations.](image)
Methods

Temperature
An Optic Stow Away data logger in each stream recorded temperature hourly from June 1997 to May 1998.

Species identification
Positive species identifications were made with laboratory-reared adult males. Larvae with darkened wing pads were collected from streams in mid-May 1999, brought to the lab, and placed in Plexiglas rearing chambers containing stream water and a small layer of rocks and leaves from the stream. Compressed air forced water to recirculate through the chambers. Larvae were reared in the dark at a constant temperature similar to ambient stream temperatures in May (14 °C). Emergent adults were collected daily and preserved in 100% ethanol. Adult males were identified under a dissecting scope, and sent to B. Kondratieff (Colorado State University, Dept. of Entomology) for confirmation.

Voltinism
Low stream flows made it impossible to use nets or other sampling devices that required water current. Therefore, larvae were obtained by collecting leaf packs from the stream and placing the leaves into a large plastic bucket. To ensure adequate sample sizes, leaves were sorted on site until ~100 larvae were counted (if possible). The remaining unsorted leaves were placed into Ziploc bags, preserved with formalin, and brought to the laboratory where the remaining larvae were removed.

In July and August when streams were very low and few leaf packs remained submerged, larvae were collected by dipping rocks into a water-filled bucket and cleaning the rocks by hand. No head width data are reported for December 1997, March 1998, or September 1998. In December 1997 and March 1998, road conditions prevented access to the streams for sampling (the road is not maintained by the county during this time). Samples were collected in September 1998 but there were very few insects (<10). It is thought that the stoneflies survive the extremely low surface flows of late summer by burrowing into the hyporheic zone.

Head widths (outside eye to outside eye) were measured using a Jandel Scientific SigmaScan Pro Measurement System version 4.01. Measurements were made to nearest 0.01mm (± 0.004mm). Since laboratory data were not available to generate a growth model from which to calculate instars, larvae were grouped into arbitrary size classes of 0.1mm. Plots made using size
classes of 0.2mm and 0.05mm did not significantly alter results. Thus, though this method does not show development of instars, it presents growth through time in an informative manner.

Adult emergence

Based on observations from the previous summer, sites were visited every other day from late May 1998 to late July 1998. Adults were captured by holding a canvas beating sheet under vegetation (usually *Rhododendron* or *Kalmia*) and disturbing the vegetation with a large stick. *Peltoperla tarteri* adults that landed on the beating sheet were stunned with a light ethanol spray, then collected and placed in plastic vials containing 100% ethanol. To ensure high quality material for DNA analysis, adults were placed at 4 °C immediately upon returning to lab, and the ethanol was replaced after 24h.

Genetic variation among cohorts

The level of genetic differentiation among cohorts was determined by comparing the distribution of haplotypes in the control region of mtDNA of adults collected May-July of 1998 with that of larvae with darkened wing pads collected in May 1999. Adults and larvae used in cohort comparisons came from the same stream (e.g., adults from Hunters Branch compared only with larvae from Hunters Branch, etc.).

Tissue preservation and preparation

Adults were preserved as previously described. Larvae were kept alive until reaching the laboratory where they were placed in 1.5 mL microcentrifuge tubes and frozen at –80 °C. DNA was isolated using a Gentra Systems Puregene® Genomic DNA Isolation Kit and the ‘mouse tail tissue protocol’ described in the kit instruction manual. Generally speaking, the procedure involves tissue homogenization, digestion of proteins with proteinase k, digestion of RNA with RNase, isopropanol precipitation of DNA, followed by ethanol washes and resuspension of the DNA in sterile ddH₂O (50µL and 100µL, respectively, for small and large DNA pellets).

PCR amplification and sequencing of control region

Two stonefly-specific primers, 12s-2 (5' - GAC CCT CTA ATT AAC CAC ACC AA -3') and dl242r (5'- GAT RCT TTC CAT TTT AGT YCT AT -3') were used to amplify a 454 base pair (bp) section of the mitochondrial control region. Each 25µL PCR reaction contained 22µL ddH₂O, 1µL of each primer (10µM), 1µL undiluted template, and a single Amersham Pharmacia Biotech ‘Ready To Go’ PCR bead. PCR reactions were cycled a total of 30 times in an MJ Research Model PTC-100 thermal cycler. The thermal cycling program was as follows: 3 min.
@ 94ºC (initial denaturation), five cycles of 30 sec @ 94ºC (denature), 30 sec @ 44ºC (anneal), and 1:15 min. @ 72ºC (extension), followed by 25 cycles of 30 sec @ 94ºC, 30 sec @ 46ºC, and 1:15 min. @ 72ºC. A final extension step was carried out at 72ºC for 4 minutes. If not visualized immediately, amplification products were stored at 8ºC.

PCR products were visualized by running 3µL on a 0.8% agarose gel stained with ethidium bromide. Qiagen’s QIAquick PCR purification kit was used to prepare PCR products for direct sequencing. All PCR products were sequenced in both directions using ABI’s Big Dye Terminator kit. Each 15µL sequencing reaction contained 3µL dye terminator reaction mix, 1µL primer (10µM), 8µL ddH2O, and 3µL template. Template dilution ranged from 1:1 to 1:6 depending upon PCR product yield.

Cycle sequencing was performed in a Hybaid PCR Express thermal cycler and consisted of 30 cycles of 15 sec @ 94ºC, 15 sec. @ 50ºC, and 4 min. @ 60ºC. Sequencing reactions were cleaned using Millipore MultiScreen HV Plates and sequences were run on an ABI 377 automated sequencer.

Genetic data analysis

Sequences were viewed and edited with ABI’s EditView version 1.01 and contigs (the consensus sequence from the region of overlap between 5' and 3' sequences) were assembled and edited with DNAstar SeqMan software version 4.03. Sequences were aligned with ClustalX (Thompson et al. 1997) and by eye. Haplotypes had variable numbers of a 74 base pair repeat, so in some cases gaps were required in order to align sequences. These were the only gaps in the alignment and the insertion/deletion of the repeat was considered a single event, therefore, individual gaps were scored as 0.014 (1/74) so that the 74bp repeat was given a weight of 1.0. Levels of genetic differentiation between cohorts were determined using Arlequin population genetic data analysis software version 2000 (Schneider et al. 2000).

Results

Temperature

The average annual stream temperature was 8.7 ºC. Maximum and minimum temperatures were 18.2 ºC (August) and 2.0 ºC (January) (Figure 3.2).
Figure 3.2: Mean monthly temperature in Sartain Branch from June 1997 – May 1998. Error bars indicate ±1 std. dev.

**Voltinism**

Larvae were present in the smallest size classes from June to November, indicating that, although the majority hatched occurs in June and July, some hatched as late as October and November (Figure 3.3). Throughout the year, two distinct cohorts were present, though the year-old cohort is less pronounced in June due to the large number of newly hatched larvae (Figure 3.3). The presence of an obvious second cohort in August further supports this idea because it is very unlikely that larvae hatched in June or July obtain half their full size in four to eight weeks.

The pattern of larval growth suggests a semivoltine life cycle for *P. tarteri* at Hunters Branch (data for Sartain Branch, not shown, give the same result). Semivoltinism is common for stoneflies at high elevations and/or northern latitudes (Danks 1992) and has been reported for
other peltoperlid species in West Virginia (Yokum et al. 1995). At the sites used in the present study, most larvae continue developing for 22-23 months after they hatch in June/July and emerge mid-May to mid-July.

Figure 3.3: Growth phenology of *Peltoperla tarteri* at Hunters Branch (Giles Co., VA). The presence of two distinct cohorts is consistent with a semivoltine life cycle, which is common for stoneflies at high elevations. Solid line at top of graph indicates adult emergence, diagonal lines and red type indicate presumptive cohorts. Cohort year is assigned according to date of hatch. Bar widths are proportional to the relative frequency of the size class. (Total n = 1406).

Adult emergence

Adults began to emerge in late-May and peaked in June (Figure 3.4). Males outnumbered females for the first 3-4 weeks of the emergence period. Numerous attempts were made to collect adults from vegetation at 5, 10 and 20m from the stream, but all successful collections of adults were from vegetation directly over the stream or on the stream bank. Most larval exuviae were found on exposed rocks in the stream, adults were never observed flying during the day, and adults not captured immediately crawled rather than flew away from the collector.
Figure 3.4: Phenology of adult emergence in Hunter’s Branch, Sartain Branch, and two nearby tributaries during summer 1998. Males outnumbered females in the first four weeks of emergence, but there were more females towards the end of the flight period. Delayed female emergence probably reflects longer development time required to reach large body size. Numbers above bars indicate sample sizes. (Total n = 90)

Patterns of genetic variation

Of the 454 base pairs that were sequenced, eleven sites were variable (Table 3.1). Seven variable sites were located outside the repeat region, and four variable sites were within the repeat, which began at site 126 (Table 3.1).
Table 3.1: Variable sites for 13 haplotypes of the control region from 93 *P. tarteri* individuals. Dots indicate similarity to haplotype 2. The dashed line indicates the beginning of the repeat region. *P. arcuata* is the sister species of *P. tarteri* and is included for reference.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Position</th>
<th>2</th>
<th>14</th>
<th>20</th>
<th>30</th>
<th>66</th>
<th>87</th>
<th>117</th>
<th>172</th>
<th>246</th>
<th>284</th>
<th>358</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2 AACACAGTGC</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>... A ...</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>.... G ....</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>..... T .....</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>........ A .-</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>6</td>
<td>.......... A</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>....... T T</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>.. .... A G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>12</td>
<td>....... T T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>15</td>
<td>.... G G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>.... G G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>T T</td>
<td></td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>-</td>
</tr>
</tbody>
</table>

*P. arcuata*: . . . . A G - - T T -

Thirteen haplotypes were identified from the control region sequences of the 93 individuals used in this study. Eight haplotypes were shared between cohorts and/or populations, and five were unique (found only in a single individual) (Table 3.2). Haplotype 2 was dominant, found at all sites, and had a frequency ranging from 0.20 to 0.69.
Table 3.2: Haplotype frequencies by cohort. Numbers in parentheses indicate sample sizes. HB = Hunters Branch, SB = Sartain Branch.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.32</td>
<td>0.42</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
<td>0.25</td>
<td>0.69</td>
<td>0.35</td>
</tr>
<tr>
<td>3</td>
<td>0.28</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.04</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>0.03</td>
<td>0.08</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>0.15</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>0.08</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>

Genetic variation among cohorts

Genetic variation among cohorts was partitioned using a nested Analysis of Molecular Variance (AMOVA) design (Table 3.3). Genetic variation between cohorts (within sites) was effectively zero, and most of the variation (77%) could be attributed to within population variation (Table 3.3).
Table 3.3: Partitioning of genetic variance according to AMOVA analysis (Weir and Cockerham 1984). Variance components were used to calculate F statistics and significance of F statistics was tested using a non-parametric permutation approach (# of permutations = 16002). Asterisk indicates significance (P < 0.001).

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>SS</th>
<th>Variance Components</th>
<th>% of variation</th>
<th>F\textsubscript{ST}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among streams (sites)</td>
<td>1</td>
<td>13.9</td>
<td>0.3</td>
<td>22.7</td>
<td>0.23</td>
</tr>
<tr>
<td>Among cohorts within sites</td>
<td>2</td>
<td>1.9</td>
<td>0.0</td>
<td>-0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Within sites</td>
<td>89</td>
<td>87.0</td>
<td>1.0</td>
<td>77.4</td>
<td>0.23*</td>
</tr>
<tr>
<td>Total</td>
<td>92</td>
<td>102.8</td>
<td>1.3</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Similarly, an exact test of population differentiation (Rousset and Raymond 1997) indicated that only between-stream comparisons showed highly significant levels of genetic differentiation (p < 0.001; Table 3.4).

\(N_m\), (effective population size \(X\) the migration rate) or the number of migrants entering a population per generation, is often used as a measure of the amount of gene flow between populations. \(N_m\) values calculated from Slatkin linearized \(F_{ST}\)’s (Slatkin 1995) indicate that gene flow between cohorts is very high (Table 3.4) and that \(N_m\) between populations in separate streams is only slightly greater than 0.5, the theoretical value below which gene flow cannot counteract the effects of genetic drift and/or selection (Slatkin 1985). The high levels of genetic differentiation between the Hunters Branch and Sartain Branch populations suggest that adult dispersal of \(P.\ tarteri\) is very limited because these streams are located in different drainages (Figure 3.1). This topic will be addressed in a forthcoming paper.

In contrast, low levels of differentiation between cohorts (Table 3.4) suggest that there is significant gene flow between \(Peltoperla tarteri\) cohorts. Therefore, for the populations of \(P. tarteri\) used in this study, larvae collected from a single site can be pooled without regard to cohort.
Table 3.4: Above diagonal: Pairwise non-differentiation exact P-values calculated with 6000 dememorisation and 100,000 Markov chain steps. P-values indicate the probability of observing a haplotype distribution equally or less likely than the observed sample configuration under the null hypothesis of panmixia (Schneider et al. 2000). Asterisks indicate that compared populations are significantly different. Below diagonal: Nm values calculated from Slatkin’s linearized FST’s (Slatkin 1985). HB = Hunters Branch, SB = Sartain Branch

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HB 1999</td>
<td>--</td>
<td>0.87 ± 0.006</td>
<td>0.00 ± 0.00*</td>
<td>0.008 ± 0.004*</td>
</tr>
<tr>
<td>HB 1998</td>
<td>∞</td>
<td>--</td>
<td>0.0002 ± 0.00*</td>
<td>0.0030 ± 0.002*</td>
</tr>
<tr>
<td>SB 1999</td>
<td>1.45</td>
<td>1.7</td>
<td>--</td>
<td>0.33 ± 0.009</td>
</tr>
<tr>
<td>SB 1998</td>
<td>2.07</td>
<td>2.8</td>
<td>∞</td>
<td>--</td>
</tr>
</tbody>
</table>

Discussion

Ecological life history data collected during the present study indicate that *P. tarteri* is semivoltine in the study streams in southwestern Virginia. Therefore, individuals from a given cohort (and their subsequent offspring) should reproduce every other year. In this study, *P. tarteri* adults were collected during two consecutive summers, and were presumably from different cohorts. Yet, the genetic data suggest that there is significant gene flow between these groups. How is this genetic homogeneity between putative cohorts maintained?

Developmental plasticity, especially in voltinism, has been well documented in a variety of stoneflies (Marten and Zwick 1989, Moreira and Peckarsky 1994, Wolf and Zwick 1989, Townsend and Pritchard 1998). On a regional scale, changes in photoperiod and climate (esp. temperature) are thought to cause the switch from univoltine to semivoltine (or longer) life cycles at northern latitudes (Tauber et al. 1986). On a local scale, voltinism can be affected a variety of factors including temperature (Moreira and Peckarsky 1994), habitat permanence (Dieterich and Anderson 1995, Williams et al. 1995), and the frequency and predictability of
disturbance (Robinson et al. 1992). For example, *Nemoura trispinosa* switched from univoltinism to semivoltinism along a gradient of habitat permanence, with the longest life cycles occurring in the most stable environments (Williams et al. 1995). Similarly, the life cycle of *Hesperoperla pacifica* was longer in a stream with a predictable flow regime than in a nearby stream that has seasonally cyclic flow (Robinson et al. 1992).

Multiple developmental pathways (e.g. changes in voltinism) also occur within the same population. Extended larval recruitment and considerable variation within size classes of *Agnetina capitata* led Peckarsky and Moreira (1994) to conclude that most of the population completed the life cycle in 3 years but a few individuals completed a 2-yr. life cycle. Likewise, in a laboratory experiment, a few *Nemoura pictetii* individuals completed larval development in 84-153 days, while most individuals required closer to 300 days to complete development (Wolf and Zwick 1989). Such ‘cohort splitting’, where the development of a few individuals is delayed or increased with respect to the rest of the population is likely the cause of differences in voltinism within populations.

Cohort splitting often results from the presence or absence of a resting stage during some point in the life cycle (Tauber et al. 1986). Diapause refers to hormonally induced state of low metabolic activity during which development is arrested and resistance to environmental extremes is increased (Tauber and Tauber 1981). Diapause is the main factor that synchronizes development of insect populations, but it can also lead to cohort splitting if a portion of the population does not enter diapause (Wolf and Zwick 1989). Variability in egg diapause results in extended larval recruitment sometimes for many months beyond the mating season (Townsend and Pritchard 1998).

In stream insects, cohort splitting may reflect a ‘bet-hedging’ strategy (see Hairston et al. 1985) whereby individuals avoid adverse conditions (such as low flows and high stream temperatures of late summer). Though there are costs associated with different developmental pathways (e.g. lowered fecundity for longer life cycles and smaller sizes associated with shorter development times), some individuals are likely to survive catastrophic events.

Such variability in development is the most likely mechanism for gene flow between consecutive cohorts of *P. tarteri*. While the majority of population is semivoltine, larval recruitment continued into October and November, several months after any adults were collected. Therefore, it seems likely that there was a cohort split, whereby the majority of the
population began developing immediately but a small portion of the eggs underwent a 3-5 month egg diapause. While egg diapause has not been reported previously in *P. tarteri* (see Ruggles and Tarter 1991), a 6-month diapause has been reported in the sister species, *Peltoperla arcuata*, and in *Tallaperla maria*, another peltoperlid stonefly (Yokum et al. 1995).

The streams used in this study have low flows from late August to October. In September 1998, flows were low and very few larvae were collected, though this was not a particularly dry summer. Larvae probably survived the low flow conditions by burrowing into the hyporheic zone. However, larvae may be unable to survive more severe conditions and since severe droughts are unpredictable, having a few individuals enter diapause every summer ensures that diapausing eggs will be present in the case of extremely dry conditions. Thus, cohort splitting may represent a ‘bet-hedging’ strategy that allows *P. tarteri* some individuals to survive very dry years.

Because they begin growing as much as 5 months later than non-diapausing individuals, *P. tarteri* individuals that undergo diapause may not be able to complete development in 2 years and may require an additional year. Such delayed development would result in adults from distinct cohorts reaching maturity and mating during the same season. Low levels of gene flow are required to counteract the effects of genetic drift and natural selection (Slatkin 1985). Therefore, only a small number of individuals with delayed development would need to reach maturity and mate with individuals that had completed development in 2 years in order to account for the low levels of genetic differentiation between cohorts observed in the present study.

The same scenario (i.e., adults from distinct cohorts mating at same time) would occur if some individuals completed development in one year rather than two. *P. tarteri* is univoltine at the type locality in West Virginia (Ruggles and Tarter 1991), so some individuals, especially those that hatch early in the season, may complete development in one year. However, the presence and/or mechanism of cohort splitting in *P. tarteri* cannot be confirmed without laboratory hatching experiments.

My results suggest that larval development of *P. tarteri* is not fixed and that cohorts are ‘leaky’ (i.e., there is enough gene flow between them to counteract the effects of genetic drift and selection). However, the role of inter-cohort genetic variation and other life history traits in
determining population genetic structure need to be examined in other stream insects so genetic studies of gene flow and dispersal can be interpreted in the appropriate context.
Chapter 4: Gene flow, dispersal, and nested clade analysis among populations of the stonefly *Peltoperla tarteri* (Plecoptera: Peltoperlidae) in the southern Appalachians

Introduction

Despite a number of recent studies (Griffith et al. 1998, Hughes et al. 1999, and Peckarsky et al. 2000) focused on the issue, patterns of stream insect dispersal in temperate streams of North America remain poorly understood. Movement of benthic invertebrates is notoriously difficult to measure directly using traditional means (see Bunn and Hughes 1997); therefore, the use of genetic markers to indirectly estimate stream insect dispersal is becoming more widespread.

The most extensive genetic studies of stream insect dispersal have been done in tropical streams of Australia (Bunn and Hughes 1997, Schmidt et al. 1995 and Hughes et al. 1998) and have led to surprising results. In Australian stream insects, the highest levels of genetic differentiation were detected at the smallest spatial scales and a number of samples were not in Hardy-Weinberg equilibrium (Bunn and Hughes 1997). The ‘patchy recruitment’ hypothesis explains these patterns as the result of widespread adult dispersal and limited larval movement coupled with the genetic effects of patchy oviposition by asynchronously emerging insects (Bunn and Hughes 1997).

The highly synchronized life cycles of temperate stream insects have led to different predictions about population genetic structure in North American streams. Patchy oviposition is not expected with synchronized emergence, so the effect of recruitment is likely to be much less pronounced in temperate streams (Bunn and Hughes 1997). As a result, genetic variation in North American stream insects is generally expected to follow a hierarchical pattern with the highest levels of genetic differentiation occurring at large spatial scales. Expectations for the overall extent of dispersal and the relative roles of larval and adult dispersal in temperate streams of North America are less clear.

All previously mentioned studies of stream insect dispersal have used allozymes to determine the genetic structure of stream insect populations. As functional products of nuclear genes, allozymes undergo recombination, can be subject to selection, and, as a result, tend to be less variable than non-coding regions of DNA (Mitton 1994, Niegel 1997, and Parker et al. 1998). In contrast, animal mitochondrial DNA (mtDNA) is haploid and uniparentally inherited
(usually maternally); this results in a four-fold reduction of the effective population size of mtDNA relative to nuclear DNA (nDNA), and makes mtDNA more subject to the effects of genetic drift and, ultimately, more variable than many nuclear markers, including allozymes (Avise 1994).

One of the most variable regions in both vertebrates and invertebrates is the ‘control’ region of mtDNA (‘d-loop’ in vertebrates) (Simon 1991). Because the d-loop is ideal for assessing levels of genetic differentiation at the population level, it has been widely used as a population-level marker in vertebrates. However, because length variation and high A-T content often complicate amplification and sequencing of this region, the control region has been characterized in only a few insect groups (Caccone et al. 1996, Clary and Wolstenholme 1987, Crozier and Crozier 1993, Rand and Harrison 1989, and Zhang et al. 1995), despite its potential usefulness for studying gene flow and dispersal.

The lack of recombination in the mitochondrion makes the extant mtDNA genotype a record of the matriarchal lineage of an organism (Avise 1998). Thus, mtDNA markers tell something about a population’s history, are ideal for phylogeographic, and nested clade analyses. Nested clade analysis (Templeton 1998) is a particularly powerful tool for studying geographic patterns of genetic variation because it can distinguish between population history and population structure with statistical support.

The objective of the present study was to determine patterns of gene flow and dispersal in several populations of the stonefly, *Peltoperla tarteri* (Plecoptera: Peltoperlidae) in the Appalachian Mountains of southwestern Virginia using the hypervariable control region of mtDNA. *Peltoperla tarteri* is a semivoltine stonefly that can be abundant in Appalachian headwater streams (Merritt and Cummins 1996). As a ‘shredder’, *P. tarteri* makes an important contribution to energy flow in headwater streams by comminuting organic matter into smaller particles that are used by other stream organisms. *P. tarteri* larvae are most often found in cool springs, seeps, and streams at higher elevations and adults are aerial (Merritt and Cummins 1996). Because adults and larvae occupy different habitats, the natural distribution of *P. tarteri* allows one to choose populations where adult flight is the most likely mechanism for dispersal among them.

*Peltoperla tarteri* is described as widely distributed throughout the Southern Appalachians (Stark and Kondratieff 1987). However, during three consecutive years of
sampling streams in southwestern Virginia (and a single visit to the type locality in West Virginia), large populations of *P. tarteri* were found only at the sites used in the current study while the sister species, *P. arcuata*, was abundant in many streams. The two species are not known to co-occur (Stark and Kondratieff 1987); thus, *P. tarteri* may have a more limited distribution than previously reported.

The limited distribution of *P. tarteri* combined with anecdotal evidence that stoneflies are poor fliers suggests that adult dispersal is limited in *P. tarteri*. In addition, larval dispersal is likely to be limited both within and among drainages because it requires movement through unsuitable habitats (warmer, high flow valley streams or terrestrial environments).

*Site Description*

The geography of southwestern Virginia provides a unique opportunity to study dispersal of *P. tarteri*. Because the Eastern Continental Divide is located in this area, several distinct drainages occur together in a relatively small geographic area (i.e., within 5 km of each other): the James and Roanoke drainages, which ultimately flow to the Atlantic Ocean, and the drainages of the New River, which ultimately flow to the Gulf of Mexico. Since the continental divide is a significant barrier to larval dispersal, gene flow between Gulf and Atlantic populations can be attributed to adult dispersal.

Hunters Branch, Pond Drain, War Spur, and Sartain Branch (N37° 22' W80° 31') are headwater streams located on Salt Pond Mountain in Giles County, Virginia, U.S.A. (Figure 4.1). Hunters Branch and Pond Drain flow into the New River of the Mississippi Drainage, and War Spur and Sartain Branch flow into the James River of the Atlantic Drainage. Aerial and aquatic distances between all sites ranged from 0.11 to 3.5 km and 0.13 to 5907 km, respectively (Figure 4.1).

The streams are located in northern hardwood forest at ~1300m above sea level. The riparian vegetation includes *Rhododendron* spp. and *Kalmia* spp. Average annual rainfall is 113 cm and is evenly distributed throughout the year (Cawley et al. *unpublished*). The proximity to Mountain Lake (the only significant high altitude lake in the Southern Appalachians) and questions about the lake’s origin have resulted in the availability of detailed geologic data for this area.

Colluvium, the rubble left over from the down slope mass movement of streambed matter that results from disruption of the substrate during repeated freezing and thawing (Ritter et. al.
1995), dating to the last full glaciation of the Pleistocene (the Wisconsinian) is abundant in streams in this area (Mills 1988). Colluvialization greatly increases substrate porosity and, as a result, streams on Salt Pond Mountain flow underground and resurface repeatedly despite being shown as perennial surface streams on topographic maps. In the present study, two locations had to be connected by continuous surface flow in order to be considered sites on the same stream. In several cases, this occurred for only short distances (e.g., 110 m) before streams went underground. Thus, because it was not possible to sample the entire length of a stream, some sites are fairly close together (Figure 4.1).

![Diagram of study sites in Giles, County, Virginia]

Figure 4.1: Study sites in Giles, County, Virginia. The dashed line represents the Eastern Continental Divide and dots indicate sampling locations. Throughout this manuscript, ‘lower’ and ‘upper’ site designations (e.g., Upper Pond Drain and Lower Pond Drain) refer to location on the stream relative to stream flow.

**Methods**

*Species Identification*

Positive species identifications were made with laboratory-reared adult males. Larvae with darkened wing pads were collected from streams in mid-May 1999, brought to the lab, and placed in Plexiglas rearing chambers containing stream water and a small layer of rocks and
leaves from the stream. Compressed air forced water to recirculate through the chambers. Larvae were reared in the dark at a constant temperature similar to ambient stream temperatures in May (14 ºC). Emergent adults were collected daily and preserved in 100% ethanol. Adult males were identified under a dissecting scope, and sent to B. Kondratieff (Colorado State University, Dept. of Entomology) for confirmation.

*Sample collection, tissue preservation and preparation*

A total of 215 individuals from 8 sites were used for the molecular analysis. Larvae were obtained by collecting leaf packs from the stream and sorting the samples on site. Larvae were kept alive until reaching the laboratory where they were placed in 1.5 mL microcentrifuge tubes and frozen at –80 ºC. DNA was isolated using a Gentra Systems Puregene® Genomic DNA Isolation Kit and the ‘mouse tail tissue protocol’ described in the kit instruction manual. Generally speaking, the procedure involves tissue homogenization, digestion of proteins with proteinase k, removal RNA with RNAse, isopropanol precipitation of DNA, followed by ethanol washes and resuspension of the DNA in sterile ddH2O (50µL and 100µL, respectively, for small and large DNA pellets).

*PCR amplification and sequencing of the control region*

Two stonefly-specific primers, 12s-2 (5'- GAC CCT CTA ATT AAC CAC ACC AA -3') and dl242r (5'- GAT RCT TTC CAT TTT AGT YCT AT -3') were used to amplify a 454 base pair (bp) section of the mitochondrial control region. Each 25µL PCR reaction contained 22µL ddH2O, 1µL of each primer (10µM), 1µL undiluted template, and a single Amersham Pharmacia Biotech ‘Ready To Go’ PCR bead. PCR reactions were cycled a total of 30 times in a Hybaid Express thermal cycler. The thermal cycling program was as follows: 3 min. @ 94ºC (initial denaturation), five cycles of 30 sec @ 94ºC (denature), 30 sec @ 55ºC (anneal), and 1:15 min. @ 72ºC (extension), followed by 25 cycles of 30 sec @ 94ºC, 30 sec @ 57ºC, and 1:15 min. @ 72ºC. A final extension step was carried out at 72ºC for 4 minutes. If not visualized immediately, amplification products were stored at 8ºC.

PCR products were visualized by running 3µL on a 0.8% agarose gel stained with ethidium bromide. Qiagen’s QIAquick PCR purification kit was used to prepare PCR products for direct sequencing. All PCR products were sequenced in both directions using ABI’s Big Dye Terminator kit. Each 15µL sequencing reaction contained 3µL dye terminator reaction mix, 1µL
primer (10µM), 8µL ddH2O, and 3µL template. Template dilution ranged from 1:1 to 1:6 depending upon PCR product yield.

Cycle sequencing consisted of 30 cycles of 15 sec @ 94ºC, 15 sec. @ 50ºC, and 4 min. @ 60ºC. Sequencing reactions were cleaned using Millipore MultiScreen HV Plates and sequences were run on an ABI 377 automated sequencer.

**Nuclear DNA markers**

Eleven enzyme systems (Est, Hk, Got, Idh, Lap, Mdh, Me, Mpi, Pgi, Pgm, and Tpi) and four buffers were screened in 80 animals. Pgi and Est were weakly polymorphic, other enzymes were either monomorphic or poorly resolved. Similarly, the ITS-1 region of nDNA of five individuals from each population was amplified and sequenced using the primers ITS-5 (5'-GGA AGT AAA AGT CGT AAC AAG G-3') and ITS-2 (5'-GCT GCG TTC TTC ATC  GAT GC-3'). No DNA sequence variation was detected. Thus, nuclear markers will not be discussed further.

**Sequence editing and alignment**

Sequences were viewed and edited with ABI’s EditView version 1.0.1 and contigs (the consensus sequence from the region of overlap between 5' and 3' sequences) were assembled and edited with DNASTAR SeqMan software version 4.03. Sequences were aligned by ClustalX (Thompson et al. 1997) and by eye. Some haplotypes had variable numbers of a 74 base pair repeat, thus, gaps were required to align sequences. These were the only gaps in the alignment and the insertion/deletion of the repeat was considered a single event, therefore, individual gaps were scored as 0.014 (1/74), so that the 74bp repeat was given a weight of 1.0.

**Genetic differentiation within and among drainages**

Hierarchical levels of genetic differentiation (i.e., within streams, among streams, and among drainages), population pairwise FST’s, and exact tests of population differentiation were performed in Arlequin population genetic data analysis software version 2000 (Schneider et al. 2000). Correlations between genetic and geographic distance matrices were tested using Arlequin’s Mantel Test feature. A Neighbor-Joining tree of populations was generated in PHYLIP version 3.6 alpha (Felsenstein 1989) from 100 bootstrap replicates of haplotype-frequency based estimates of Nei’s genetic distance.

**Phylogeography and Nested clade analysis**

A haplotype network was constructed by hand using the absolute number of pairwise nucleotide differences, then confirmed with Minimum Spanning Network in Arlequin and TCS
alpha version 1.0 (Templeton et al. 1992). Nested clade analysis (Templeton 1998) was used to test the null hypothesis of no association between haplotype distribution and geography. Clades were nested according to rules outlined in Templeton et al. (1992) and GeoDis version 2.0 (Posada et al. 2000) was used to test for significant associations between haplotypes and geography. Where significance was detected, the inference key of Templeton (1998) was used to determine the likely cause of the associations.

**Results**

*Patterns of genetic variation*

Of the 454 base pairs that were sequenced, eleven sites were variable (Table 4.1). Seven variable sites were located outside the repeat region, and four variable sites were within the repeat, which began at site 126 (Table 4.1). The transition to transversion ratio was 5:7; an unusual ratio for control region sequences where transitions normally outnumber transversions fifteen to one (Vigilant et al. 1991). However, overall nucleotide diversity was 0.004, which is average for non-coding regions (Li 1997).
Table 4.1: Variable sites for 21 haplotypes of the control region from 215 *P. tarteri* individuals. Dots indicate similarity to haplotype 2. The dashed line indicates the beginning of the repeat region. *P. arcuata* is the sister species of *P. tarteri* and is included for reference.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Position 2</th>
<th>14</th>
<th>20</th>
<th>30</th>
<th>66</th>
<th>87</th>
<th>117</th>
<th>172</th>
<th>246</th>
<th>284</th>
<th>358</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td>.</td>
</tr>
<tr>
<td>4</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td>.</td>
</tr>
<tr>
<td>5</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>T</td>
<td>.</td>
<td>C</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*P. arcuata* . . . . A G - - T T -

Twenty-one haplotypes were identified from the control region sequences of the 215 individuals used in this study. Fourteen haplotypes were shared between at least 2 populations, and 7 were unique (found only in a single individual) (Table 4.2). Haplotype 2 was dominant, found in 7 out of 8 sites, and had a frequency ranging from 0.22 to 0.75. The mean number of haplotypes per site (*h*) was highest in the Gulf streams (*h* = 7.75 vs. 5 for Atlantic streams). However, this was not simply an artifact of sample size because the second highest number of haplotypes was reported from Upper Pond Drain, which had half the sample size of Sartain Branch (Table 4.2).
Seven haplotypes were confined to either the Gulf or Atlantic streams. Three of these were reported in only one or two individuals, but the remaining four (Haplotypes 3, 6, 7 and 17) were common in the drainages they were reported from (Table 4.2). Most notably, haplotype 3, which accounted for 5-45% of individuals from Gulf streams, was absent from the Atlantic streams. Similarly, haplotypes 7 and 17 were found only in Sartain Branch where (combined) they accounted for ~30% of individuals.

Table 4.2: Haplotype frequencies by site. Numbers in parentheses indicate sample sizes. HB = Hunters Branch, UHB = Upper Hunters Branch, PD = Pond Drain, UPD = Upper Pond Drain, SB = Sartain Branch, LSB = Lower Sartain Branch, WS = War Spur, LWS = Lower War Spur. Green and blue indicate haplotypes found only in the Gulf and Atlantic drainages, respectively.

<table>
<thead>
<tr>
<th>Site</th>
<th>HB (54)</th>
<th>UHB (20)</th>
<th>PD (20)</th>
<th>UPD (20)</th>
<th>SB (39)</th>
<th>LSB (20)</th>
<th>WS (22)</th>
<th>LWS (20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>haplotype</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.36</td>
<td>0.30</td>
<td>0.15</td>
<td>0.08</td>
<td>0.09</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.22</td>
<td>0.35</td>
<td>0.40</td>
<td>0.46</td>
<td>0.45</td>
<td>0.67</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.22</td>
<td>0.45</td>
<td>0.20</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.06</td>
<td>0.10</td>
<td>0.10</td>
<td>0.25</td>
<td>0.20</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.04</td>
<td>0.10</td>
<td>0.10</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.02</td>
<td>0.10</td>
<td>0.05</td>
<td>0.15</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td>12</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.02</td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.02</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td>0.18</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Genetic differentiation within and among drainages**

Genetic differentiation within sites accounted for most of the variation, followed by differentiation among streams within drainages (Table 4.3). Levels of genetic differentiation both within streams and among drainages were not significant. However, the level of genetic differentiation among streams within drainages was high ($F_{ST} = 0.15; p < 0.001$; Table 4.3) and may explain the lack of significant differentiation at the drainage level.

Table 4.3: Partitioning of genetic variance according to AMOVA analysis (Weir and Cockerham 1984). Variance components were used to calculate F statistics and significance was tested using 16002 non-parametric permutations. Asterisk indicates significance ($p < 0.001$), NS = not significant.

<table>
<thead>
<tr>
<th></th>
<th>Within sites</th>
<th>Among sites within streams</th>
<th>Among streams within drainages</th>
<th>Among drainages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall $F_{ST}$</td>
<td>0.15*</td>
<td>0*</td>
<td>0.14*</td>
<td>0.02 NS</td>
</tr>
<tr>
<td>% of variation</td>
<td>84</td>
<td>0</td>
<td>14</td>
<td>2</td>
</tr>
</tbody>
</table>

High levels of differentiation among streams within drainages were also indicated by the population pairwise $F_{ST}$ comparisons and exact tests of non-differentiation. With few exceptions, (4 and 2 out of 28 comparisons for the pairwise $F_{ST}$ and exact tests, respectively) sites were significantly differentiated from all others except those on the same stream (Table 4.4). For the most part, this was true regardless of whether sites were located in the same drainage; however, some populations in different drainages were not significantly differentiated from each other (e.g., Pond Drain and War Spur; Table 4.4).

Population pairwise estimates of $N_m$ ranged from 1.1 to infinity (Table 4.4). Although all estimates of $N_m$ values were above the level necessary for gene flow to overcome the effects of genetic drift (> 1 migrant every other generation; Slatkin 1985), there was still significant genetic variation among almost all sites (Table 4.4).
Table 4.4: Above diagonal: Pairwise $F_{ST}$ values estimated from Tajima-Nei distances and 16002 permutations. * = Significant result from permutation test ($p < 0.05$), e = significant pairwise comparisons according to exact tests of non-differentiation calculated with 6000 dememorisation and 100,000 Markov chain steps (Schneider et al. 2000). Below diagonal: Population pairwise $N_m$ values calculated from Slatkin’s linearized $F_{ST}$’s (Slatkin 1985).

<table>
<thead>
<tr>
<th>Site</th>
<th>HB</th>
<th>UHB</th>
<th>PD</th>
<th>UPD</th>
<th>SB</th>
<th>LSB</th>
<th>WS</th>
<th>LWS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HB</td>
<td></td>
<td>0.006</td>
<td>0.02</td>
<td>0.09*</td>
<td>0.21*</td>
<td>0.28*</td>
<td>0.05*</td>
<td>0.03*</td>
</tr>
<tr>
<td>UHB</td>
<td>83.5</td>
<td></td>
<td>0.04</td>
<td>0.11*</td>
<td>0.22*</td>
<td>0.30*</td>
<td>0.13*</td>
<td>0.12*</td>
</tr>
<tr>
<td>PD</td>
<td>20.5</td>
<td>12.3</td>
<td></td>
<td>0.01</td>
<td>0.14*</td>
<td>0.20*</td>
<td>0.02*</td>
<td>0.03*</td>
</tr>
<tr>
<td>UPD</td>
<td>5.1</td>
<td>4.1</td>
<td>56.6</td>
<td></td>
<td>0.07*</td>
<td>0.10*</td>
<td>0.07*</td>
<td>0.08*</td>
</tr>
<tr>
<td>SB</td>
<td>1.8</td>
<td>1.8</td>
<td>3.1</td>
<td>6.7</td>
<td></td>
<td>0.0*</td>
<td>0.17*</td>
<td>0.18*</td>
</tr>
<tr>
<td>LSB</td>
<td>1.3</td>
<td>1.2</td>
<td>2.1</td>
<td>4.6</td>
<td>∞</td>
<td></td>
<td>0.26*</td>
<td>0.28*</td>
</tr>
<tr>
<td>WS</td>
<td>9.6</td>
<td>3.2</td>
<td>21.3</td>
<td>6.9</td>
<td>2.4</td>
<td>1.4</td>
<td></td>
<td>-0.04</td>
</tr>
<tr>
<td>LWS</td>
<td>13.9</td>
<td>3.6</td>
<td>16.6</td>
<td>6.0</td>
<td>2.6</td>
<td>1.3</td>
<td>∞</td>
<td></td>
</tr>
</tbody>
</table>

The Neighbor Joining tree indicates that the Gulf populations (HB, UHB, PD and UPD) group together into a well-supported clade (Figure 4.2). However, the Atlantic populations (WS, LWS, SB and LSB) do not group together, and Lower War Spur groups more closely with the Gulf streams (HB and PD) than with Sartain Branch.
Figure 4.2: Neighbor-joining tree indicating genetic differentiation among populations. Numbers indicate bootstrap values (100 replicates). Green and blue indicate Gulf and Atlantic Drainages, respectively. The Gulf streams form a well-supported group. However, the War Spur populations grouped with the Gulf streams rather than with Sartain Branch. Blue and green indicate Atlantic and Gulf streams, respectively.

The high levels of genetic differentiation among streams within drainages, lack of differentiation among drainages, and close association of the Gulf streams and War Spur (as indicated by the NJ tree) suggest that contemporary gene flow and dispersal may not fully explain the observed patterns of genetic structure. If larval dispersal is predominant, low levels of genetic differentiation within drainages, high levels among drainages, and a significant correlation between genetic and aquatic distance are expected. Conversely, if dispersal is mainly by adults, low levels of genetic differentiation within and among drainages and a significant correlation between genetic and aerial distance are expected.

However, the Mantel test indicated that there was no significant correlation between either aerial or aquatic distance and genetic distance (as pairwise $F_{ST}$'s) (Figure 4.3). Therefore,
the patterns of genetic structure observed in \textit{P. tarteri} are more complex than would be predicted if dispersal-mediated gene flow was the primary determinant of population genetic structure.

![Figure 4.3](image_url)

**Figure 4.3:** Pairwise FST’s vs. aerial and aquatic distances (km). Distances were log transformed to allow plotting all pairwise comparisons on a single graph. There was no significant correlation between genetic distance (as FST) and either aerial or aquatic distance (p >> 0.05 for Mantel tests with 1000 permutations).

**Phylogeography and nested clade analysis**

To determine the historical relationships among haplotypes, a network was constructed using the 21 haplotypes shown in Table 4.1. The resulting network had many ‘closed’ loops (indicating homoplasy; see Smouse 1998) and a maximum parsimony tree based on these data was uninformative (Figure 4.4).
Figure 4.4: Unrooted phylogram indicating strict consensus of 654 most parsimonious trees of 21 mtDNA control region haplotypes. Trees were constructed with a general heuristic search using simple addition of sequences and tree bisection-reconnection branch swapping. Numbers indicate haplotypes. Most relationships between haplotypes were unresolved.

However, the presence of variable numbers of repeats in the control region and lack of recombination in the mitochondrion has led others (e.g., Zhang et al. 1995) to suggest that mechanisms other than gene flow may be responsible for genetic variation within repeats in the control region. Thus, a haplotype network that included only sites outside the repeat (sites 1-125; Table 4.1) was constructed; this resulted in a reduction in the overall number of haplotypes from 21 to 8 with haplotypes 1-3 remaining the dominant haplotypes (Table 4.5; Figure 4.5).
Table 4.5: Haplotype frequencies used in the nested clade analysis. Exclusion of variation in the repeat reduced the overall number of haplotypes to eight. Numbers in parentheses indicate sample sizes. HB = Hunters Branch, UHB = Upper Hunters Branch, PD = Pond Drain, UPD = Upper Pond Drain, SB = Sartain Branch, LSB = Lower Sartain Branch, WS = War Spur, LWS = Lower War Spur.

<table>
<thead>
<tr>
<th>Site</th>
<th>HB</th>
<th>UHB</th>
<th>PD</th>
<th>UPD</th>
<th>SB</th>
<th>LSB</th>
<th>WS</th>
<th>LWS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(54)</td>
<td>(20)</td>
<td>(20)</td>
<td>(20)</td>
<td>(39)</td>
<td>(20)</td>
<td>(22)</td>
<td>(20)</td>
</tr>
<tr>
<td>Haplotype 1</td>
<td>0.37</td>
<td>0.30</td>
<td>0.15</td>
<td>0.10</td>
<td>0.08</td>
<td>0.14</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Haplotype 2</td>
<td>0.39</td>
<td>0.25</td>
<td>0.60</td>
<td>0.80</td>
<td>0.59</td>
<td>0.65</td>
<td>0.77</td>
<td>0.75</td>
</tr>
<tr>
<td>Haplotype 3</td>
<td>0.24</td>
<td>0.45</td>
<td>0.20</td>
<td>0.10</td>
<td>0.15</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haplotype 7</td>
<td>0.15</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haplotype 10</td>
<td>0.045</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haplotype 11</td>
<td>0.045</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haplotype 14</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haplotype 17</td>
<td>0.18</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some would argue that ‘pruning’ the data by excluding the repeat results in a less reliable haplotype network because the ‘true’ data set was not used. However, Wills (1996) demonstrated that pruning of human control region data using a ‘topiary pruning’ model resulted in a closer approximation of the ‘true’ tree, which led him to suggest that ‘unpruned’ (homoplasic) data sets are distorted because homoplasies make it difficult to discern the true relationship between haplotypes.

In the present study, the results of the AMOVA, Mantel, and pairwise tests did not change when the ‘new’ data set was used (results available from the author upon request). This strongly suggests that variation in the repeat is not an important determinant of population structure but a source of homoplasy that confounds phylogeographic analysis (Figure 4.4). Thus, exclusion of the repeat region probably increased the likelihood of determining the true relationship between haplotypes.

Numerous studies have shown that haplotypes located at the tips of a network (tip clades) are younger than internal haplotypes (internal clades) (Templeton et al. 1995). In the present analysis, all of the tip clades are either found in a single stream (Haplotypes 10, 11, 14, 7, and 17; Figure 4.5) or are confined to a single drainage (Haplotype 3). This arrangement
suggests that Haplotypes 1 and 2 are ancestral haplotypes, and 3, 10, 11, 14, 17 and 7 are relatively younger. The Atlantic clades (SB and WS) appear to have arisen independently (Figure 4.5), which may reflect multiple foundings or other historical processes.

Figure 4.5: Haplotype network showing 95% most probable connections as calculated by TCS alpha. Large numbers inside circles indicate haplotype number, and small numbers indicate sample sizes. Haplotypes 10 and 14 were found in only a single individual; thus no sample size is indicated. Sizes of circles are not to scale, but do represent relative sample sizes (e.g., largest and smallest circles contain the most and least individuals, respectively).

Because historical processes can have a profound effect on genetic structure, they are often invoked to explain the geographic distribution of haplotypes in a network. The usual ‘phylogeographic’ approach is to overlay a haplotype network onto the geographic sampling area and infer explanations for the observed geographic arrangement through visual inspection. However, this approach lacks statistical support.

Nested clade analysis (Templeton 1998) allows for statistical testing of the null hypothesis of no association between haplotype and geographic location. In nested clade analysis, a haplotype tree is used to define a nested series of branches (clades) that are used in the evolutionary analysis of the spatial distribution of genetic variation (Templeton 1998). The nested clade information (location: interior vs. tip), sample size for each haplotype, and geographic location of each clade (latitude and longitude) are entered into the computer program GeoDis (Posada et al. 2000) which tests for significant associations between haplotype and geographic location. Where statistically significant associations are detected, a set of criteria
(outlined in the ‘inference key’; Templeton 1998) is used to discriminate between the effects of population structure (e.g., gene flow) and population history (e.g., range expansion).

The nested cladogram comprises three 1-step clades (1-1, 1-2, and 1-3) and the total cladogram (4.6). Nested clades 1-1 and 1-2 are tip clades, and clade 1-3 is an interior clade.

![Total Cladogram Diagram]

Figure 4.6: Nested clade diagram of haplotype network shown in Figure 4.5. Clades are nested according to rules outlined in Templeton et al. (1995). Haplotypes are 0-step clades, solid lines and italics indicate 1-step clades, and the dotted line encompasses the total cladogram.

Statistically significant associations were detected at several nested clade levels (Table 4.6). According to the inference key in Templeton (1998), each association had a unique explanation. At the 0-step clade level, significant associations were explained by fragmentation (Haplotype 1) and restricted gene flow with isolation by distance (Haplotypes 2 and 3) (Table 4.5). At the 1-step clade level, they were explained by contiguous range expansion (Haplotypes 7 and 17) (Table 4.6).
Table 4.6: Results of GeoDis analysis showing clade (Dc), nested (Dn), and interior-tip clade (I-T) distances. Gray highlights indicate associations that are significant at the 5% level. L= Significantly large, S=Significantly small. ‘Chain of inference’ refers to key in Templeton (1998), numbers indicate steps followed in the inference key to reach the conclusion of population fragmentation (FR), restricted gene flow with isolation by distance (IBD), or range expansion (RE). Dc, Dn, and I-T represent the geographic dispersion of a clade, the distance between all individuals with Haplotype X from all those with Haplotype Y, and the average distance between tip and interior clades, respectively.

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>1-Step Clades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>No.</td>
</tr>
<tr>
<td>Int.</td>
<td>1</td>
</tr>
<tr>
<td>Int.</td>
<td>10</td>
</tr>
<tr>
<td>Tip</td>
<td>11</td>
</tr>
<tr>
<td>I-T</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chain of Inference</th>
<th>1-2-3-4-9-No: FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tip</td>
<td>1-1</td>
</tr>
<tr>
<td>Int.</td>
<td>17</td>
</tr>
<tr>
<td>I-T</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location</th>
<th>No.</th>
<th>Dc</th>
<th>Dn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tip</td>
<td>7</td>
<td>0.22</td>
<td>0.30</td>
</tr>
<tr>
<td>Int.</td>
<td>17</td>
<td>0.28</td>
<td>0.30</td>
</tr>
<tr>
<td>I-T</td>
<td></td>
<td>0.05</td>
<td>-0.006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chain of Inference</th>
<th>1-2-3-4-9-No: FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tip</td>
<td>1-2</td>
</tr>
<tr>
<td>Int.</td>
<td>1-3</td>
</tr>
<tr>
<td>I-T</td>
<td></td>
</tr>
</tbody>
</table>

Discussion

The results shown in Table 4.6 suggest that both historical processes and population structure have played an important role in determining the geographic pattern of genetic variation in *Peltoperla tarteri*. The conclusion of fragmentation is based on the restricted geographic distributions of Haplotypes 1, 10 and 11 (Table 4.6). Haplotype 1 is present in 7 of 8 sites, but in lower frequencies in the Atlantic populations (Table 4.5). The geographic distribution of Haplotype 1 accounts for the significantly small Dc. According to the inference key in Templeton (1998), this pattern follows predictions under historical fragmentation and suggests allopatric fragmentation of the streams on Salt Pond Mountain.
Several processes could have resulted in allopatric fragmentation of *P. tarteri* populations on Salt Pond Mountain: 1) colluvialization during the last full glaciation of the Pleistocene (the Wisconsinian; 30,000-13,000 years BP), 2) seismic activity, and 3) habitat fragmentation due to climatic oscillations or ‘microvicariance’ (Templeton et al. 1995).

Colluvium from Pond Drain has been dated to the Wisconsinian (Mills 1988) and has undoubtedly affected the streambed and increased its porosity (see Site Description). However, colluvial boulders have not moved great distances and tend to be found in the valley from which they originated (Mills 1988, Cawley et al. *unpublished*). Thus, it is unlikely that colluvialization played a substantial role in the allopatric fragmentation of streams on Salt Pond Mountain.

Seismic activity has been reported in the area and a fault underlying Mountain Lake was recently discovered (Cawley et al. *unpublished*). However, all the streams in the present study are north of the fault which runs SE to NW. Therefore, while a possible role of historical seismic events cannot be ruled out, the orientation of the fault does not explain the fragmented distribution of Haplotype 1.

Palynological studies indicate that, during the Wisconsinian, much of the unglaciated Southern Appalachians above 1200m was dominated by arctic tundra (Delcourt and Delcourt 1998 and Godt et al. 1996). Today, high elevation bogs are rare in the Southern Appalachians, and are thought to be relicts formed from tundra as the climate warmed (Pitillo 1994). At one time, a large bog(s) may have covered much of the higher elevations of Salt Pond Mountain, which is relatively flat-topped, and a more or less continuous habitat may have been available for *P. tarteri*. As the climate became drier, habitat ‘islands’ may have been formed from shrinking wetlands.

Although only a few small bogs (e.g., Little Spruce Bog) remain there today, the flora and fauna of Salt Pond Mountain suggest that more extensive wetlands may have been present in the past. Many salamander species are known from Salt Pond Mountain and in some areas the forest understory is dominated by two very large fern species that are usually found in wet marshy woods, swamps, ditches, and stream banks (cinnamon and interrupted fern; *Osmunda cinnamomea* and *O. claytoniana*). Thus, ‘microvicariance’ is the most likely explanation for the fragmented pattern of genetic variation among the streams of Salt Pond Mountain. Climatic oscillations and subsequent habitat fragmentation during the Wisconsinian have been implicated in the distribution and/or genetic diversity of many organisms, including the stoneflies of the
Black Hills of South Dakota and Wyoming, (Huntsman et al. 1999), Appalachian salamanders (Tilley 1997), and rare plants of the Southern Appalachians (Godt et al. 1996).

The genetic variation in Sartain Branch fits the pattern predicted for range expansion (Table 4.6). Under contiguous range expansion, some older haplotypes from the source population will be confined to the pre-expansion area and, as a result, will have significantly small $D_c$s and $D_n$s. Younger haplotypes that arose in the expanding populations will be more restricted geographically and will have significantly large $D_c$s and $D_n$s (Templeton 1998).

Haplotypes 7 and 17, which were found only in Sartain Branch, show the pattern predicted for younger haplotypes, while Haplotype 3, which was found only in Gulf populations, shows the pattern described for ancestral haplotypes (Table 4.6; Figures 4.5 and 4.6). These results indicate that a Gulf source population expanded into Sartain Branch. The most likely candidate is Hunters Branch since its headwaters are very close to the headwaters to Sartain Branch (Figure 4.1).

The genetic effects of fragmentation and range expansion are overlaid upon a contemporary population structure that is influenced by isolation by distance with restricted gene flow (Table 4.6). Under recurrent, restricted gene flow, the geographic extent of a haplotype is strongly correlated with its age (Templeton et al. 1995 and references therein). Ancestral clades will be widely distributed and have significantly large $D_c$s and $D_n$s, and younger clades will initially reside within the range of the ancestral haplotypes and have significantly small $D_c$s and $D_n$s (Templeton 1998). Hence, Haplotype 3 has a limited distribution because it is relatively young and has not spread far due to restricted gene flow while Haplotype 2 is an ancestral haplotype that has had many generations to become widespread (Tables 4.5 and 4.6). It is important to note that the terms ‘ancestral’ and ‘young’ refer to the historical event being considered. Thus, Haplotype 3 is ancestral relative to the range expansion into Sartain Branch, but is relatively young when considered in the context of allopatric fragmentation.

Several conclusions about dispersal of *P. tarteri* can be drawn from the nested clade analysis. First, in order for the patterns caused by fragmentation and range expansion to be maintained, contemporary gene flow must be limited. Second, expansion of Gulf populations into Sartain Branch without subsequent contact suggests that adult dispersal is a rare over short distances. Anecdotal evidence that supports this conclusion includes the limited distribution of *P. tarteri* and the observation that when adults attempt to evade capture, they often crawl...
(sometimes falling as far as 1 m to the ground) rather than fly away. Finally, the distinct evolutionary trajectories of Sartain Branch and War Spur (Figure 4.5) suggest low levels of gene flow between these streams; thus, larval dispersal of *P. tarteri* is also limited.

Results from the pairwise F<sub>ST</sub> analysis and exact tests of population differentiation largely agree with the conclusions from the nested clade analysis. In pairwise comparisons, most sites were significantly differentiated from all others except those on the same stream (Table 4.4). Although this was most pronounced in the Atlantic streams, the overall conclusion from the pairwise tests is that both larval and adult dispersal are limited in *Peloperla tarteri*. However, since only 2% of the total variation was attributed to drainages and most of the variation was found at smaller spatial scales, one might conclude from the AMOVA analysis that adult dispersal is widespread (Table 4.3).

In addition, if only the F<sub>ST</sub> based analyses were used, some results that would be difficult to interpret. For example, low levels of differentiation between Pond Drain and War Spur (the most distant sites in the study) and much higher levels between Hunters Branch and Sartain Branch (the shortest distance between sites on different streams) are not readily explained by dispersal alone. Interpretation of these data is further confounded by the lack of significant correlations between aquatic and aerial distance according to the Mantel test (Figure 4.3). We know from the nested clade analysis that fragmentation and range expansion have had a significant impact on the genetic structure of these populations. However, because F<sub>ST</sub> tests cannot distinguish between the effects of population structure and population history, interpreting these data would be difficult without the aid of nested clade analysis.

Nonetheless, there are some benefits to using F<sub>ST</sub> based analyses. A primary advantage is that the hierarchical nature of F<sub>ST</sub> and exact tests makes it possible to determine the magnitude of genetic differentiation between populations. For example, using nested clade analysis, the only conclusion that can be made about larval dispersal is that it is limited (see above), while the F<sub>ST</sub> analysis gives an indication of the scale at which movement by larvae is an effective dispersal mechanism. The pairwise F<sub>ST</sub> tests suggest that larval dispersal is common at distances ranging from 0.11 to 2.5 km (range between sites where larval dispersal was possible and genetic differentiation was not detected; see Figure 4.1 and Table 4.4) but not at distances greater than 2.8 km (minimum distance at which significant genetic differentiation was detected between sites where larval dispersal was possible; see Figure 4.1 and Table 4.4).
A further advantage of hierarchical $F_{ST}$ analysis is that the relative importance of adult and larval dispersal in the maintenance of $P. \text{tarteri}$ populations can be inferred from this approach. Even though the Eastern Continental Divide separates them, the headwaters of Hunters Branch and Sartain Branch are very close together (Figure 4.1). Thus, the distance between sites HB and SB (0.71 km) is very similar to the distance between the two sites on Sartain Branch (0.68 km).

Since adult flight is the primary way in which $P. \text{tarteri}$ can move between Hunters Branch and Sartain Branch, comparing the levels of genetic differentiation between HB and SB to that of SB and LSB allows one to determine the extent of larval and adult dispersal over similar geographic distances. The high level of genetic differentiation between HB and SB ($F_{ST} = 0.21$; Table 4.4) compared to the lack of differentiation between SB and LSB ($F_{ST} = 0$; Table 4.4) indicates that larval movement is the primary mechanism for short-distance dispersal of $P. \text{tarteri}$.

Thus, the most complete picture of dispersal in $P. \text{tarteri}$ is obtained when nested clade and hierarchical $F_{ST}$ analyses are combined. Together, these analyses lead to the following conclusions about dispersal in $\text{Peltoperla tarteri}$: 1) Contemporary gene flow is sufficiently restricted so that historical processes still have a profound effect on population genetic structure. 2) Adult dispersal is extremely rare but can be an effective mechanism for dispersal. 3) Larval dispersal is the primary mechanism for short distance (< 3 km) dispersal, and 4) Long distance dispersal is infrequent and probably the result of adult movement.

Western Montana populations of $\text{Yoraperla brevis}$ (Plecoptera: Peltoperlidae) showed patterns of genetic differentiation similar to those observed in $P. \text{tarteri}$. In $Y. \text{brevis}$, sites within streams were not significantly differentiated and there were high levels of genetic differentiation among streams (Hughes et al. 1999). According to Hughes et al. (1999), this pattern results from adult dispersal that is confined to the natal stream system. Larval dispersal was ruled out because the pattern of genetic variation did not fit expectations under unidirectional dispersal and Hughes et al. (1999) felt that larvae were unlikely to move large distances upstream. However, our results indicate that Appalachian peltoperlid larvae can disperse up to 3 km, and since $P. \text{tarteri}$ and $Y. \text{brevis}$ are members of the same family and are similar in size and ecology, it is likely that larval dispersal, not adult flight, is at least in part responsible for the genetic patterns observed in $Y. \text{brevis}$. 
Although the overall $F_{ST}$ among streams for *Yoraperla brevis* was very similar to that of *P. tarteri* (0.156 and 0.14, respectively), there is considerably more geographic distance between the sites used in the *Y. brevis* study. Thus, *P. tarteri* has the same level of genetic differentiation as *Y. brevis* at a much smaller spatial scale. This may be an artifact of the genetic markers that were used since Hughes et al. (1999) used allozymes to determine levels of genetic differentiation in *Y. brevis*, while the current study used a hypervariable region of mtDNA.

However, differences in genetic structure between *Y. brevis* and *P. tarteri* may also be a result of the unique population histories of the two species. The study areas Hughes et al. (1999) used were fully glaciated during the Wisconsinian; thus, the *Y. brevis* populations were established post-glaciation. Numerous examples from other organisms have shown that populations from previously glaciated areas tend to have lower levels of genetic variation than their unglaciated counterparts (e.g. freshwater fishes, Bernatchez and Wilson 1998; salamanders, Tilley 1997; and terrestrial insects, Reiss et al. 1999). Thus, population history may explain the smaller spatial scale at which significant genetic differentiation was detected in *Peltoperla tarteri*.

The only other North American stonefly for which genetic data are available is *Pteronarcys proteus* (Plecoptera: Pteronarcyidae). The large size of *P. proteus* and low overall $F_{ST}$ (0.04 using allozymes; White 1989) suggest that adult flight is a more effective dispersal mechanism for *P. proteus* than for the smaller peltoperlid stoneflies.

Finally, it is possible that the genetic patterns of *Peltoperla tarteri* and *Yoraperla brevis* are unique to headwater systems. If higher order streams serve as downstream barriers to larval dispersal by stream insects, headwater systems may represent ‘island’ habitats. However, further study is needed to determine whether the pattern of genetic variation observed in *Peltoperla tarteri* is common among headwater organisms and to improve our understanding of dispersal and recruitment of stream insects in temperate streams of North America.
Chapter 5: Conclusions

In conclusion, this study has provided several insights into the molecular population genetics and ecology of the stonefly, *Peltoperla tarteri*:

1) The arrangement of the control region in *P. tarteri* is similar to that described for other insect taxa. The presence of conserved structural elements in Plecoptera, Orthoptera, and Diptera strongly suggest that structural elements, rather than primary sequences, are conserved throughout the control region of insect mitochondrial DNA.

2) The possibility of recurrent mutations and convergent evolution in tandem repeats of the *P. tarteri* control region could potentially confound evolutionary studies. Thus, the repeat region should be avoided in studies of population history.

3) Populations of *Peltoperla tarteri* located in southwestern Virginia are semivoltine. This life cycle is longer than that reported for populations of *P. tarteri* in West Virginia. The slower development time in southwestern Virginia is attributed to the colder climate at the study sites.

4) Low levels of genetic differentiation among cohorts suggests that larval development of *P. tarteri* is not fixed and that some individuals complete development in one or three years rather than two. This ‘cohort splitting’ probably results in individuals from distinct cohorts maturing at the same time and mating with one another.

5) Contemporary gene flow is sufficiently restricted so that historical processes still have a profound effect on the population genetic structure of *Peltoperla tarteri*. Restricted gene flow probably reflects the limited realized dispersal of this stonefly as both larvae and adults.

6) Although adult dispersal is an effective dispersal mechanism, it is extremely rare. Therefore, larval dispersal is thought to be the primary mechanism for short distance (< 3 km) dispersal of *P. tarteri*. Long distance dispersal is infrequent and probably the result of adult movement.
Literature Cited


Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin ver. 2.000: A software for population genetics analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.


Alicia Slater Schultheis  
Department of Biology  
Virginia Polytechnic Institute and State University  
Blacksburg, VA 24061-0406  
(540) 231-9053  
aschultheis@vt.edu

BORN: 13 May 1970  Knoxville, Tennessee

ACADEMIC DEGREES:  
B.S. Biology, 1994  
Georgia Institute of Technology  
GPA: 3.7/major

M.S. Biology, 1996  
Thesis: “Recovery of macroinvertebrate community structure and function following use of a constructed wetland to abate copper pollution”  
Virginia Polytechnic Institute and State University  
GPA: 4.0

Ph.D. Biology, December 2000  
Dissertation: “Gene flow and dispersal among populations of the stonefly, Peltoperla tarteri (Plecoptera: Peltoperlidae) in the Southern Appalachians”  
Virginia Polytechnic Institute and State University  
GPA: 4.0

HONORS AND AWARDS:  
2000 International Postdoctoral Research Fellowship  
National Science Foundation  
2000 University Commendation Teaching Award for Outstanding Teaching as a Graduate Teaching Assistant, Virginia Tech  
2000 NABS Endowment Travel Award (based on abstract quality)  
1999 Wildco Award “Best Oral Presentation in Basic Research” at North American Benthological Society Meeting, Duluth, Minnesota  
1999 NABS Endowment Travel Award (based on abstract quality)  
1998 Departmental Nominee, University Graduate Student Teaching Award  
1997 Departmental Nominee, University Graduate Student Service Award  
1996 Hobbs Scholarship, Mountain Lake Biological Station, University of Virginia
**CURRENT RESEARCH INTERESTS:** Dispersal, population genetics and molecular ecology of stream insects, PCR and DNA sequencing, effect of habitat fragmentation on dispersal of adult aquatic insects, conservation genetics, stream ecology.

**RESEARCH EXPERIENCE:**

Virginia Polytechnic Institute and State University
Ph.D. candidate (August 1996-present)
Advisor: Dr. Albert C. Hendricks
Developed technique for amplifying and sequencing hypervariable region of mitochondrial DNA, studied relative roles of adult and larval dispersal and life history of stoneflies.

Advisor: Dr. Albert C. Hendricks
Responsibilities included: basic water quality analyses, heavy metal and sulfate detection, macroinvertebrate survey and leaf decomposition experiments, analysis and publication of results, ordering equipment and supplies, maintenance of equipment), supervision of undergraduates

Georgia Institute of Technology
Undergraduate Research Assistantship
June 1993-March 1994
Advisor: Dr. Edward Yeargers
Responsibilities included: maintenance of *C. elegans* culture, data recording and analysis of survival of *C. elegans* following exposure to gamma radiation

**PUBLICATIONS:**


Schultheis, A.S. *in prep*. Genetic evidence for overlapping cohorts in the semivoltine stonefly, *Peltoperla tarteri* (Plecoptera: Peltoperlidae. To be submitted to Freshwater Biology


**LIMITED DISTRIBUTION**


**PUBLICATIONS:**

**RESEARCH GRANTS:**

International Postdoctoral Research Fellowship

National Science Foundation

July 2000 $53,808

Virginia Academy of Science

May 1999 $1750

Graduate Research Development Project

Virginia Polytechnic Institute and State University

Jan 1997: $800

Jan 1998: $500

Aug. 1999: $1000

Graduate Student Association Travel Fund

Virginia Polytechnic Institute and State University

June 1997 $250

June 1999 $300

June 2000 $300

Sigma Xi

May 1997 $1100

Jan. 1998 $1750

Virginia Dept. of Mines, Minerals, and Energy

Nov. 1997 $4500


Slater, A. Recovery of macroinvertebrate community structure and leaf processing following use of a constructed wetland to abate copper pollution. Virginia Polytechnic Institute and State University. May 1996. Georgia Institute of Technology, Atlanta, GA. May 1996. (Invited seminar)

INVITED SEMINARS:


TEACHING EXPERIENCE:
Virginia Polytechnic Institute and State University
Graduate teaching assistantships
Freshwater Ecology Laboratory
Fall 1996, Fall 1997-Spring 1999
Average Evaluation: 3.7/4.0
Because I have taught this lab with several different professor, my responsibilities have ranged from co-teaching the lab with the professor to ultimate responsibility for design and implementation of the lab and occasionally lecturing for the professor.

*Responsibilities:* Lab set-up, ordering and maintenance of supplies and equipment, organizing and leading field trips, mentoring undergraduate and graduate students during research projects, and guiding student development of scientific-meeting style presentations using PowerPoint. Designed and implemented laboratory exercises including: freshwater invertebrate collection, bathymetric map construction, and presentation and defense of research proposals to a panel modeled after the NSF system. Also designed and maintained laboratory web page: www.biol.vt.edu/department/graduate/gradpages/tapages/schultheis/freshwater which has class materials downloadable as PDF files, and allows students to check their current grades by entering their ID numbers.

Principles of Biology Laboratory
Fall 1994-Spring 1995, Spring & Summer 1997
Average Evaluation: 3.5/4.0
*Responsibilities:* Pre-lab lectures and demonstrations, supervision of students during lab sessions, group projects, and oral presentations, grading, and other miscellaneous tasks related to course administration.
Honors Biology Laboratory
Fall 1995-Spring 1996
Average Evaluation: 3.8/4.0

Responsibilities: I had ultimate responsibility for the laboratory portion of this course. Standard duties were as listed above under Freshwater Ecology and Principles of Biology Laboratory. As part the Honors course, I developed a lab exercise that used molecular techniques (PCR and RFLP’s) to help freshman broaden their definition of a species.

PROFESSIONAL SERVICE:

University GTA workshop August 1997, 1998, 1999
Moderator, “GTA Dual Role Panel”
Presenter, “Teaching Life Sciences Labs” (1998 and 1999 sessions received highest overall evaluation 3.9/4.0)

Endowment Committee, North American Benthological Society
May 1998-present

Graduate Resources Committee, North American Benthological Society
May 1998-present

Commencement Committee, Virginia Tech
Graduate Student Representative August 1999-present

Commission on Faculty Affairs, Virginia Tech
Graduate Student Representative August 1998-May 1999

Biology Graduate Student Association, Virginia Tech;
President, December 1996 - December 1997
Treasurer, December 1995 - December 1996
GSA Delegate, December 1994 - December 1995

Virginia Junior Academy of Science
Judge, Environmental Science Student Paper Competition.
June 1997.

Biology Dept Virginia Tech

PROFESSIONAL MEMBERSHIPS:

North American Benthological Society
International Society of Plecopterologists
Sigma Xi
Virginia Academy of Sciences
Society for Comparative and Integrative Biology
Phi Kappa Phi Honor Society
Biology Graduate Student Association