2010-12-02

Gene Flow and Dispersal of the Flatworm, *Polycelis coronata*: A Multiscale Analysis

Jeffrey N. Moore
Brigham Young University - Provo

Gene Flow and Dispersal of the Flatworm, *Polycelis Coronata*:

A Multiscale Analysis

Jeffrey N. Moore

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Russell B. Rader
Jerald B. Johnson
Lee F. Braithwaite

Department of Biology

Brigham Young University

December 2010

ABSTRACT

Gene Flow and Dispersal of the Flatworm, *Polycelis Coronata:*

A Multiscale Analysis

Jeffrey N. Moore

Department of Biology

Master of Science

We determined genetic variance and gene flow across multiple scales (reaches, headwater segments, and catchments) to examine the dispersal ability of the flatworm *Polycelis coronata* along the Wasatch Mountains of Utah. Multiple models predict patterns of genetic differentiation in stream invertebrates based on dispersal traits and the spatial connectivity of the habitat. The stream hierarchy model predicts genetic differentiation to be low and gene flow to be high between reaches nested in segments, moderate among segments within catchments, and differentiation to be highest and gene flow lowest among catchments, whereas the headwater model predicts the greatest differentiation between headwater segments. Our objective was to determine which model best described genetic patterns observed in *P. coronata*. Using a nested hierarchical sampling design ensured that if limitations to dispersal had an effect on genetic differentiation, we would be able to identify at what scale these processes operate. We hypothesized genetic variation would be small within headwater segments and reach maximum levels between headwater segments with no increase in differentiation with increasing distance between headwater patches or between drainages. We do not expect high dispersal along the stream network or across the terrestrial environment (actively or passively).

We generated DNA sequence data (mitochondrial COI) from 50 sites nested within 24 segments, which were nested in four adjacent catchments. We identified 134 haplotypes from 506 individuals using a 763 bp fragment of mtDNA.

Genetic patterns did not conform to the SH model. Evidence from one drainage (Provo River) was consistent with the headwater model. However, high differentiation within sites suggested that the genetic patterns we uncovered may be representative of high ancestral polymorphism among pre-fragmented populations that were historically widespread. Large effective population sizes and no evidence of bottleneck events suggest incomplete lineage cannot be discounted as an explanation of high differentiation at the smallest scales.

Key words: genetic differentiation, aquatic dispersal, flatworms, cytochrome oxidase I

TABLE OF CONTENTS

List of Tables

List of Figures

INTRODUCTION

Dispersal is one of the most important processes determining the ecology and evolution of aquatic organisms (Bohonak & Jenkins, 2003). Understanding the processes that control evolution and diversification (e.g. dispersal limitations) are critical to protecting and sustaining biodiversity (DeSalle & Amato, 2004, Lande, 1988).

The dispersal of aquatic invertebrates is difficult to measure directly because traditional methods, such as mark and recapture (Southwood & Henderson, 2000) and stable isotopes (Hershey et al., 1993) are often impractical or ineffective (Bilton et al., 2001). However, population molecular data can be used to infer dispersal patterns based on the spatial distribution of genetic differentiation among conspecific populations following the assumption of selective neutrality (Slatkin, 1985, Hughes et al., 2008).

Multiple models have been used to predict patterns of genetic differentiation in stream invertebrates based on the interaction between dispersal traits and the spatial connectivity of the habitat (Hughes and Finn, 2009). The most basic expectation is isolation by distance where the degree of differentiation increases as the geographic distance between populations increases (Wright, 1943, Slatkin, 1993). On its own, the isolation by distance model (IBD) does not incorporate the spatial structure of the stream hierarchy (e.g. Frissell et al. 1986). Meffe and Vrijenhoek (1988) developed the Stream Hierarchy (SH) model to examine isolation by distance where movements of individuals are confined to the stream network. This model best applies to species that complete their life cycle in the water (e.g. shrimp and fish) or for species that primarily disperse along the stream corridor (e.g. some aquatic insects). For example, genetic variation in Pacific blue-eye fish, *Pseudomugil signifier* (McGlashan and Hughes 2002), and the midge, *Elporia barnardi* (Wishart and Hughes 2003) was low between reaches nested in

segments, moderate among segments within catchments, and reached maximum levels at the greatest distances, among catchments.

Meffe and Vrijenhoek (1988) also described the Death Valley model (DVM), for species that disperse among distinct habitat patches separated by an inhospitable matrix, such as springs in a desert landscape. However, this model can also apply to stream organisms that are habitat specialists. Even though streams are continuous, they can have a patchy structure to species that require specific habitat conditions (e.g. cool water temperatures and stable substrates). Patch-type models similar to DVM such as the "sky island" (Finn et al. 2006) and "headwater" hypotheses (Finn et al. 2007), have been applied to headwater specialists in streams that cannot disperse through the stream network because of inhospitable conditions in downstream segments (e.g. warm temperatures and frequently disturbed substrate). According to these models we expect population genetic differentiation between but not within patches. Depending on the dispersal ability of specific species there may or may not be a relationship between levels of differentiation and the Euclidean distance between patches. For example, genetic differentiation increased with distance between headwater patches in the giant water bug, *Abedus herberti* (Finn et al., 2007) and for the crayfish, *Euastacus* spp. (Ponniah and Hughes (2004, 2006). For species that cannot disperse across the terrestrial landscape, vector mediated dispersal (e.g. ducks feet) may maintain low levels of gene flow between patches (Figuerola & Green, 2002).

The freshwater flatworm *Polycelis coronata* (Turbellaria, Planariidae; Ehrenberg, 1831, Girard, 1891) is a habitat specialist in cold springs and headwater streams in mountainous regions. Observations from the Wasatch Mountains of Utah suggest that this species cannot tolerate water temperatures approximately > 10° C (Beck 1954). Although *P. coronata* populations can reach large numbers in headwater segments in mountainous regions, they are

seldom found in warmer downstream reaches where the substrate is frequently disturbed (Beck 1954). Nixon and Taylor (1977) used protein electrophoresis to show high dissimilarity between distant populations of *P. coronata* in the state of Washington. These authors also speculated that warmer downstream water temperatures may act as a barrier to gene flow between headwater populations. Thus, we expected limited movement between patches through the stream network. Also, this species‟ life cycle is entirely aquatic. Consequently, we did not expect dispersal through the terrestrial environment except rarely by passive means on various types of transport vectors (birds, mammals, etc.).

The objective of this study was to examine the dispersal of *Polycelis coronata* using estimates of gene flow across multiple scales in the Wasatch Mountains of Utah. We used mtDNA to address the following questions: 1) how is genetic variation partitioned across multiple scales (reaches, segments, catchment basins), 2) at what scale(s) do we begin to detect the effects of dispersal limitations on gene flow, 3) which model (stream hierarchy or headwater patches) best describes the geographic pattern of population differentiation? We hypothesized that genetic variation in *P. coronata* would be small within headwater segments and reach maximum levels between headwater segments with no increase in differentiation with increasing distance between headwater patches or between drainages. Thus, we do not expect this species to disperse along the stream network or through the terrestrial environment (actively or passively).

METHODS

Focal Species

Polycelis coronata is the only species of this genera found in North America. It occurs in springs and cold creeks of the western United States (Kenk, 1973). Like most flatworms, it is a

wandering predator (Calow et al., 1981) that uses tactile and chemosensory (Collins & Gerald, 2009) cues to locate prey such as the larvae and pupae of aquatic insects (Wrona and Dixon, 1991). The ecology of this flatworm is poorly understood (Kenk, 1973), and only one study has used molecular techniques to examine dispersal (Nixon and Taylor, 1977).

Rate of Dispersal Experiment

We constructed six artificial stream channels in the South Fork of the Provo River to estimate the maximum upstream dispersal rate of *P. coronata*. Channels were constructed using 183 cm by 13 cm sections of vinyl rain gutter which were secured in the middle of the stream side by side using steel rebar. The top portion of each channel was exposed 2 cm above the surface of the water, and the upstream and downstream ends of each channel were covered with fine nylon mesh (1 mm diameter). Two to three centimeters of gravel sized substrate was washed in a bucket remove all aquatic invertebrates and placed in each channel. Because *P. coronata* is thought to be more active at dawn and dusk (Kenk 1976), each channel was covered with thick black plastic for the duration of the experiment to simulate low light conditions. We randomly selected three channels as treatment replicates, and three channels as control replicates. The treatment replicates each received 2 grams of beef liver secured to the upstream end of the channel. We haphazardly collected 160 *P. coronata* by hand from the stream and starved them for 24 hours in a stream-side cage. Following the starvation period, twenty individuals were haphazardly chosen for each replicate and placed in a small holding cage at the downstream end of the channel and allowed to acclimate for ten minutes. Individuals were then released for one hour. We recorded the distance each individual moved upstream from the starting location to estimate average rates of dispersal. We used a Wilcoxon Rank Sum test (Wilcoxon, 1960) to test the hypothesis that an olfactory stimulus (beef liver) will positively affect rates of dispersal.

Sampling Design

We used a spatially nested hierarchical sampling design (Frissell et al., 1986) to test the SH model within genetic differentiation is lowest between reaches nested in segments, moderate among segments within catchments and reach maximum levels at the largest scales, among catchments.

We collected individuals from 50 sites nested within 24 segments, nested in four adjacent catchments along the Wasatch Mountains of central Utah (Fig. 1; Appendix 1). The four catchments selected were the Provo, American Fork, Little Cottonwood Canyon, and Weber River catchments (Fig. 1). Provo River is a fourth order catchment, while American Fork, Little Cottonwood, and Weber Rivers are third order catchments. Within each catchment, we conducted a thorough search for *P. coronata* in both headwater segments and larger downstream reaches, except in the Weber River catchment because of limited access to headwater reaches.

Within each catchment we selected up to 8 segments, and within each segment we chose an upstream and downstream site. Sites were a stream reach of 50 m to 100 m in length and were separated by at least 500 meters within a segment. At each site we collected at least 20 flatworms by hand from the undersides of several rocks. Individuals were initially stored in 95% ethanol; in the lab they were transferred into 99% ethanol and stored at -70 C until genetic analysis was performed. We also collected physico-chemical data from each site (pH, conductivity, and water temperature) and made visual estimates of substrate stability (presence of bryophytes, armoring, and siltation). We obtained *P. coronata* specimens from several geographically distant locations (outgroups) in order to get a clearer picture of the relative levels of divergence (Idaho, Wyoming, Colorado, New Mexico). We expected these sequences from

distant sites to show greater divergence from populations in the Wasatch Mountains compared to levels of divergence among populations within the Wasatch Mountains.

DNA extraction, amplification, and sequencing

We extracted genomic DNA from tissue samples using the DNeasy Tissue Kit (Qiagen, Hilden, Germany). The mitochondrial DNA gene cytochrome c oxidase I is one of the most commonly used markers in flatworm molecular studies (Álvarez-Presas et al., 2008, Brändle et al., 2007, Pongratz et al., 2003, Sunnucks et al., 2006, Vilas et al., 2005). Mitochondrial DNA evolves at a relatively fast rate and is not subject to excessive recombination (Vilas et al., 2005), making it an appropriate choice given the geographic scale of our study. We initially used the universal flatworm primers Pr-a2 and pr-b (Bessho et al., 1992) for use in amplifications of mtDNA, with very little success. Based on the limited sequence data we obtained using Pr-a2 and pr-b, we designed our own primers, FW14F (5‟-ACACCTGATATGATWTTYCCTCG-„3) and FW.COI.2R (5‟-GCTTTAGACAAAACTATTCCAG -3‟). Using these primers, we amplified a 763 base segment of the COI gene.

Final concentrations for polymerase chain reaction (PCR) components per 25 µL reaction were as follows: 25 ng template DNA, 0.25 µM of each primer, 0.625 units of Taq DNA polymerase, 0.1 mM of each dNTP, 2.5 µL of 10X reaction buffer and 2.5mM $MgCl_2$. Amplification parameters were as follows: 94°C for 2 min followed by 35 cycles of 94°C for 30 s, 48°C for 30 s, and 72°C for 75 s, and 72°C for 7 min. We examined PCR products on a 1% agarose gel using SYBR safe DNA gel stain (Invitrogen, Eugene, OR, USA). We purified PCR products using a Montage PCR 96 plate (Millipore, Billerica, MA, USA). Sequences were obtained via cycle sequencing with Big Dye 3.0 dye terminator ready reaction kits using 1/16th reaction size (Applied Biosystems, Foster City, CA). Sequencing reactions were run with an

annealing temperature of 52°C following the ABI manufacturer"s protocol. We purified sequenced products using sephadex columns. Sequences were obtained using an Applied Biosystems 3730 XL automated sequencer at the Brigham Young University DNA Sequencing Center.

DNA sequences were edited using Chromas Lite 2.0 (Technelysium, Tewantin, Queensland, Australia) and imported into BioEdit 7.0.5.2 (Hall, 1999) then aligned by eye and checked using MAFFT version 6 (Katoh et al. 2002). COI sequences were checked for unexpected frame shift errors or stop codons in Mega 4.0 (Tamura *et al.* 2007). Editing resulted in a 763 base pair (bp) fragment for each individual included in our study. We amplified a total of 541 unambiguous sequences from over 50 sites, with each site represented by at least 10 individuals. This is consistent with similar molecular studies on stream invertebrates (De Praz et al., 2008, Lehrian et al., 2010, Garrick et al., 2004, Wishart & Hughes, 2003). Our preliminary work revealed sufficient genetic variation to suggest adequate sample size (Björklund & Bergek, 2009). All sequences obtained in this study will be deposited in GenBank, accession numbers pending.

Molecular analysis

We used a variety of analytical approaches to address our hypotheses regarding genetic differentiation and historical gene flow at multiple scales, focusing on patterns and potential underlying processes. We employed a phylogenetic approach using maximum parsimony and maximum likelihood methods. We then used Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992) to reveal patterns of genetic differentiation at multiple scales. This was coupled with coalescent approaches to reveal historical gene-flow estimates and population size.

These approaches vary in their underlying assumptions, and by employing these different techniques it was our intent to provide robust evidence to support our conclusions.

To test the SH model hypothesis that catchments are barriers to gene flow, we reconstructed phylogenetic trees using both maximum parsimony and maximum likelihood methods using Tree Analysis with New Technology (TNT) (Goloboff et al., 2000) and RAxML (Stamatakis, 2006). Since high levels of differentiation were previously found using protein electrophoresis (Nixon & Taylor, 1977), we chose to represent the data with a phylogenetic tree instead of a haplotype network because haplotype networks do not show haplotype relationships as clearly when divergences are deep. To reconstruct a phylogenetic tree using these approaches, we first condensed all redundant sequences using MacClade v4.06 (Maddison and Maddison 2003). This reduced our total dataset to 149 haplotypes, plus two outgroup taxa, *Polycelis felina* and *Polycelis tenius* (GenBank accession no. DQ666049 and AF178321). To reconstruct a phylogenetic tree under the assumptions of maximum parsimony, we performed a traditional search in TNT. Nodal support was evaluated with 10,000 bootstrap replicates. To reconstruct a phylogenetic tree under the assumptions of maximum likelihood we first used the Akaike information criterion in the software package ModelTest vc.06 (Posada & Crandall, 1998) to select the appropriate model of molecular evolution. ModelTest selected the Hasegawa-Kishino-Yano model, with rates among sites conforming to a gamma distribution with invariant sites (HKY (G+I)). This model of evolution was assumed to then reconstruct a phylogenetic tree in RAxML under the operating procedures of maximum likelihood. Nodal support was evaluated with 1000 bootstrap replicates.

We performed Analysis of Molecular Variance (AMOVA) using Arlequin 3.5 (Excoffier & Lischer, 2010) to test the hypothesis that genetic differentiation would be lowest between

reaches nested in segments, moderate among segments within catchments and reach maximum levels at the largest scales, among catchments according to the SH model. This analysis shows the spatial scales that account for the greatest amount of total genetic variance. Arlequin is unable to simultaneously analyze more than 3 nested hierarchical levels, whereas we have four levels in our sampling design (within-site variation, variation among sites nested in segments, variation among headwater segments nested in catchments, and variation among catchments). Given this constraint, we chose to divide our AMOVA into two separate steps. First, for each catchment we examined how variation was partitioned within a site (F_{ST}), between sites within each segment (F_{SC}), and then between segments (F_{CT}). Thus, we completed four separate AMOVA‟s for each catchment. Second, we examined how molecular variation was partitioned within segments, among segments within a catchment, and among catchments. We also conducted a Mantel test using the statistical package R 2.8.1 (Team, 2008)to test the null hypothesis that pair-wise genetic distances (F_{ST}) were not correlated with spatial Euclidian distances between populations.

Sequences were also analyzed using a coalescent approach to reveal the underlying processes that may explain the genetic patterns. The coalescent approach has the added advantage of using genetic data to look backward into time (Pearse & Crandall, 2004). We first used the program MIGRATE-N vers 3.1.3 (Beerli and Felsenstein 2001). This program uses tree based information to estimate the magnitude and direction of migration rates and gene flow between groups (Beerli & Felsenstein 2001). We performed three separate runs to test the SH model versus the headwater hypothesis. We estimated gene flow among catchments, among segments within catchments, and among reaches within segments. According to the SH model gene flow will be rare between catchments, moderate among segments within catchments, and

maximum between reaches within segments. By contrast, the headwater model predicts the lowest gene flow among stream segments with no additional decrease in gene flow between catchments. We used the following parameters: 15 short chains of $2X10^5$ steps followed by 4 long chains of $2X10^6$ steps, each chain was sampled every 20 steps after an initial burn-in of 10000 trees. Adaptive heating with the following initial relative temperatures [6.00, 3.00, 1.20, and 1.00] was included, where acceptance-rejection swaps were tried with every step. Each run also applied the Gelman"s convergence criterion that extends the last run until the convergence criterion is met. Nucleotide frequencies were estimated from the data, and initial estimates of theta and gene flow were obtained using F_{ST}. Multiple runs were conducted to check for convergence. We excluded individuals from 3 segments within the Provo catchment from the analysis of sites within segment due to a small sample size.

We acknowledge the possibility that incomplete lineage sorting may complicate inferences of dispersal based on genetic data. To address this issue, we used Beast 1.5.3 (Drummond & Rambaut, 2007) to construct a Bayesian skyline plot (BSP) used to estimate demographic history. We also used Arlequin to estimate effective population sizes in the distant (Θ_S) and recent (Θ_π) past. A decline in effective population size is evidence for a bottleneck event. Evidence of a bottleneck would indicate a reduction in effective population size, causing an increase the rate of lineage sorting attributed genetic drift, which would reduce the number of rare haplotypes and thus, the importance of incomplete sorting in explaining patterns of genetic divergence. To ensure adequate sample size, we combined data from all catchments. We used a substitution rate of 1.5%./My, based on a genetic study of terrestrial flatworms (Garrick et al., 2004). We ran five initial short runs of $1 X 10^6$ to ensure adequate tuning and optimization of

the operators and then ran multiple longer runs with a chain length of 5 X 10^8. We used Tracer 1.4 (Rambaut & Drummond, 2007) to visualize and combine results, and calculate the BSP.

RESULTS

Physico-chemical conditions

Physico-chemical conditions were similar among sites where *P. coronata* was collected. The mean water temperature was cold (7.4^0 C; SE ± 0.68), whereas and the mean pH (7.7; SE ± 0.06) and conductivity (311.6 µS/cm; SE ± 17.6) were higher than typical downstream, mainstem segments for mountainous streams in this region. Also, 52% of the sites showed evidence of stable substrate (presence of bryophytes, armoring, and siltation). These data confirm previous results indicating that this species is a cold stenotherm reaching highest densities where the substrate is stable because of relatively stable flows attributed to high groundwater and spring inflows. These measurements and the fact that this species was rare or absent from the mainstem and downstream reaches in each catchment confirms the assertion that *P. coronata* is a headwater specialist.

Rate of Dispersal Experiment

The maximum rate of dispersal for a single individual was 6 cm/min, whereas the mean rate for the treatment was 1.87 cm/min, which was significantly different (P = 0.046) from the control rate (1.18 cm/min). Thus, on average these flatworms can move 8 km per year and they do increase their rate of movement in response to food. These results indicate that *P. coronata* does have the capacity to move from one stream to another through the stream network over the course of a year if the intervening environmental conditions permit (cold temperatures and stable substrate).

Geographic distribution

We sequenced 503 individuals from 50 sites/reaches nested in 26 segments, nested in four catchments using a fragment of the mtDNA COI gene (763 bp). We also sequence 38 additional individuals from distant sites (Appendix 2). Separate haplotypes were identified based on variation in 227 base pairs. Two patterns were used to test our hypotheses: 1) the geographical distribution of haplotypes across scales and 2) the phylogenetic relatedness among haplotypes. We also used nested AMOVA and coalescent analyses (MIGRATE N, BEAST.) to determine potential processes contributing to the formation of these patterns.

Polycelis coronata was rare or absent from the lowest elevations of each catchment and from four sub-basins in the Provo River catchment (Fig. 1). The streams in these four sub-basins are on private land and are heavily diverted for agriculture. Thus, the absence of *P. coronata* in this part of the Provo catchment is probably a recent extirpation caused by anthropogenic dewatering.

On average, there were 5 haplotypes out of a total of 134 that occurred at a site in this study. Seventy-one percent (95 haplotypes) were rare and only occurred in a single reach, whereas 5% (6 haplotypes) were abundant and collected in between 8 and 19 sites (Fig. 2). Although the rare haplotypes were evenly distributed across all catchments, the abundant haplotypes primarily clustered within a single catchment with much fewer individuals scattered throughout the remaining three basins. This pattern among the abundant haplotypes suggests greater gene flow within than between catchments, whereas numerous rare haplotypes evenly distributed across all catchments suggests the importance of historical events (see Discussion).

Haplotype relatedness

Analyses based on both maximum parsimony (MP) and maximum likelihood (ML) showed similar phylogenetic trees composed of one large clade (Clade 1) and five smaller clades (Fig. 3). Over 87% of all haplotypes and 85% of all individuals were contained within Clade 1. Most haplotypes in Clade 1 were only divergent from each other by 1 or 2 base pairs, and 3% were scattered across all four drainages. As expected, Clades 2, 3, and 5 were separate from Clade 1 because they were composed of individuals from distant sites (Wyoming, Colorado, New Mexico, and Logan, UT). However, a few of the individuals from these distant sites were most closely related to and occurred in Clade 1. This indicates that some of the haplotypes in Clade 1 collected from the Wasatch Mountains were closely related to haplotypes from distant site that must have a broad geographical distribution. Also, it is interesting that the two most distantly related clades (1 and 6) were primarily comprised of individuals from the Wasatch Mountains. Furthermore, all haplotypes in Clade 6 co-occurred in the same reach with haplotypes from Clade 1. This suggests the possibility of cryptic speciation.

Potential Processes

The AMOVA analysis showed that *P. coronata* populations did not conform to either the SH model or the headwater model. The SH model predicts genetic differentiation to be low between reaches nested in segments, moderate among segments within catchments, and highest among catchments, whereas the headwater model predicts the greatest differentiation between headwater segments. Instead, genetic variation reached maximum levels among individuals within sites, whereas differences among catchments and difference among headwater segments explained less of the total amount of genetic variance. When all drainages were analyzed together, differences among drainages only accounted for a small percentage of the total variance

(7.5%), because most (54.4%) was accounted for by differences among (39%) and within especially within segments (54.4%, Table 1). When catchments were analyzed separately and sites/reaches were nested in segments, differences among individuals within sites/reaches accounted for the greatest variation except in the Provo River, which did follow the headwater model (Table 1). In the Provo River, differences among segments accounted for the greatest variation (50.1%). However, the Mantel test found a positive correlation between genetic differences and spatial differences between all populations ($p = .001$), indicating isolation by distance and suggesting that passive overland dispersal may be in important factor in explaining the genetic patterns observed. Isolation by distance and maximum genetic variation within sites explaining the greatest proportion of total variance are inconsistent with both the SH model and headwater model.

There was no difference in effective population sizes in the distant (Θ_S) and recent (Θ_π) past (Table 6). Plus, the Bayesian skyline plot showing the variation in effective population size through time was flat. Thus, there is no evidence of a past population bottleneck or population expansion. Large effective population sizes and the absence of a bottleneck suggest incomplete lineage sorting as a possible explanation of high within site differentiation as uncovered by the AMOVA. If high within site differentiation is a historical signature that is not indicative of current processes, incomplete lineage sorting might weaken support for either the SH or HM by diluting the amount of variation attributed to higher scales.

Historical estimates of gene flow (MIGRATE N) revealed low gene flow between catchments consistent with the SH model (Table 3), but also showed that patterns of gene flow between headwater segments were more consistent with the headwater model because gene flow between adjacent segments was not always greater than between non-adjacent segments (Tables

4-7). Within headwater segments, we did not detect any obvious patterns of movement between upstream and downstream reaches (Tables 8-11). For example, downstream movement was at least four times greater in magnitude than upstream movement in 3 segments of the Provo catchment, but the reverse was also true of 3 other segments in this catchment. Under the SH model, gene flow should be lowest between catchments, moderate among segments within catchments and highest between reaches nested in segments. Estimates were consistent with this general pattern providing support for the SH model. However, when we compared gene flow between adjacent stream segments versus non-adjacent segments, say on opposite ends of a drainage, we found greater gene flow between non-adjacent versus adjacent segments. This is not consistent with the SH model.

DISCUSSION

Our results suggest two processes may explain the patterns of many of rare haplotypes evenly distributed across all catchments, abundant haplotypes primarily found in one catchment, and high differentiation primarily within sites: 1) limited dispersal between segments consistent the headwaters model, or 2) incomplete lineage sorting. Of these two processes, incomplete lineage sorting was best supported by the data. There was little evidence to support the SH model, because there was little differentiation among catchments and because phylogenetic clades did not group by catchments.

According to the headwater model, the highest amount of genetic variation should be attributed to differences between headwater segments, because headwater specialists do not easily disperse through the stream network from one segment to the next. This pattern, however, was only detected in the Provo catchment. In the Weber catchment, high levels of differentiation between segments explained a meaningful portion of total genetic variance, but it was not

significant at the 0.05 level ($p = 0.059$). The remaining two catchments showed significant and meaningful differentiation among sites nested in segments, which was consistent with the DVM version of the headwater model, but again the majority of the variation was explained by within site differences.

Gene flow between adjacent segments was not always greater than between non-adjacent segments, which suggests dispersal has occurred between patches of headwater segments, rather than through the stream network. However, we acknowledge that these estimates of movement could also be an artifact of high within population variation due to incomplete lineage sorting, because MIGRATE N cannot distinguish between short divergence times with low gene flow and longer divergence times with higher gene flow. Incomplete lineage sorting and the presence of highly divergent individuals may have led MIGRATE N to calculate nonsensical results, such as the extremely high rates of movement calculated between a select number of segments. Considering all the results, there is not enough evidence to clearly support the headwater model over incomplete lineage sorting.

Incomplete lineage sorting can occur when populations have only recently been isolated and/or when effective populations are extremely large. The effect of incomplete lineage sorting is insufficient time to establish equilibrium between drift and migration. Small invertebrates such as *P. coronata* that have a relatively short generation time and can reproduce asexually (De Meester et al., 2002) can build very large effective population sizes under favorable conditions. Under conditions of infinite population size, the rate at which novel haplotypes are introduced via mutation can be considered to be at its highest and the effects of genetic drift are thought to be minimal, which could result in elevated levels of polymorphism (Templeton, 2006). For headwater specialist stenotherms such as *P. coronata*, it is possible that ancestral populations

could have been extremely large, extensively widespread, and much more connected among streams of western North America, including the Wasatch Mountains (Harper & Petersen, 1990), during the cooler and wetter climate of the Pleistocene. During the warming period following the end of the Pleistocene, these populations would have become isolated, as suitable habitats constricted to present day headwater streams. Although dispersal between these now isolated populations is unlikely, if these populations were able to maintain high effective population sizes, drift may not have had enough time to sort out the historical signal of high ancestral polymorphism. Estimates of past effective population sizes in *P. coronata* were extremely high, which is consistent with our current observations. We also did not find evidence of a bottleneck. Thus, the genetic patterns uncovered in this study may be relics of the past and not reflect current processes of dispersal.

Incomplete lineage sorting may account for several key patterns we detected. The fact that clades that did not cluster by catchment could be a result of insufficient time for populations to diverge from one another, and shared haplotypes across large distances may simply be a signature of much higher connectivity in the past that has not yet been erased. The high levels of within site differentiation uncovered by the AMOVA, may be an artifact of drift not yet having reduced a large portion of the genetic "noise" of ancestral polymorphism. Finn and Alder (2006) found strong among and within population divergence and shared haplotypes among populations in a headwater specialist black fly *Metacnephia coloradensis* and postulated that large population sizes have slowed the rate of among population divergence.

We acknowledge that high within population genetic differentiation may be the product of non-dispersal processes. Studies that report high within population genetic differentiation are rare but often invoke such explanations (Table 12). Explanations in these studies include

multiple founding events, rapid rates of nucleotide substitution, and systems of mating. We cannot discount these possibilities at this time. In addition, some of the total variation among haplotypes may be attributed to c ryptic speciation. Cryptic speciation in freshwater flatworms has been documented in at least one other case (Casu & Curini-Galletti, 2006). We suspect that haplotypes from clade 6 may be a cryptic species because of their high degree of divergence from the rest of the haplotypes (over 16% sequence divergence), and because they occurred in sympatry with many individuals from clade 1. Although we do not have multiple lines of evidence to support this hypothesis at this time, it does warrant future investigation.

LITERATURE CITED

Álvarez-Presas, M., Baguñà, J. & Riutort, M. 2008. Molecular phylogeny of land and freshwater planarians (Tricladida, Platyhelminthes): From freshwater to land and back. *Molecular Phylogenetics and Evolution* **47**: 555-568.

Austerlitz, F., Mariette, S., Machon, N., Gouyon, P. H. & Godelle, B. 2000. Effects of colonization processes on genetic diversity: Differences between annual plants and tree species. *Genetics* **154**: 1309-1321.

Bessho, Y., Ohama, T. & Osawa, S. 1992. Planarian Mitochondria.1. Heterogeneity of Cytochrome-C-Oxidase Subunit-I Gene-Sequences in the Fresh-water Planarian, *Dugesia-Japonica*. *Journal of Molecular Evolution* **34**: 324-330.

Bilton, D. T., Freeland, J. R. & Okamura, B. 2001. Dispersal in Freshwater Invertebrates. *Annual Review of Ecology and Systematics* **32**: 159-181.

Björklund, M. & Bergek, S. 2009. On the relationship between population differentiation and sampling effort: is more always better? *Oikos* **118**: 1127-1129.

Blouin, M. S., Dame, J. B., Tarrant, C. A. & Courtney, C. H. 1992. Unusual Population Genetics of a Parasitic Nematode: mtDNA Variation Within and among Populations. *Evolution* **46**: 470-476.

Blouin, M. S., Yowell, C. A., Courtney, C. H. & Dame, J. B. 1995. Host Movement and the Genetic Structure of Populations of Parasitic Nematodes. *Genetics* **141**: 1007-1014.

Bohonak, A. J. & Jenkins, D. G. 2003. Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters* **6**: 783-796.

Brändle, M., Heuser, R., Marten, A. & Brandl, R. (2007) Population structure of the freshwater flatworm Crenobia alpina (Dana): old lineages and low gene flow. Vol. 34. pp. 1183-1192. Blackwell Publishing Ltd.

Bunn, S. E. & Hughes, J. M. 1997. Dispersal and recruitment in streams: Evidence from genetic studies. *Journal of the North American Benthological Society* **16**: 338-346.

Calow, P., Davidson, A. F. & Woollhead, A. S. 1981. Life-cycle and feeding strategies of freshwater triclads: a synthesis. *Journal of Zoology* **193**: 215-237.

Casu, M. & Curini-Galletti, M. 2006. Genetic evidence for the existence of cryptic species in the mesopsammic flatworm Pseudomonocelis ophiocephala (Rhabditophora: Proseriata). *Biological Journal of the Linnean Society* **87**: 553-576.

Collins, A. M. & Gerald, G. W. 2009. Attraction of Flatworms at Various Hunger Levels Toward Cues from an Odonate Predator. *Ethology* **115**: 449-456.

Davison, A. & Clarke, B. 2000. History or current selection? A molecular analysis of ˜area effects" in the land snail Cepaea nemoralis. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**: 1399-1405.

De Meester, L., Gomez, A., Okamura, B. & Schwenk, K. 2002. The Monopolization Hypothesis and the dispersal-gene flow paradox in aquatic organisms. *Acta Oecologica-International Journal of Ecology* **23**: 121-135.

De Praz, A., Cordellier, M., Hausser, J. & Pfenninger, M. (2008) Postglacial recolonization at a snail's pace (Trochulus villosus): confronting competing refugia hypotheses using model selection. Vol. 17. pp. 2449-2462. Blackwell Publishing Ltd.

DeSalle, R. & Amato, G. 2004. The expansion of conservation genetics. *Nat Rev Genet* **5**: 702-712.

Drummond, A. J. & Rambaut, A. 2007. Beast: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**: 214.

Excoffier, L. & Lischer, H. E. L. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Vol. 10. pp. 564-567. Blackwell Publishing Ltd.

Excoffier, L., Smouse, P. E. & Quattro, J. M. 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* **131**: 479-491.

Figuerola, J. & Green, A. J. (2002) Dispersal of aquatic organisms by waterbirds: a review of past research and priorities for future studies. In: *Freshwater Biology*, Vol. 47. pp. 483-494. Wiley-Blackwell.

Finn, D. S. & Adler, P. H. (2006) Population genetic structure of a rare high-elevation black fly, *Metacnephia coloradensis*, occupying Colorado lake outlet streams. In: *Freshwater Biology*, Vol. 51. pp. 2240-2251. Wiley-Blackwell.

Finn, D. S., Blouin, M. S. & Lytle, D. A. 2007. Population genetic structure reveals terrestrial affinities for a headwater stream insect. *Freshwater Biology* **52**: 1881-1897.

Frissell, C. A., Liss, W. J., Warren, C. E. & Hurley, M. D. 1986. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environmental Management* **10**: 199-214.

Garrick, R. C., Sands, C. J., Rowell, D. M., Tait, N. N., Greenslade, P. & Sunnucks, P. (2004) Phylogeography recapitulates topography: very fine-scale local endemism of a saproxylic giant springtail at Tallaganda in the Great Dividing Range of south east Australia. In: *Molecular Ecology*, Vol. 13. pp. 3329-3344. Wiley-Blackwell.

Goloboff, P., Farris, S. & Nixon, K. (2000) TNT (Tree analysis using New Technology) (BETA) ver. xxx pp. Published by the authors, , Tucuman, Argentina.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95-98.

Harper, K. T. & Petersen, M. S. 1990. Natural History of Little Cottonwood Canyon, Utah. *Bulletin of the Ecological Society of America* **71**: 381-386.

Hershey, A. E., Pastor, J., Peterson, B. J. & Kling, G. W. 1993. Stable Isotopes Resolve the Drift Paradox for Baetis Mayflies in an Arctic River. *Ecology* **74**: 2315-2325.

Hsiao, J. Y. & Lee, S. M. 1999. Genetic diversity and microgeographic differentiation of Yushan cane (Yushania niitakayamensis; Poaceae) in Taiwan. *Molecular Ecology* **8**: 263-270.

Hughes, J., Schmidt, D., McLean, A. J. & Wheatley, A. (2008) Population genetic structure in stream insects: what have we learned? pp. CABI Publishing.

Hughes, J. M., Bunn, S. E., Hurwood, D. A. & Cleary, C. 1998. Dispersal and recruitment of Tasiagma ciliata (Trichoptera : Tasimiidae) in rainforest streams, south-eastern Australia. *Freshwater Biology* **39**: 117-127.

Kenk, R. 1973. *Freshwater triclads (Turbellaria) of North America, V: the genus Polycelis*. Smithsonian Institution Press; [for sale by the Supt. of Docs., U.S. Govt. Print. Off.], Washington.

Lande, R. 1988. Genetics and Demography in Biological Conservation. *Science* **241**: 1455-1460.

Lehrian, S., BÃ¡lint, M. s., Haase, P. & Pauls, S. U. 2010. Genetic population structure of an autumn-emerging caddisfly with inherently low dispersal capacity and insights into its phylogeography. *Journal of the North American Benthological Society* **29**: 1100-1118.

Michalski, S. G. & Durka, W. (2007) High selfing and high inbreeding depression in peripheral populations of Juncus atratus. Vol. 16. pp. 4715-4727. Blackwell Publishing Ltd.

Moritz, C. & Heideman, A. 1993. The Origin and Evolution of Parthenogenesis in Heteronotia binoei (Gekkonidae): Reciprocal Origins and Diverse Mitochondrial DNA in Western Populations. *Systematic Biology* **42**: 293-306.

Nixon, S. E. & Taylor, R. J. 1977. Large genetic distances associated with little morphological variation in polycelis coronata and Dugesia tigrina (Planararia). *Systematic Zoology* **26**: 152-164.

Pearse, D. E. & Crandall, K. A. 2004. Beyond F_{ST}: Analysis of population genetic data for conservation. *Conservation Genetics* **5**: 585-602.

Pongratz, N., Storhas, M., Carranza, S. & Michiels, N. 2003. Phylogeography of competing sexual and parthenogenetic forms of a freshwater flatworm: patterns and explanations. *BMC Evolutionary Biology* **3**: 23.

Posada, D. & Crandall, K. 1998. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* **14**: 817 - 818.

Rambaut, A. & Drummond, A. J. 2007. Tracer 1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Ribeiro, M. M., Plomion, C., Petit, R., Vendramin, G. G. & Szmidt, A. E. 2001. Variation in chloroplast single-sequence repeats in Portuguese maritime pine (Pinus pinaster Ait.). *TAG Theoretical and Applied Genetics* **102**: 97-103.

Slatkin, M. 1985. Rare Alleles as Indicators of Gene Flow. *Evolution* **39**: 53-65.

Slatkin, M. 1993. Isolation by Distance in Equilibrium and Non-Equilibrium Populations. *Evolution* **47**: 264-279.

Southwood, R. & Henderson, P. A. (2000) Ecological methods. Blackwell Science, Oxford; Malden, Mass.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688-2690.

Sunnucks, P., Blacket, M. J., Taylor, J. M., Sands, C. J., Ciavaglia, S. A., Garrick, R. C., Tait, N. N., Rowell, D. M. & Pavlova, A. (2006) A tale of two flatties: different responses of two terrestrial flatworms to past environmental climatic fluctuations at Tallaganda in montane southeastern Australia. In: *Molecular Ecology*, Vol. 15. pp. 4513-4531. Wiley-Blackwell.

Team, R. D. C. (2008) R: A Language and Environment for Statistical Computing. pp. R Foundation for Statistical Computing.

Templeton, A. R. 2006. *Genetic Drift in Large Populations and Coalescence*. John Wiley & Sons, Inc.

Vilas, R., Criscione, C. D. & Blouin, M. S. 2005. A comparison between mitochondrial DNA and the ribosomal internal transcribed regions in prospecting for cryptic species of platyhelminth parasites. *Parasitology* **131**: 839-846.

Watanabe, Y. & Chiba, S. (2001) High within-population mitochondrial DNA variation due to microvicariance and population mixing in the land snail Euhadra quaesita (Pulmonata: Bradybaenidae). Vol. 10. pp. 2635-2645. Blackwell Science Ltd.

Wilcoxon, F. 1960. *Individual comparisons by ranking methods*. Bobbs-Merrill, College Division, Indianapolis, Ind.

Wishart, M. J. & Hughes, J. M. (2003) Genetic population structure of the net-winged midge, *Elporia barnardi* (Diptera: Blephariceridae) in streams of the south-western Cape, South Africa: implications for dispersal. In: *Freshwater Biology*, Vol. 48. pp. 28-38. Wiley-Blackwell.

Wright, S. 1943. Isolation by Distance. *Genetics* **28**: 114-138.

TABLE 1—AMOVA RESULTS. Results of the hierarchical analysis of molecular variance (AMOVA; Excoffier et al, 1992). for all catchments together and for each separately. Segments were nested within catchments and sites within segments. Significance was tested for each analysis using 16,000 permutations. F_{SC} =among segments within catchments and among sites within segments. F_{CT}= among catchments and among segments.

Source of variation	**df**	**Percentage of variation**	**Fixation indices**	**p**
ALL CATCHMENTS				
Among catchments	3	07.49	F_{CT}: 0.075	<0.0001
Among segments within catchments	20	39.08	F_{SC}: 0.422	<0.0001
Within segments	430	**53.44**	F_{ST}: 0.466	<0.0001
PROVO CATCHMENT				
Among Segments	10	**50.14**	F_{CT}: 0.501	<0.001
Among sites within segments	11	11.87	F_{SC}: 0.238	<0.0001
Within sites	173	37.99	F_{ST}: 0.620	<0.0001
AMERICAN FORK CATCHMENT				
Among Segments	4	21.24	F_{CT}: 0.212	0.09
Among sites within segments	5	32.23	F_{SC}: 0.409	<0.0001
Within sites	90	**46.53**	F_{ST}: 0.535	<0.0001
LITTLE COTTONWOOD CATCHMENT				
Among Segments	5	-5.19	F_{CT}: -0.052	0.6
Among sites within segments	6	29.3	F_{SC}: 0.279	<0.0001
Within sites	108	**75.89**	F_{ST}: 0.241	<0.0001
WEBER CATCHMENT				
Among Segments	4	44.96	F_{CT}: 0.450	0.059
Among sites within segments	5	02.09	F_{SC}: 0.038	0.025
Within sites	73	**52.95**	F_{ST}: 0.471	0.0001

TABLE 2—THETA ESTIMATES. Average effective population sizes measured by theta, population genetic diversity, in the distant (Θ_S) and recent (Θ_π) past for each catchment.

Catchment	Θ_S	Θ_π
Provo	13.26	13.35
Weber	26.23	33.32
American Fork	17.81	22.75
Little Cottonwood	16.39	19.06

TABLE 3—MIGRATE RESULTS, ALL CATCHMENTS. Gene flow estimates for each catchment calculated in MIGRATE N (Beerli and Felsenstein 2001). Values show the number of individuals per generation (*Nm*) moving from catchments in the first column (source populations) to catchments in the top row (destination populations).

	Provo	American Fork	Little Cottonwood	Weber
Provo		3.3	1.7	6.8
American Fork	12.9		<0.1	3.5
Little Cottonwood	0.8	8.3		<0.1
Weber	0.9	4.9	0.6	

TABLE 4—MIGRATE RESULTS, PROVO. Gene flow estimates for each segment within the Provo catchment calculated in MIGRATE N (Beerli and Felsenstein 2001). Values show the number of migrants per generation (*Nm*) moving from segments in the first column (source populations) to segments in the top row (destination populations).

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1		2.7	1.0	1.1	0	0.8	10.4	0.2	0.2	3.3	0.2
2	9.0		6.9	1.3	0.1	8.9	0	0.1	0	0	0
3	3450	0		7.8E+10	0	0	1.1E+10	0	0	8.2E+6	0
4	0	6.4	0.8		0	0	0.2	0.2	1.6	0	0
5	0	0	1.0	0		0	4.2	0	0.1	5.4	1.4
6	0.9	9.2	0	0	0		2.2	0	0	0.9	0.5
7	7.6	24.4	0	2.2	0	4.2		0	0	1.2	0
8	0	0	0	0.5	0	0	5.5		0	3.2E+10	0
9	0	0	0	0	0.1	0	0.3	0		0.9	0.1
10	3	9.5	1.1	0.4	0	3.2	0	0.3	0		0
11	0	0	0.6	0	3.7	0.2	0	3.2	1.5	0	

TABLE 5—MIGRATE RESULTS, AMERICAN FORK. Gene flow estimates for each segment within the American Fork catchment calculated in MIGRATE N (Beerli and Felsenstein 2001). Values show the number of migrants per generation (*Nm*) moving from segments in the first column (source populations) to segments in the top row (destination populations).

	1	**2**	**3**	**4**	**5**
1		16.2	0	0	0
2	44.5		0	27.9	0
3	2.3	2.1		0	25.1
4	0.4	0	0		1.2
5	0	0	31.9	0	

TABLE 6—MIGRATE RESULTS, WEBER. Gene flow estimates for each segment within the Weber catchment calculated in MIGRATE N (Beerli and Felsenstein 2001). Values show the number of migrants per generation (Nm) moving from segments in the first column (source populations) to segments in the top row (destination populations).

	1	**2**	**3**	**4**	**5**
1		3.4	0	0	0
2	0.9		0.8	0	0
3	0	0.6		44.1	0.4
4	0	0	16.8		0.9
5	0	4.5E+07	1.3E+08	3.2	

TABLE 7—MIGRATE RESULTS, LITTLE COTTONWOOD. Gene flow estimates for each segment within the Little Cottonwood catchment calculated in MIGRATE N (Beerli and Felsenstein 2001). Values show the number of migrants per generation (Nm) moving from segments in the first column (source populations) to segments in the top row (destination populations).

	1	**2**	**3**	**4**	**5**	**6**
1		27.4	0	0	8.6	0
2	0		6.5	2.4	13.7	17.3
3	6.5	17.7		0.8	0	0
4	0	15.1	0		1.9	7.8
5	0	0	0	0.4		14.7
6	0	781.0	0	57.2	23.2	

TABLE 8—MIGRATE RESULTS, PROVO REACHES. Gene flow estimates for movement between upstream and downstream reaches within each headwater segment of the Provo catchment, calculated in MIGRATE N. Values show the number of migrants (Nm) per generation moving upstream or downstream.

Segment	Upstream	Downstream
1	0	515.0
2	7.0E+08	68.3
3	2.8E+05	56.7
4	0	2.2E+09
5	0	20.1
6	9.8E+09	61.9
7	0	6.0E+09
8	0	3.9E+09

TABLE 9—MIGRATE RESULTS, LITTLE COTTONWOOD REACHES. Gene flow estimates for movement between upstream and downstream reaches within each headwater segment of the Little Cottonwood catchment, calculated in MIGRATE N. Values show the number of migrants (Nm) per generation moving upstream or downstream.

Segment	Upstream	Downstream
1	64.9	2.6E+15
2	4.9E+11	36.3
3	67.4	2.3E+10
4	72.0	1.2E+16
5	1.5E+13	83.2
6	1.7E+15	53.5

TABLE 10—MIGRATE RESULTS, AMERICAN FORK REACHES. Gene flow estimates for movement between upstream and downstream reaches within each headwater segment of the American Fork catchment, calculated in MIGRATE N. Values show the number of migrants (*Nm*) per generation moving upstream or downstream.

Segment	Upstream	Downstream
1	6.62E+11	98.4
2	66.3	2.4E+10
3	6.6E+10	67.3
4	2.0E+10	38.4
5	87.1	118.0

TABLE 11—MIGRATE RESULTS, WEBER REACHES. Gene flow estimates for movement between upstream and downstream reaches within each headwater segment of the Weber catchment, calculated in MIGRATE N. Values show the number of migrants (*Nm*) per generation moving upstream or downstream.

Segment	Upstream	Downstream
1	1.9E+13	81.1
2	9.5E+09	33.1

TABLE 12—LITERATURE SUMMARY. Summary of literature reporting high genetic differentiation within populations. A few articles are not included because they fall outside the scope of this analysis (e.g. used different genetic markers).

Citation	Organism	Explanation
(Austerlitz et al., 2000)	Annual plants and tree species	Delayed reproduction allows for a large increase in the number of initial founders representing differing lineages for a given population before reproduction begins.
(Blouin et al., 1992)	Nematode	1. Populations represent a mix of worms from previously isolated populations, including cryptic species. 2. Accelerated rate of nucleotide substitution. 3. Large effective population sizes.
(Blouin et al., 1995)	Nematodes	1. Large effective population sizes. 2. Accelerated rate of nucleotide substitution.
(Davison & Clarke, 2000)	Land snail	Populations are mixed descendants of multiple founding events.
(Finn & Adler, 2006)	Blackfly	Large effective population sizes and recent fragmentation events can cause current population structure to retain a signature of historical patterns.
(Hsiao & Lee, 1999)	Yushan Cane plant	Out-crossing mating system.
(Hughes et al., 2008, Bunn & Hughes, 1997, Hughes et al., 1998)	Aquatic insects Plecoptera, Trichoptera, Ephemeroptera	Patchy recruitment hypothesis…a stream reach is re-populated each generation by the offspring of only a few females from a small subset of the total number of haplotypes in the region. Populations are rare and do not move far within a stream segment.
(Michalski & Durka, 2007)	Black Rush plant	High levels of inbreeding depression maintain allelic richness by favoring heteroszygosity.
(Moritz & Heideman, 1993)	Gecko	Multiple lineages with distinctive ancestry have arisen at different places and times.
(Ribeiro et al., 2001)	Pine tree	Human‟s augment dispersal and colonization of various haplotyes to local patches.
(Watanabe & Chiba, 2001)	Land snail	Repetition of isolation and mixing has resulted in fine scale variation within populations.

FIGURE LEGENDS

Figure 1. Map showing the sites (reaches) within catchments of this study (Provo River, Weber River, American Fork River, and Little Cottonwood River). Closed circles represent sites where *Polycelis coronata* occurred, whereas open circle show sites where this species was absent. Dashed lines indicate the boundaries between catchments.

Figure 2. Chart showing the relative frequencies of the six most abundant haplotypes in each catchment. Colors denote haplotypes, and are listed in order from most frequent.

Figure 3. Maximum likelihood phylogram of all COI haplotypes of *Polycelis coronata*, with major clades denoted by numbers 1-6). The substructure of clade 1 is shown to the right. Each haplotype is followed by the catchment(s) it was found in: Provo (P), Weber (W), American Fork (AF), and Little Cottonwood (LC). Sites outside the focal sampling area are designated as follows: WY = Wyoming, CO = Colorado, ID = Idaho, Log = Logan, UT, PG = Pleasant Grove, UT. When haplotypes are shared, catchments are listed in order of most abundant. * indicates ML/MP bootstrap = 90-100,** indicates ML/MP bootstrap = 50-89.

FIGURE 1—SAMPLING MAP

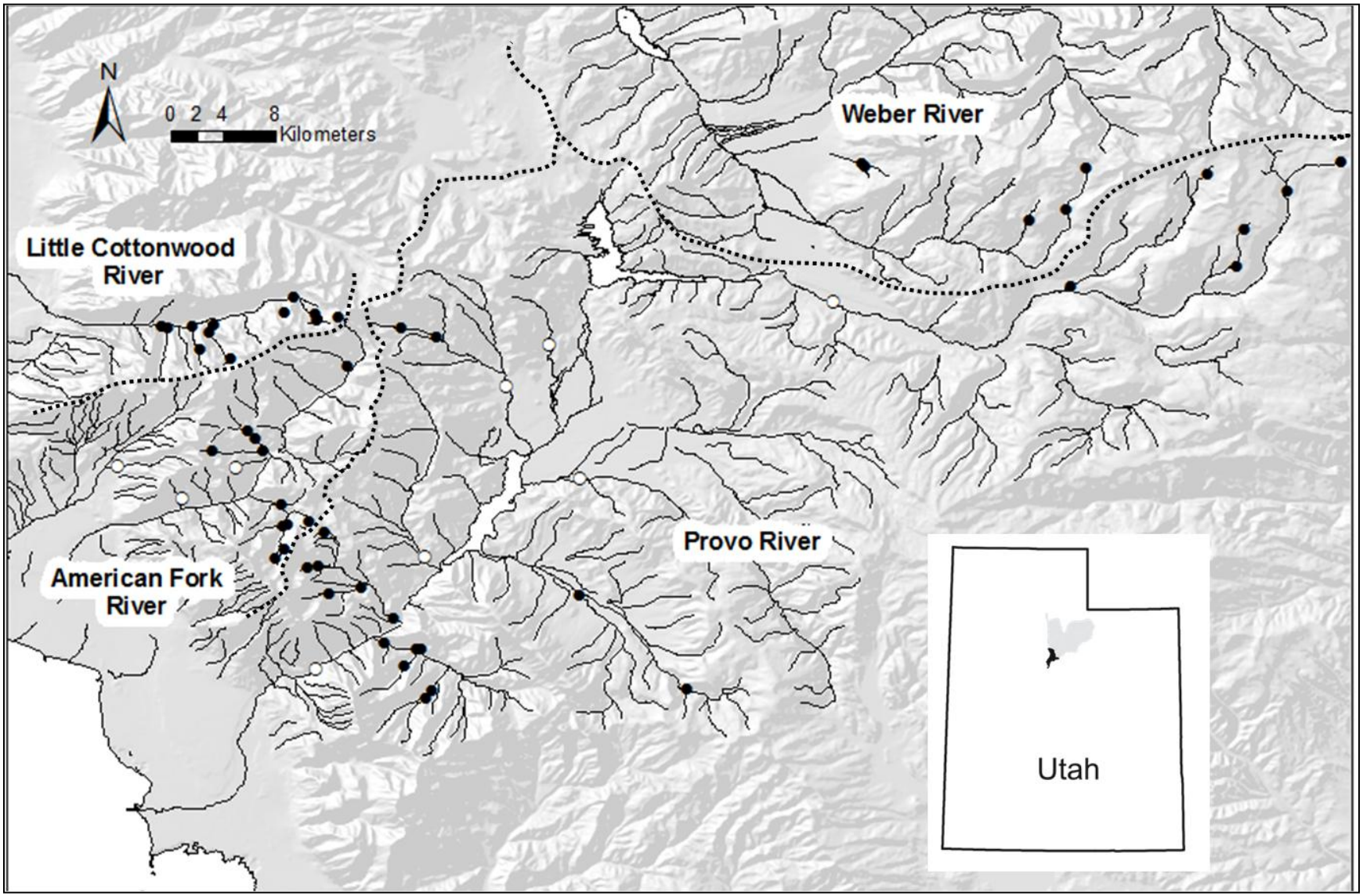


FIGURE 2—FREQUENCY AND ABUNDANCE OF MOST COMMON HAPLOTYPES.

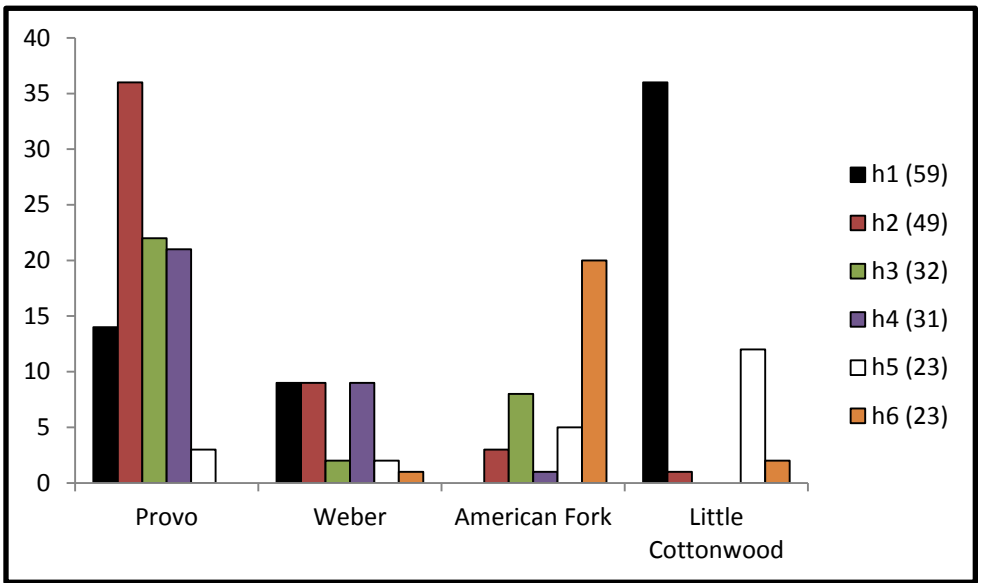
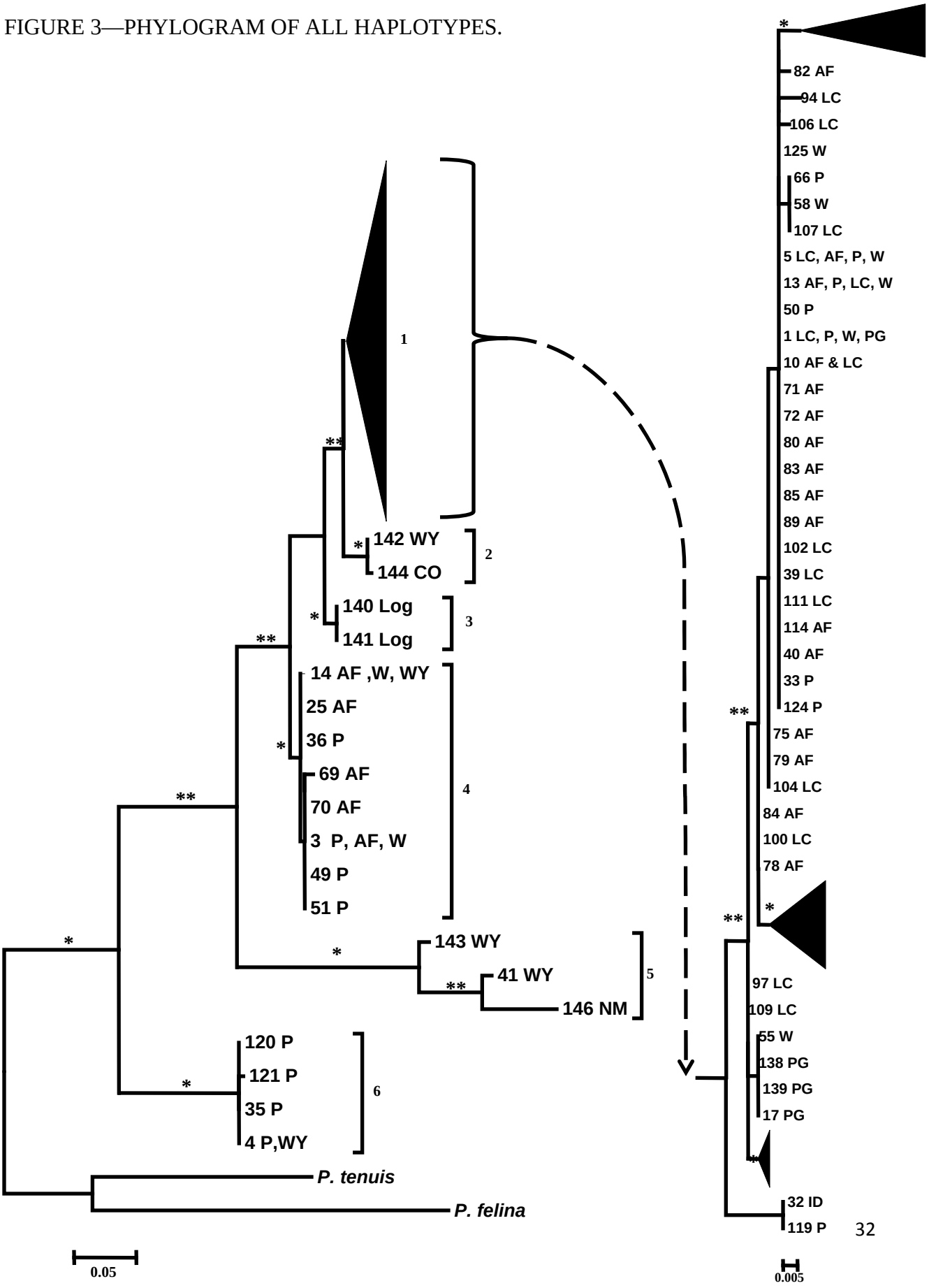


FIGURE 3—PHYLOGRAM OF ALL HAPLOTYPES.



Appendix 1. GPS coordinates and elevation (m asl) for sites sampled in this study. Site numbers correspond to those used in Appendix 2.

Site	Catchment	Segment	GPS Coordinates	Elevation (m)
1	Provo	1	N 40°34.135' W111°33.414'	2386
2	Provo	1	N 40°34.099' W111°33.311'	2377
3	Provo	2	N 40°24.240' W111°37.245'	2315
4	Provo	2	N 40°24.268' W111°36.780'	2181
5	Provo	3	N 40°23.167' W111°36.370'	2241
6	Provo	3	N 40°23.424' W111°35.055'	1907
7	Provo	4	N 40°25.669' W111°36.581'	2419
8	Provo	4	N 40°22.135' W111°33.721'	1624
9	Provo	5	N 40°19.295' W111°24.995'	2089
10	Provo	5	N 40°20.192' W111°33.284'	2158
11	Provo	6	N 40°20.858' W 111°32.728'	1831
12	Provo	6	N 40°18.820' W111°32.368'	1683
13	Provo	7	N 40°19.183' W111°32.103	1921
14	Provo	7	N 40°20.842' W111°32.582'	1914
15	Provo	8	N 40°21.111' W111°34.057'	1656
16	Provo	8	N 40°40.861' W111°54.835'	1596
17	Provo	9	N 40°39.783' W111°56.792'	2793
18	Provo	9	N 40°40.562' W111°59.485'	2904
19	Provo	10	N 40°35.832' W111°05.743'	3077
20	Provo	10	N 40°40.900' W111°04.439'	2303
21	Provo	11	N 40°40.046' W111°05.988'	2862
22	Provo	11	N 40°38.151' W110°58.540'	2793
23	Weber	1	N 40°38.151' W110°58.540'	3054
24	Weber	1	N 40°36.676' W110°58.900'	2901
25	Weber	2	N 40°40.749' W111°14.226'	2224
26	Weber	2	N 40°40.917' W111°14.369'	2202
27	Weber	3	N 40°40.749' W111°14.226'	2618
28	American Fork	1	N 40°32.563' W111°35.597'	2353
29	American Fork	2	N 40°29.841' W111°39.737'	2182
30	American Fork	2	N 40°29.524' W111°39.398'	2085
31	American Fork	3	N 40°29.063' W111°41.155'	2680
32	American Fork	3	N 40°29.078' W111°39.117'	2051
33	American Fork	4	N 40°24.636' W111°38.589'	2635
34	American Fork	4	N 40°26.025' W111°38.083'	2389
35	American Fork	5	N 40°25.000' W111°38.130'	2389
36	American Fork	5	N 40°26.835' W111°38.335'	2023
37	American Fork	6	N 40°43048' W111°63886'	2636
38	American Fork	6	N 40°26.025' W111°38.083'	2230
39	Little Cottonwood	1	N 40°35.220' W111°37.490'	2615
40	Little Cottonwood	1	N 40°34.272' W111°41.098'	2615
41	Little Cottonwood	2	N 40°34.590" W111°36.009"	2853
42	Little Cottonwood	2	N 40°34.739' W111°38.195'	2805
43	Little Cottonwood	3	N 40°34.451" W111°36.882'	2842
44	Little Cottonwood	3	N 40°34.716' W111°36.923'	2821
45	Little Cottonwood	4	N 40°32.839' W111°40.394'	2892
46	Little Cottonwood	4	N 40°33.959' W111°41.330'	2484
47	Little Cottonwood	5	N 40°33.211' W111°41.709'	2663
48	Little Cottonwood	5	N 40°34.202' W111°42.019'	2184
49	Little Cottonwood	6	N 40°34.121' W111°43.029'	2014
50	Little Cottonwood	6	N 40°34.172' W111°43.282'	1997

Appendix 1. continued

Distant Sites

Site	Location	GPS Coordinates
51	Santa Fe, New Mexico	N 35°45.24' W105°48.32'
52	Fort Bridger, WY	N 41°18.50' W110°31.07'
53	Afton, WY	N 42°42.02' W110°51.00'
54	Fort Hall,ID	N 43°03.44' W112°31.03'
55	Thompson creek, Colorado	N 40°27.17' W105°26.04'
56	Logan, UT	N 41°42.31' W111°43.05'
57	Pleasant Grove, UT	N 40°23.402' W111°41.725'

Appendix 2. Haplotypes generated from COI sequences. The number of sites is shown below each catchments and the number of individuals from each site with a particular haplotype are shown in the body of the table. Empty cells indicate that no individuals representing the specified haplotype occurred at that site. The total number of individuals sharing each haplotype is indicated in the last column.

Localities

Haplotypes	Provo																					Weber					American Fork												Little Cottonwood												total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	
1	3	9							1					1						2			1			5	1													7	1	6	1		5	1	2	7	4	1	1	59
2	1						1										8	3	7		7	9	2	1			6	3																	1						49
3			8	8		1	5																1			1		1					4			2													32
4					2			1	4						2	4				8				9				1																						21
5	1					2	1										2											1			2				2				1		1				7	1				2	23
6																		1													1	7	5		1	5	1										2		23
7											6	6				1																	1																	14
8																																							1	1	1	5	1	3						2	14
9																													6	2		1				1						1							2	13
10																														2	1		2	1						1	1		2				1	1		12
11													5	4	2																																			11
12					1	2		5								1												1																							10
13					1	1				1									1													4								2										10
14																																			3	1		4													9
15																										1					6									1											7
16																																							1				5					1		7
17										4																																								4
18												1	3	2																																				6
19																									2	1	1																								4
20																																														3		1	1	1	6

Appendix 2. continued

Localities

Haplotypes	Provo																						Weber					American Fork											Little Cottonwood												total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	
21									1		1		1	1																					1																5
22															2	2																																	1		5
23																		5																																	5
24																					5																														5
25																																			1	3		1													5
26												1		2																					1																4
27						2		2																																											4
28																																								2								1		1	4
29								1							2																																				3
30																		1			2																														3
31									2																		1																								3
32										2																																									2
33	3																																																		3
34												1																																1							2
35															2																																				2
36							1	1																																											2
37				2																																															2
38																																	1	1																	2
39																																															2				2
40																																																			2
41																																			1																1
42													1																																						1
43												1																																							1
44											1																																								1
45											1																																								1

Appendix 2. continued

Localities

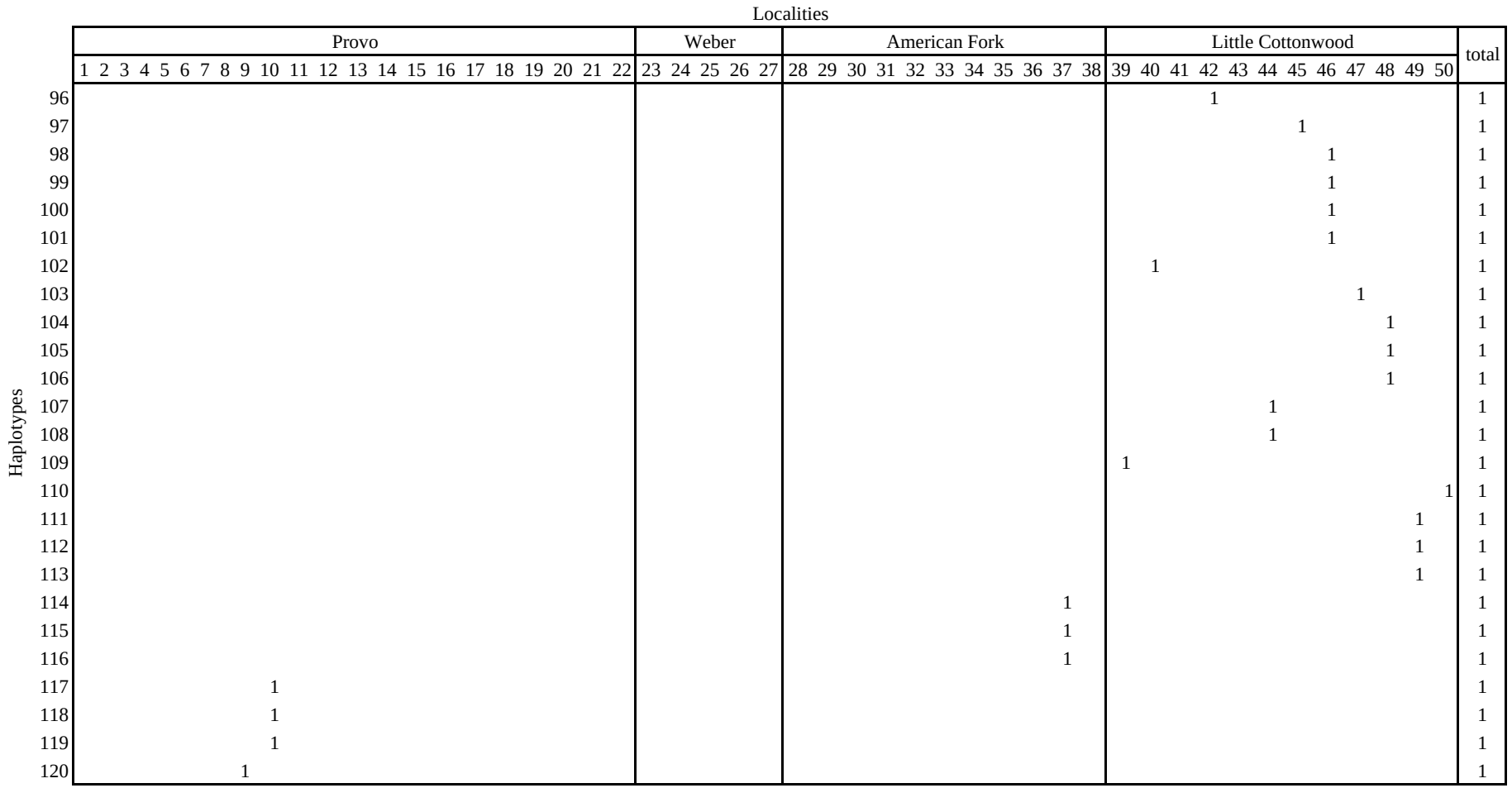
Haplotypes	Provo																					Weber					American Fork													Little Cottonwood													total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	
46											1																																								1
47																1																																			1
48																1																																			1
49							1																																											1
50							1																																											1
51			1																																															1
52			1																																															1
53					1																																													1
54					1																																													1
55																			1																															1
56																			1																															1
57																							1																											1
58																										1																								1
59																					1																													1
60																						1																												1
61					1																																													1
62					1																																													1
63					1																																													1
64					1																																													1
65					1																																													1
66					1																																													1
67																																		1																1
68																																		1																1
69																																					1													1
70																																					1													1

Appendix 2. continued

Localities

Haplotypes	Provo																						Weber					American Fork											Little Cottonwood												total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	
71																																				1															1
72																																				1															1
73																															1																				1
74																															1																				1
75																													1																						1
76																													1																						1
77																													1																						1
78																													1																						1
79																														1																					1
80																														1																					1
81																														1																					1
82																														1																					1
83																														1																					1
84																														1																					1
85																												1																							1
86																												1																							1
87																												1																							1
88																												1																							1
89																																1																			1
90																																																			1
91																																											1								1
92																																											1								1
93																																									1										1
94																																										1									1
95																																										1									1

Appendix 2. continued



Appendix 2. continued

Localities

Haplotypes	Provo																						Weber					American Fork											Little Cottonwood												total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	
121									1																																										1
122	1																																																		1
123	1																																																		1
124		1																																																	1
125																									1																										1
126																									1																										1
127																									1																										1
128																									1																										1
129																									1																										1
130																									1																										1
131																									1																										1
132																									1																										1
133																										1																									1
134																										1																									1