# Population Structure of *Tetragnatha versicolor* (Araneae: Tetragnathidae) in Northern California

**Hillary Thomas**

**Abstract** The genus *Tetragnatha* is an abundant group of orb-weaving spiders, often found to have riparian associations. *Tetragnatha versicolor* is a member of this genus that is widespread and found in a diversity of habitats in North America, including California. I examined mitochondrial sequence data (Cytochrome oxidase I) from 21 individuals from eight different localities to assess population structure for *T. versicolor*. A grouped AMOVA showed that variation between populations (45% of the total variation) was likely due to larger regional difference. An estimation of migration from $F_{ST}$ suggested that approximately 1 individual moves among populations every second generation. A surprising amount of genetic variation was present in this species, and significant genetic differences were found among populations. However, divergent haplotypes of *T. versicolor* were more closely related to each other than to other species within the genus, which is consistent with the status of *T. versicolor* as a species.

## Introduction

The genus *Tetragnatha* (Araneae: Tetragnathidae) is a widespread and abundant group of orb-weaving spiders with a cosmopolitan distribution (Gillespie, *pers. com*; Levi, 1980). While their common name, the Long-jawed orb-weavers, points outwardly to their morphologically distinct chelicerae, it is both their ecological associations and reproductive behavior that best characterize them. Members of this genus are most commonly found building their webs either near or over the water's surface (Foelix, 1996; Gillespie, 1987; Williams et al, 1995).

Several characteristics may serve to account for their association with water, including an increased risk of desiccation due to dehydration with comparison to other spiders (Gillespie, 1987), or potential prey preferences (Nyffeler, 1999). Relative to other species within this genus, *T. versicolor* has been found to have a more diverse habitat distribution, and likely exhibits greater tolerance in non-riparian environments (Aiken & Coyle, 2000). As with most spiders, the primary mode of dispersal is through "ballooning," whereby spiderlings (immature spiders) let silk out loosely into the wind and essentially imitate flying, although the distance traveled can vary greatly, depending on weather conditions (Foelix, 1996). The variety of distributions exhibited by spiders, coupled with their unique and successful dispersal mechanisms, allow them to serve as interesting subjects for studies of phylogeography and population genetic structure.

Studies in population structure have long been concerned with looking at patterns of genetic distribution and movement of alleles at the intraspecific level (Costa & Ross, 1994); however, it is only recently that tools have been available which integrate the interests of both population geneticists and systematists. Traditionally, population geneticists have been concerned with microevolutionary processes such as natural selection, migration, mutation and genetic drift, as microevolutionary processes are some of the driving forces governing the structure of phylogenetic trees at the interspecific scale. Despite this understanding, differences in models and methods have historically divided the two disciplines (Avise, 1989). With newly available methods and a greater ease of obtaining sequence data (Slatkin, 1994; Roderick, 1996; Emerson et al, 2001), the methodologies for analyzing microevolutionary processes are converging with those used by systematists in tree reconstruction.

In particular, the use of "gene trees," (a genealogy of haplotypes or alleles) has been a bridge between the two disciplines. For example, traditional phylogenetic methods can assume the

absence of ancestral haplotypes (different allelic variants of the same gene), an assumption that is often violated at the population level because ancestral haplotypes are frequently the most common intraspecific allele variants (Clement et al, 2000; Galacatos et al, 2002). The construction of gene trees generated from sequence data at homologous sites, coupled with new modes of analysis more appropriate at the population level, such as nested clade analysis, have allowed us to differentiate between historical processes, such as fragmentation and range expansion, and recurrent processes such as gene flow (Alexandrino et al, 2002; Roderick, 1996). These methods can circumvent problems that would otherwise be encountered when using phylogenetic methods. Nested clade analysis (Templeton et al, 1992), the method employed in the current study, allows us to test for geographic associations of haplotypes at different hierarchical levels of a gene tree (Alexandrino et al, 2002). When this method has been tested on known data sets, it has been shown to accurately reflect the history and structure of populations (Templeton et Al, 2000). In this manner, a null hypothesis of no geographic association is tested, and further analyses do not proceed unless this can be rejected (Templeton et al, 2000).

These developments in available tools and types of data have been part of the foundation of phylogeography, a discipline concerned with relating the phylogeny of different genotypes to geography (Avise, 2000; Hillis & Morowitz, 1996). Findings in this field have been both insightful and controversial, as for example, Wake's (1997) study on incipient species formation of salamanders in California, which added substantial questions to the already controversial topic of species concepts. While in the past, many species have been defined by such criteria as their morphological characters and reproductive compatibility, Wake's paper shows the true complexity of species definition, by illustrating how range expansion and contraction, and fluctuating levels of gene flow can affect the process of speciation. Still other papers serve to illustrate how social behavior can affect genetic variation (Costa & Ross, 1994; Johannesen et al, 1998), and the application of these tools has also extended into analysis of human populations, again serving to separate historical events from recurrent ones (Templeton, 2002).

Using mitochondrial sequence data from a single locus, this study seeks to infer the recent history of *T. versicolor* by testing for structure in the distribution of alleles both within and between populations. If structure exists, this study also asks at what spatial scale subdivision becomes observable. Using population genetic analyses, I inferred the existing level of migration that characterizes the population structure, assuming a state of genetic equilibrium. I

also employ sequence data from other members of the genus, to determine which mechanisms may have contributed to the observed genetic population structure of *T. versicolor*.

## Methods

**Collection** Specimens were collected between June 1, 2002 and August 10, 2002 from the following localities: Inyo National Forest (Mono county), Shasta National Forest (Shasta county), Six Rivers National Forest (Trinity county), Stanislaus National Forest (Stanislaus county), Hopland Field Station (Mendocino County), El Dorado county, and Sonoma county (See Map, Figure 1).

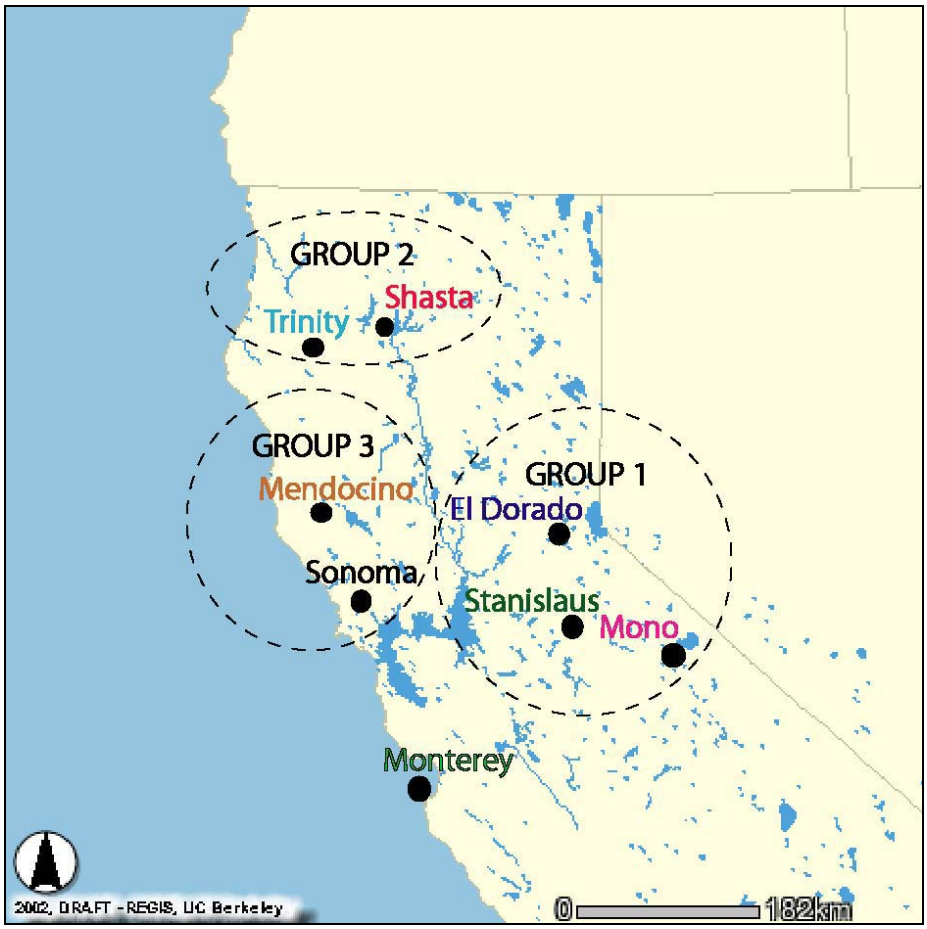



Figure 1: Map of Collection Sites. Source: UC Berkeley REGIS, 2002.

All specimens were collected by hand or with use of a sweep net, labeled, and preserved in 95% ethanol. Specimens were collected primarily along rivers and lakes, or in marshes. Fishing permit #046814-05 was obtained for collection of specimens from National Forests. Specimens were identified using a key by Levi (1981).

**Sequencing** DNA samples were extracted from individuals in each population using a commercial tissue extraction kit (DNEasy, Qiagen, Inc). A 40 cycle protocol was established and a 400 base pair portion of the mitochondrial gene *Cytochrome-oxidase* I (CO I) was amplified through the polymerase chain reaction (PCR) using the primer LCO 1628: 5' ATAATGTAATTGTTACTGCTCATGC paired with HCO 2396: 5'ATTGTAGCTGAGGTAAAATAAGCTCG, both designed in the Roderick and Gillespie laboratories, or with HCO 2198: 5' TAAACTTCAGGGTGACCAAAAAATCA (Folmer, 1994 #285). Cleaned PCR products were cycle sequenced using a Big Dye Terminator cycle sequencing kit v.3.0 (Applied Biosystems). PCR products were sequenced in both directions, and sequences were edited and aligned automatically using the program Sequencher v4.1 (Gene Codes Corporation).

**Statistical Techniques** Haplotypes were determined by calculating a pairwise distance matrix using the program Paup 4.0b10 (Swofford, 1994). Gene trees were constructed using parsimony and maximum likelihood criteria as well as neighbor joining methods. One goal of the use of trees in this study was to examine relationships with other species within the genus, to better understand the relationship of populations and alleles in a broader context. A network was generated using the program TCS v.1.13 (Clement et al, 2001), which first determines the limits of parsimony and then computes a network of the statistically relevant relationships between alleles within the limits of parsimony criterion. The resulting graph represents the reconstructed gene genealogy of the haplotypes.

Two analyses of molecular variation (AMOVA) were also performed using the program Arlequin v.2.000 (Schneider et al, 2000), Variation was first assessed both among and between populations through conventional methods from haplotype frequencies, and the resulting $F_{ST}$ was calculated.

| Group 1 | Group 2 | Group 3 |
|---|---|---|
| Stanislaus | Trinity | Mendocino |
| Mono | Shasta | Sonoma |
| El Dorado | | |

Table 1: Summary of population groupings for grouped AMOVA of variance.

Populations were then divided into three groups by geographic regions as summarized in Table 1, and an AMOVA was performed which calculated variation among groups, among populations within each group, and within populations, also using conventional F-statistics.

Significance tests and resulting p-values were calculated with a minimum of 1000 permutations for the population analysis, and 10,000 permutations for the grouped population analysis. The number of migrants per generation (Nm) was inferred from $F_{ST}$, using the following equation proposed by Wright (1951):

$$F_{ST} \approx \frac{1}{1 + 2Nm} \quad (1)$$

A global differentiation test was also performed. This method constructs a contingency table of the populations (rows) and haplotypes (columns) and then compares this to randomly constructed contingency tables of the data set, and calculates the probability that the observed results would be obtained by chance.

**Results**

Twenty-one sequences representing 10 haplotypes of *T. versicolor* were obtained. Two haplotypes, H7 and H2 were not included in the TCS analysis because their magnitude of difference compared to the other haplotypes in a pairwise data matrix was too great for the analysis to include them in the network. Of those haplotypes included in the computations, five groups still exhibited levels of variation great enough that it could not be resolved with statistical parsimony. These are represented by the rectangular boxes which lack resolution in their relationship to other groups within the TCS network (Appendix 1).

The relationships of the remaining three haplotypes were established within the limits of parsimony criterion. Haplotype 1, which represented the largest number of individuals within the data set, was connected to each of haplotypes 4 and 10 by two mutations (represented by lines). This assumes one internal node, an intermediate haplotype between haplotypes 1 and 4, and between haplotypes 1 and 10, which is not represented within the data set.

The gene trees reconstructed using parsimony, maximum likelihood, and neighbor joining were identical and consequently only the tree reconstructed using parsimony criterion is presented.

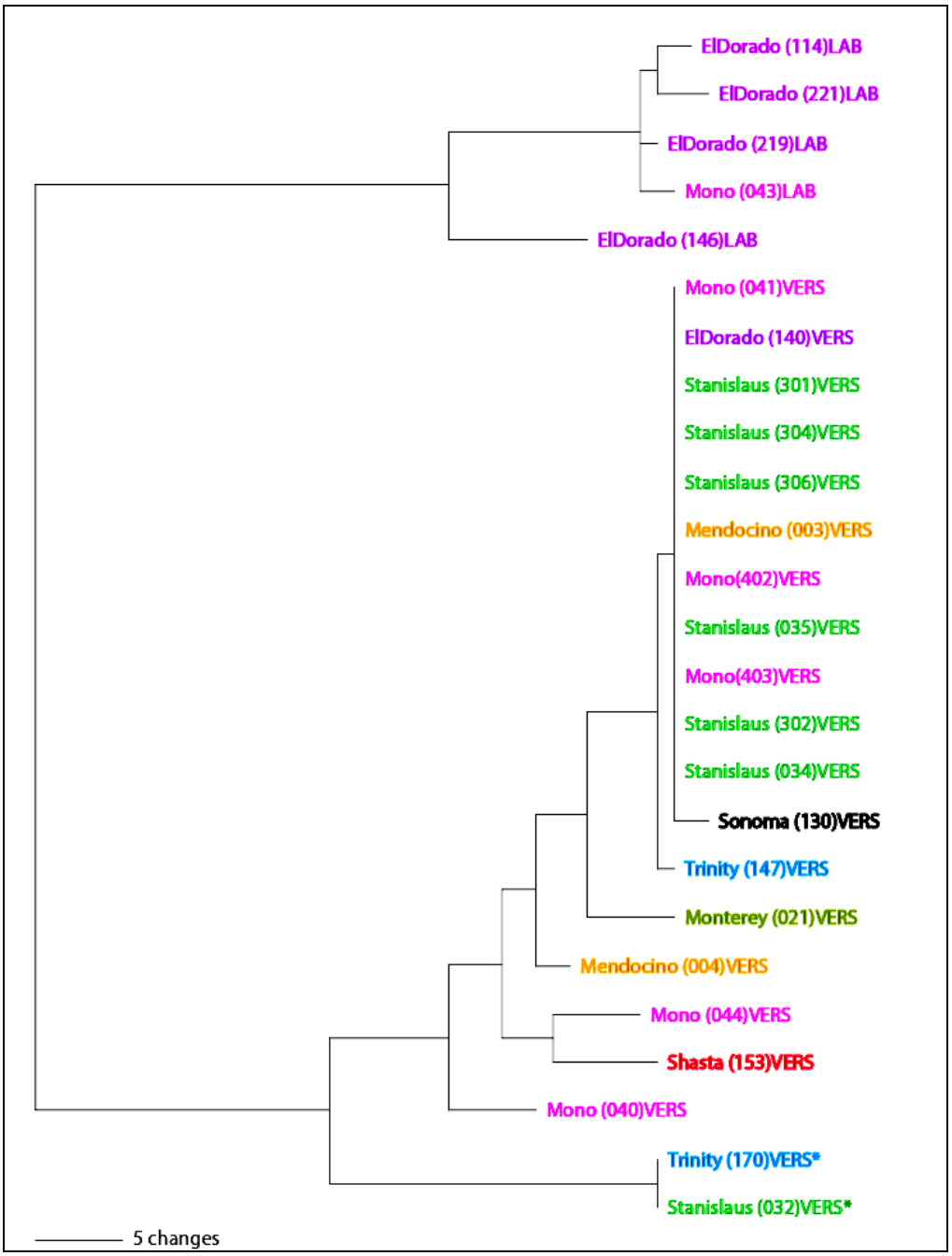



Figure 2: Parsimony tree of sequences, color coded by location. Individual names containing "LAB"are species *T. laboriosa*, while those labeled "VERS" refer to *T. versicolor*.

The phylogram (Figure 2) shows that the two species, *T. versicolor* and *T. laboriosa* clearly group separately. The two haplotypes, H2 and H7, which were omitted from the TCS cladogram due to the level of variance from the other haplotypes, are shown in the parsimony tree to group with the *T. versicolor* clade, despite the observed extent of variation. The tree also suggests that

very different haplotypes are represented within geographical populations, for example Mono(040) falls out separately from Mono(403), Mono(041) and Mono(402). Likewise, very distinct geographical locations are represented within haplotypes, in particular Stanislaus, Mono, El Dorado and Mendocino populations were all represented within H1.

Variation among populations comprised 45.2% of total variation while variation within populations accounted for 54.8% of total variation (Table 2).

| Source of Variation | d.f | Sum of Squares | Variance Components | Percentage of Variation |
|---|---|---|---|---|
| Among Populations | 7 | 3.54 | 0.157 (Va) | 45.23 |
| Within Populations | 10 | 1.90 | 0.190(Vb) | 54.77 |
| Total | 17 | 5.44 | 0.347 | |
| **$F_{st}$ : 0.45** **P(rand. value>=obs. value) = .012** | | | | |

Table 2: AMOVA results for variation among and within populations.

The resulting fixation index value for the haplotypes is .45 for $F_{ST}$, indicating that if random mating occurred between a single population and more encompassing level of hierarchical grouping, the loss of heterozygosity for the population would be 45%. This value is relatively high, and is consequently indicative of pronounced genetic differentiation (Hartl & Clark, 1997). Using Equation 1, Nm = .605, suggesting that there is only 1 migrant among populations every second generation. The global population differentiation test showed that there is virtually no probability that this distribution of haplotypes among populations would occur by chance (p = 0.001).

| Source of Variation | d.f | Sum of Squares | Variance Components | Percentage of Variation |
|---|---|---|---|---|
| Among groups | 2 | 1.69 | 0.11 (Va) | 28.58 |
| Among population within groups | 5 | 1.85 | 0.09 (Vb) | 22.45 |
| Within populations | 10 | 1.9 | 0.19(Vc) | 48.97 |
| Total | 17 | 5.44 | 0.39 | |

Table 3: Sources of variation in grouped structure AMOVA.

| Variance Component | P(rand.>=obs.) | P(rand.<=obs.) |
|---|---|---|
| Va | 0.02 | - |
| Vb | .32* | - |
| Vc | - | 0.01 |

Table 4: Associated P-values for variance components of grouped structure AMOVA.

The grouped structure AMOVA for the three groupings specified by geographic region (Table 3) indicated that (28.6%) of variation was due to larger groups and 49% of variation was due to variance within the populations. However, when regional differences were considered, genetic variation among populations within the groups was not significant.

**Discussion**

One possible source for such pronounced genetic variation within one species could be hybridization between closely related species. This possibility is unlikely given the distinct grouping of *T. versicolor* relative to other species within the genus (Figure 2). The magnitude of genetic differentiation exhibited between haplotypes of *T. versicolor* suggests that the population size must be large enough to maintain such levels of genetic variation over time. That these are most likely old, persistent haplotypes is the most plausible explanation for the observed variation.

There is also a significant level of geographic subdivision within the species, as indicated by both the test of global differentiation and the AMOVA analyses (Tables 2, 3). The significance of variation among populations is further expounded by the results of grouping the populations into larger geographic regions. The statistically significant percentage of variation due to differences among groups (45%), coupled with the lack of significant variation within populations of the same group suggests that there is some level of geographic subdivision, visible at a larger scale than the population level sampling. It is also likely that this accounts for the observed variation between populations, using haplotypes without grouped comparisons.

The inferred migration rate, on the order of 1 migrant every second generation between populations, is low enough to validate geographic subdivision, as a homozygous group requires around 1 migrant per generation between paired populations (Wright, 1931). However, the migration rate is also substantially high enough to overcome the effects of genetic drift; while there is differentiation between populations geographic subdivision, the species has not diverged

enough for evolution to act independently on the populations. This is supported by the presence of diversely distributed haplotypes within a single population.

The presence of some geographic structure of *T. versicolor* populations is of interest in the larger scheme of patterns of genetic exchange, in that it illustrates the capacity for several old haplotypes to persist in a group of populations that appear to represent a reproductively isolated species. It is possible that ecological conditions in the different geographical regions, particularly with respect to risk of dessication, may be selecting for different haplotype frequencies and that the coexistence of diverse haplotypes within the same population is likely the result of dispersal mechanisms through somewhat rare long distance ballooning event.

In order to begin to explain why this structure exists, the nature of their dispersal mechanism must be taken into consideration. Of interest to this topic is a study by Ramirez and Haakonseen (1999) which assessed gene flow for *Argiope trifasciata* among habitat patches and found that ballooning was contributing to genetic exchange between populations, but there was no evidence of equal exchange, as assumed in Wright's (1931) Island Model. They further explicate these results by suggesting that ballooning may not be as sufficient a mode of long distance dispersal as previously thought, due to the high levels of observed heterozygosity between populations.

This interpretation may undermine the role of causative factors in determining both when and why a spider balloons, particularly when members of the genus *Tetragnatha* are taken into consideration. *Tetragnatha* have demonstrated their ability to disperse over long distances (Gillspie, *pers. com*), although the frequency of long distance dispersal has not been addressed. In both agricultural and other habitat recolonization studies, members of the family Linyphiidae have been recognized as the most frequent immigrants (Weyman, 1993; Crawford et al, 1995).

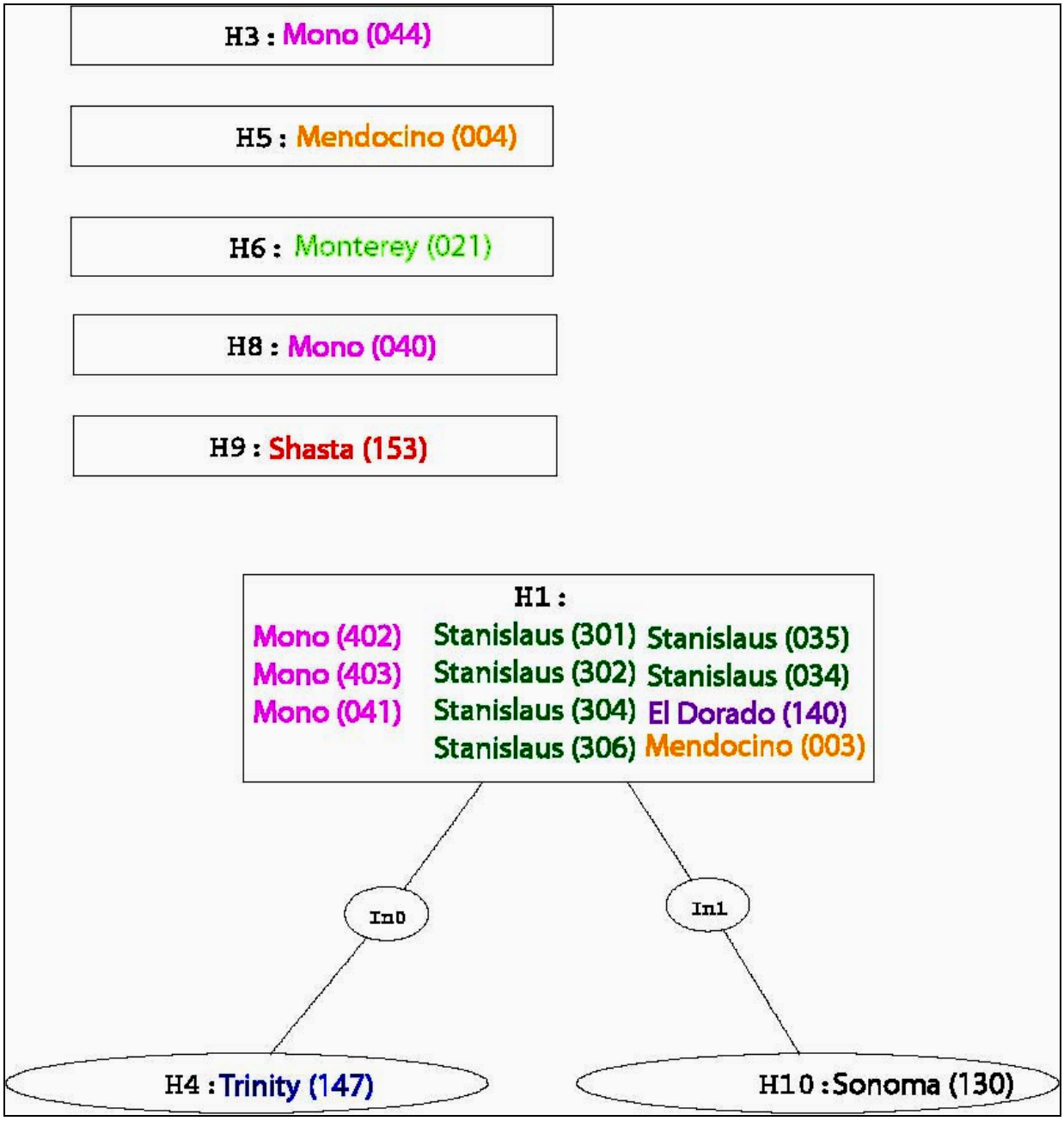
While there have been many studies on meteorological factors that may affect the choice to initiate ballooning, and some general conclusions have been made such as the maximum wind speed, even the results of these studies differ in their characterization of other optimal conditions (Weyman, 2002). It may be found that there is a great deal of variation in causation amongst different groups of spiders. Despite confounding meteorological data, both practical impediments to testing the frequency of long range dispersal, and a general absence of studies in this area, have limited our understanding of the more ecological and behavioral aspects of ballooning (Weyman, 1993).

One must consider that the potential benefits of dispersal outweigh the risks involved, such as mortality. However, in discussions of gene flow, it seems we are making assumptions that there is consistently an attempt for longer range of dispersal. Of interest is that members of this genus with riparian associations have been shown to aggregate for prey availability (Gillespie, 1987). As spiders are territorial, a high density of conspecifics, or a low prey availability are both factors which may necessitate movement, but different scales of population density, and prey availability might call for different ranges of dispersal distance. Spiderlings have been shown to make decisions about the weather conditions under which they will balloon (Foelix, 1996), so it is plausible that under various conditions, ballooning may provide multiple scales of dispersal. The question remains as to what degree spiders may be discriminating in their dispersal efforts.

Another ecological factor which might affect population structure is the level of microhabitat association exhibited by the species. One study in the Great Smoky Mountain National Park (Aiken and Coyle, 2000) remarked on *T. versicolor* as a habitat generalist. While this may be true, as they clearly can be found in a variety of habitats, a study in Ontario (Williams, Browning & Ambrose, 1995) found that *T. versicolor* and its more specialized relative, *T. elongata*, constituted 91% of all spider species observed along the banks of the study site. Their important role as predators occupying niches in various ecological systems deems their genetic differentiation more meaningful, particularly with respect to the maintenance of intraspecific genetic diversity. When one considers their abundance, it is also possible to conjecture the integral role they play in controlling insect populations in a variety of habitats, including aquatic freshwater ecosystems, rice fields, and forested areas.

At a time when biological diversity is being increasingly threatened by human induced environmental changes, it is important to obtain a broad understanding of the factors shaping the state of species, and the potential susceptibility of these species to alterations of the environment. One implication of the relatively high level of genetic variation and differentiation for *T. versicolor* can be understood by their key role as abundant generalist predators in trophic systems (Williams et al, 1995). Further inquiries into the population structure of other species within the genus may shed light on the ecological dynamics shaping their genetic structure, particularly when compared to more specialized species such as *T. elongata* or in comparison to *T. laboriosa* which has a similarly broad North American distribution.

## Appendices




Appendix A: TCS network of haplotypes.


## Acknowledgements

Special thanks to Rosemary Gillespie and George Roderick who supervised this project, and the entire Roderick-Gillespie lab for contributions and support. I would also like to thank Matt

Orr, John Latto and Manish Desai for comments, my editor, Jennifer Irwin, and my collecting assistant, Robert Michener. Additional funding provided by a Biology Fellow Program 2002 Undergraduate Research Grant.

**References**

Aiken, M. Coyle, F. 2000. Habitat Distribution, Life History and Behavior of *Tetragnatha* Spider Species in the Great Smoky Mountains National Park. Journal of Arachnology 28: 97-106.

Alexandrino, J. Arntzen, J. Ferrand, N. 2002. Nested Clade Analysis and the Genetic Evidence for Population Expansion in the Phylogeography of the Golden-Striped Salamander, *Chioglossa lusitanica* (Amphibia: urodela). Heredity 88: 66-74.

Avise, J. 1989. Gene Trees and Organismal Histories: a phylogenetic approach to population biology. Evolution 43(6): 1192-1208.

Avise, J. 2000. Phylogeography: the history and formation of species. Harvard University Press, Cambridge.

Clement,M. Posada, D. Crandall, K.A. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657-1659.

Costa, J. Ross, K. 1994. Hierarchical Genetic Structure and Gene Flow in Macrogeographic Populations of the Eastern Tent Caterpillar. Evolution 48: 1158-1170.

Crawford, R. Sugg, P. Edwards, J. 1995. Spider Arrival and Primary Establishment on Terrain Depopulated by Volcanic Eruption at Mount St. Helens, Washington. American Midland Naturalist 133 (1): 60-75.

Emerson, B. Paradis, E. Thébaud, C. 2001. Revealing the Demographic Histories of Species Using DNA Sequences. Trends in Ecology and Evolution 16(12): 707-716.

Excoffier, L. Smouse, P. Quattro, J. 1992. Analysis of Molecular Variance Inferred from Metric Distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491.

Foelix, R.F. 1996. Biology of Spiders. Oxford University Press, New York.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol Marine Biol Biotechnol* 3: 294-299.

Galacatos, K. Cognato, A. Sperling, F. 2002. Population Genetic Structure of Two Water Strider Species in the Ecuadorian Amazon. Freshwater Biology 47: 391-399.

Gillespie, R. 1987. The Mechanism of Habitat Selection in the Long-Jawed Orb-Weaving Spider *Tetragnatha elongata* (Araneae: tetragnathidae). Journal of Arachnology 15: 81-90.

Harl, D. Clark, A. 1997. Principles of Population Genetics. 3rd Ed. Sinauer Assoc: Massachusetts.

Hillis, D. Moritz, C. Mable, B. 1996. Molecular Systematics, 2nd Ed. Sinauer Assoc: Massachusetts.

Johannesen, J. Baumann, T. Seitz, A. Veith, M. 1998. The Significance of Relatedness and Gene Flow on Population Genetic Structure in the Subsocial Spider *Eresus cinnaberinus* (Araneae: eresidae). Biological Journal of the Linnean Society 63: 81-98.

Levi, H. 1981. The American Orb-Weaver Genera *Dolichognatha* and *Tetragnatha* North of Mexico. Bulletin of the Museum of Comparative Zoology 149(5): 271-317.

Mun, J. Bohanak, A. Roderick, G. 2003. Population Structure of the Pumpkin Fruit Fly *Bactrocera depressa* (Tephritidae) in Korea and Japan: Pliocene allopatry or recent invasion? *In press.*

Nyffeler, M. 1999. Prey Selection of Spiders in the Field. Journal of Arachnology 27: 317-324.

Ramirez, M. Haakonsen, K. 1999. Gene Flow Among Habitat Patches on a Fragmented Landscape in the Spider *Argiope trifasciata* (Araneae: Araneidae). Heredity 83: 580-585.

Roderick, G. 1996. Geographic Structure of Insect Populations: gene flow, phylogeography and their uses. Annu. Rev. Entomol. 41: 325-352.

Schneider, S,. Roessli, D. Excoffier, L. 2000. *Arlequin: a software program for population genetics data analysis*, Ver.2.0, Geneva, Switzerland.

Slatkin, M. 1994. *in* Ecological Genetics. Edited by Leslie Real. Princeton University Press: USA.

Swiffold, D. 1989-1994. *Paup*, version 4.0beta for Macintosh. Sinauer Assoc, Inc: Massachusetts.

Templeton, A. Crandall, K. Sing, C. 1992. A Cladistic Analysis of Phenotypic Associations with Haplotypes Inferred from Restriction Endonuclease Mapping and DNA Sequence Data. Genetics 132 (2): 619-633.

Templeton, A. Maskas, S. Cruzan, M. 2000. Gene Trees: a powerful tool for exploring the evolutionary biology of species and speciation. Plant Species Biology 15: 211-222.

Templeton, A. 2002. Out of Africa again and again. Nature: 45-51.

Wake, D. 1997. Incipient Species Formation in Salamanders of the *Ensatina* Complex. Proc. Natl. Acad. Sci. 94: 7761-7767.

Weyman, G. 1993. A Review of the Possible Causative Factors and Significance of Ballooning in Spiders. Ethology, Ecology, and Evolution 5: 279-291.

Weyman, G. Sunderland, K.D. Jepson, P.D. 2002. A Review of the Evolution and Mechanisms of Ballooning by Spiders Inhabiting Arable Farmland. Ethology, Ecology and Evolution 14: 307-326.

Williams, D. Ambrose, L. Browning, L. 1995. Trophic Dynamics of Two Sympatric Species of Riparian Spider (Araneae: tetragnathidae). Can. J. Zool. 73: 1545-1553.

Wright, S. 1931. Evolution and the Genetics of Populations. Volume 1. University of Chicago Press.

Wright, S. 1951. The Genetical Structure of Populations. Annu. Eugenics 15: 323-354.