

Detecting cryptic species in phylogeographic studies: Speciation in the California Slender Salamander, *Batrachoseps attenuatus*



Richard Highton*

Department of Biology, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Article history:

Received 13 August 2013
Revised 24 October 2013
Accepted 3 November 2013
Available online 22 November 2013

Keywords:

California Slender Salamander
Batrachoseps attenuatus
Plethodontidae
Speciation
Cytochrome-b DNA sequences
Phylogeography

ABSTRACT

A study of DNA sequence variation in the plethodontid salamander *Batrachoseps attenuatus* by [Martínez-Solano et al. \(2007\)](#) revealed more species than acknowledged by the authors. They sequenced 677 base pairs of the cytochrome-b mitochondrial gene in 178 individuals from 123 populations of the currently recognized species *B. attenuatus* from throughout most of its known range in southwestern Oregon and northern and central California. Their data show that the common ancestor of the species diverged into five clades during the late Miocene Epoch, an estimated 9.2–5.5 mya, with subsequent divergences producing at least 39 living lineages that replace each other geographically. These groups have been diverging independently from each other throughout the Pleistocene Epoch and many of them have probably reached the species level of divergence.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Two simple methods useful in determining how many species are present in a genetic distance data set are the Good–Wake method and the histogram method, reviewed by [Sites and Marshall \(2004\)](#). The first was suggested by [Good and Wake \(1992\)](#) and utilizes paired comparisons of genetic distances and geographic distances among populations. Within a species, the two distances are expected to be correlated because of gene flow between neighboring populations throughout the range of a continuously distributed species. A regression line showing the relationship between these two variables should intercept the y-axis near the origin. If a study includes multiple species, because of reproductive isolation between biological species, little or no correlation between genetic and geographic distances is expected. The regression line for between-species comparisons is usually parallel to the X-axis (geographic distance) with an intercept on the Y-axis (genetic distance) close to the mean genetic distance between the species. This estimates approximately how long a period of isolation there has been since the divergence of the two species.

I suggested a related method to distinguish between molecular data sets of one or more species ([Highton, 1998, 2000](#)). If a histogram of variation of all pairwise genetic distances includes a single species, the distribution is expected to be monomodal, while if two species are represented the distribution is expected to be

bimodal. If more than two species are present, the distribution is either bimodal or polymodal, depending on the number of included species and the relative times of the various speciation events. In four large allozyme data sets comprising a total of 50 species of plethodontid salamanders in [Highton \(1998: Fig. 2\)](#), polymodal distributions were present, and the cut-off between within- and between-species distributions in 302 pairwise comparisons in the four studies was close to a Nei distance of 0.15, indicating that all pairs of species within each of these four groups of salamanders have been evolving independently for >2 my ([Highton, 1989](#)). The histogram method has seldom been used to evaluate molecular phylogenetic data for species-detection studies, but [Marshall et al. \(2006\)](#), and [Puillandre et al. \(2012\)](#) found it useful.

The usual cut-off between genetic distances based on allozyme data within- and between-species was suggested earlier by [Baverstock et al. \(1977\)](#) and [Thorpe \(1982\)](#), the latter a comprehensive review of 100 allozyme studies of vertebrates. He found that, excluding birds, 97% of [Nei \(1972\)](#) *I*-values between species are <0.85 ($D > 0.16$), while within species 98% of *I*-values within species are >0.85, similar to the plethodontid data sets cited above. Using the calibration of [Maxson and Maxson \(1979\)](#), this amount of genetic divergence represents a time estimated to be about 2.1 million years of independent evolution, which apparently approximates the time it takes two isolated plethodontid salamander populations to diverge sufficiently to have evolved effective reproductive isolation (e.g., [Highton, 1989, 1995, 1997, 1998, 1999; Highton and Peabody, 2000](#)).

* Fax: +1 301 314 9358.

E-mail address: rhighto1@umd.edu

I applied the above methods to allozyme data for 27 other amphibian groups (Highton, 2000) and found that in those with large data sets there is about the same cut-off level of divergence between intraspecific versus interspecific comparisons (Nei D of 0.15) as that suggested in Thorpe's review. In their analysis of geographic DNA sequence variation in the mitochondrial gene cytochrome-b (cyt-b) of *Batrachoseps attenuatus*, Martínez-Solano et al. (2007) applied both the Good–Wake and histogram methods in their Fig. 3. These methods had been suggested earlier by Slatkin and Hudson (1991) and Mantel (1967) for other purposes in analyzing within-species variation. They called these methods “isolation by distance” and “mismatch distributions,” respectively. They did not use either method for their entire data set but instead they divided *B. attenuatus* into five of the most divergent groups, their “major mtDNA clades,” and then applied both methods only to comparisons within each of these five groups. Except for the Bodega Bay clade, which appears to be a single species, their histogram analyses for the other four major clades in their Fig. 3 clearly indicate that within each there are multiple species. Their Good–Wake distributions for the four major clades thus measure spatial patterns of cladogenesis superimposed upon intraspecies gene flow with isolation by distance.

In this paper the histogram and Good–Wake methods are used to estimate the number of species in the cyt-b sequence data of Martínez-Solano et al. (2007). This is a preliminary estimate of the number of taxa, and for several reasons, it needs to be confirmed before naming the numerous undescribed species. Although only a single mitochondrial gene was sequenced for their entire data set, reciprocal monophyly of geographic populations for mitochondrial haplotypes is used as the criterion for hypothesizing separate species lineages. Because of the maternal inheritance of mitochondrial genes, at contact or overlap zones between taxa it is difficult to distinguish (1) intrapopulation polymorphisms, (2) between-subspecies variation, (3) intergradation between subspecies, (4) sympatry of reproductively isolated sympatric species, and (5) hybridization between species. Moreover, 82% of their populations and 33% of the probable species included in their data set are represented by only a single individual, greatly increasing the likelihood of missing polymorphisms or cases of sympatry or hybridization between the taxa.

Many taxonomists now utilize mitochondrial DNA data as an aid in distinguishing species, discussed by Baker and Bradley (2006), who review the use of morphological and molecular data in defining species limits. They also discuss different species concepts and predict that use of their genetic species concept will result in a large increase in the number of species recognized by taxonomists. On the other hand, some molecular systematists consider mitochondrial sequence data not sufficient to define species (e.g. Moritz et al., 1992; Rubinoff et al., 2006). They require confirmation of species boundaries by nuclear genomic markers. In fact, Martínez-Solano et al. (2007) sequenced three other mitochondrial genes in 42 (24%) of their salamanders and found the results fully consistent with those of their cyt-b sequences. They also sequenced a nuclear gene (Rag-1) in 42 populations, but its genetic variation was not correlated with geography or the mtDNA data. It evolves too slowly to detect species lineages formed during the Pleistocene. They also compared their results with an earlier allozyme study of 13 populations by Yanev (1978). Despite the limitations of their data set, it provides a preliminary estimate of the number of species in the *B. attenuatus* species group. It is extremely unlikely that additional molecular markers would support the conclusion that *B. attenuatus* represents a single species.

2. Methods

The *B. attenuatus* haploid cyt-b sequences of Martínez-Solano et al. (2007) were used to calculate the pairwise sequence

differences using the MEGA program (Tamura et al., 2007). Phylogenetic trees were calculated by the maximum likelihood (ML) (Felsenstein, 1981) and neighbor-joining (NJ) (Saitou and Nei, 1987) methods using the MEGA program. Both the histogram and Good–Wake methods were used to compare groups at various levels of divergence in order to look for geographic patterns of haplotype variation and evidence of present gene exchange or genetic admixture between populations. Using the MEGA program, the ML tree based on the cyt-b data (677 bp) is shown in Fig. 1, but the 173 individuals are numbered in the order they appear on the NJ tree since its topology is more highly supported by the bootstrap method (Felsenstein, 1985) with 2000 replications. The locality numbers used by Martínez-Solano et al. (2007) are given in Table 1 to compare with the numbers of individuals and haplotype groups (potential species) used in this paper.

The often-used estimate of the rate of evolution in the cyt-b gene is about 1 my for each 1% sequence divergence (%sd) (Johns and Avise, 1998), which is similar to the estimates of divergence time used by Martínez-Solano et al. (2007). Since they sequenced 677 bp, a sequence divergence of 7 bp is equivalent to about 1.03 million years (my) of divergence. I found there are 39 different mtDNA cyt-b lineages of probable species, which I call haplotype groups, that are estimated to have been diverging from each other for >1.8 my. The five ancient, most distantly related haplotype groups are called “major mtDNA clades” and “main clades,” by Martínez-Solano et al. (2007), but I call them species complexes using the same common names that they used: Northern, Eastern, Bodega Bay, Southern (North) and Southern (South).

It should be emphasized that the data for 178 individuals include only that many independent observations. Using a data set of 15,753 pairwise comparisons among these individuals includes only 178 independent observations. However, the larger data set is more enlightening because it eliminates much of the sampling error that would occur if only a small sample of randomly selected pairs of comparisons between each pair of haplotype groups is used. No statistical tests rely on these comparisons.

3. Results

3.1. Bootstrap support for haplotype groups

Bootstrap support for the NJ tree provides an indication of confidence for the 26 haplotype groups that include >1 individual. Nineteen are supported at the >95% level (mean = 89.7%) (Table 1). The mean bootstrap support would be higher (96.5%) except for three groups (24, 31, 39) with low support (51%, 34%, 28% respectively). The mean bootstrap support for the 26 haplotype groups in the ML tree (Fig. 1) was similar (89.3%) to that of the NJ tree and the same 39 haplotype groups are present. They are also present in the tree in Martínez-Solano et al. (2007). Although the content of most of the haplotype groups is strongly supported, the relationships of the groups within species complexes were generally weakly supported, as in the tree in Martínez-Solano et al. (2007). There is strong support (>97%) for three of the four species complexes with multiple species (Northern, Eastern, and Southern [North]) similar to that in the tree in Martínez-Solano et al. (2007).

3.2. Detecting species by the histogram method

Fig. 2A is a histogram of the 15,753 pairwise comparisons of cyt-b base-pair differences among the 178 individuals in the study of *B. attenuatus* of Martínez-Solano et al. (2007). These histograms do not include the important geographical component of variation, which is discussed below. The histogram is polymodal with numerous groups that are estimated to have diverged at several

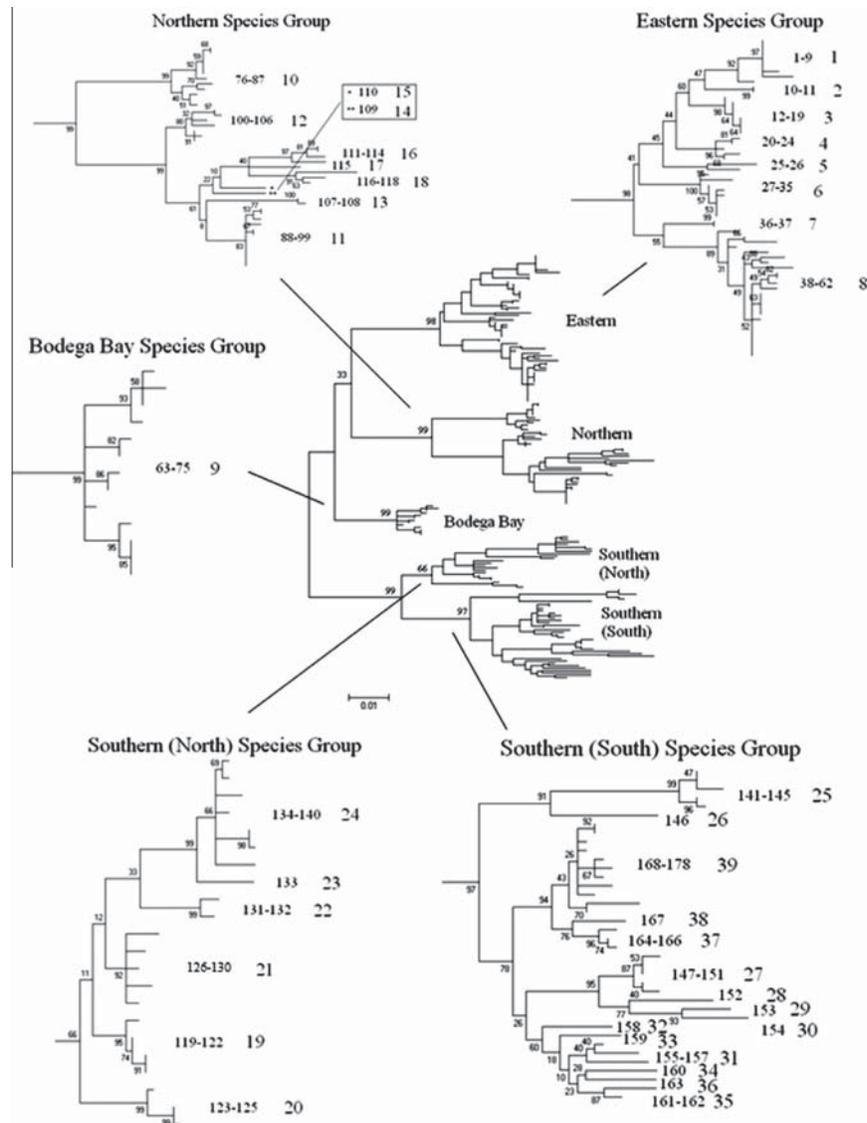


Fig. 1. Maximum likelihood tree of the *cyt-b* sequence data of Martínez-Solano et al. (2007) for 178 individuals in 123 populations of *B. attenuatus*. The small numbers to the right of each haplotype group represent the 178 individuals in their study rearranged in the order they appear on the NJ tree, and the larger numbers are those of the 39 numbered haplotype groups delimited in this study. Support values for each node are based on 2000 bootstrap trees.

different times during the late Miocene and throughout the Pliocene Epochs. These include 39 haplotype groups that have been diverging from each other for at least 1.8 my (number of bp difference >12) with little or no evidence of gene exchange among the groups throughout most of the Pleistocene Epoch. This suggests that many, if not all, are reproductively isolated and have reached the species level of divergence. These candidate species are arranged into five species complexes, which diverged from each other between 9.5 and 5.5 mya (the “major clades” of Martínez-Solano et al., 2007) during the latter half of the Miocene Epoch.

There is an almost complete separation of the within- and between-haplotype group distributions at 12 bp. Within the 39 haplotype groups, 772 of 777 pairs (99.4%) have a sequence difference ≤ 12 bp (Fig. 2D), while in the between-haplotype group pairs, 14,904 of 14,976 (99.5%) have a sequence divergence >12 bp (Fig. 2B). A difference of 12 bp is estimated to represent about 1.8 million years (my) of divergence, close to the cut-off of within- and between-species comparisons in other vertebrate genera based on allozyme data discussed above (Thorpe, 1982), as well as in many studies of plethodontid genera (e.g., *Plethodon* and *Ensatina*, Highton, 1989, 1998, 1999; Highton and Peabody, 2000).

The within-haplotype group distribution (those comparisons usually <12 bp) includes both within- and between-population haplotype variation. The difference between within-population variation (mean = 1.8; range, 0–9; $n = 161$, Fig. 2F) and between-population variation (mean = 4.9; range, 0–14, $n = 607$, Fig. 2E) within the haplotype groups is partially responsible for the bimodality of the overall within-haplotype group distribution with modes at 1 and 4 bp in Fig. 2D. The 39 numbered haplotypes or haplotype groups are terminal branches in the tree (Fig. 1) and are estimated to have been diverging from each other independently throughout the Pleistocene Epoch. Table 1 shows the amount of variation within each haplotype group. Fig. 2G shows the within-population distribution without taking into account haplotype groups and shows additional sequence differences in populations where more than one haplotype group is represented at a single locality, the latter with the divergence between the haplotypes of the different haplotype groups ranging from 14 to 21 bp (see below). It is unlikely that at these four localities there are within-population variants of a single species because in all four cases, the different haplotypes are identical or well within the known range of variation of a parapatric haplotype group. The se-

Table 1

Mean and range of sequences differences in within-haplotype group comparisons of the 39 groups of the *Batrachoseps attenuatus* complex. Bootstrap percentages are based on 2000 trees.

Species Complex	Haplotype Group	Individuals	Mean	Range	% Bootstrap Support on NJ Tree	Numbers of populations in Appendix of Martínez-Solano et al. (2007)
Eastern	1	1–9	2.4	0–9	92	69–73
	2	10–11	0	0	100	67–68
	3	12–19	0.2	0–2	99	59 (DBW 6375–79, 6382–84)
	4	20–24	3.4	0–6	96	60–64
	5	25–26	10.0	10	95	65–66
	6	27–35	2.1	0–5	99	54, 56, 57, 59 (DBW 6380–81)
	7	36–37	0	0	100	55
	8	38–62	4.6	0–14	96	37–53, 58
Bodega Bay	9	63–75	6.6	0–11	100	29–36
Northern	10	76–87	4.7	0–10	100	1–12
	11	88–99	1.5	0–4	88	25–26, 28 (DBW 6351–57, DBW 6360–61, MVZ 221030)
	12	100–106	4.6	0–7	95	24, 27, 28 (DBW 6358)
	13	107–108	1.0	1	100	22–23
	14	109	n/a	n/a	–	13
	15	110	n/a	n/a	–	21 (MVZ 151997)
	16	111–114	5.3	0–8	97	14–16, 21 (MVZ 152018)
	17	115	n/a	n/a	–	20
	18	116–118	8.7	4–12	95	17–19
Southern (North)	19	119–122	2.8	0–5	97	87, 88 (MVZ 224437, 224446), 89
	20	123–125	4.0	1–6	100	90–92
	21	126–130	6.4	4–9	90	74, 85–86, 88 (MVZ 24438)
	22	131–132	5.0	0–13	100	83–84
	23	133	n/a	n/a	–	81
	24	134–140	6.8	1–12	51	75–80, 82
Southern (South)	25	141–145	4.0	0–8	100	105–108
	26	146	n/a	n/a	–	104
	27	147–151	2.6	0–4	97	110–111
	28	152	n/a	n/a	–	112
	29	153	n/a	n/a	–	113
	30	154	n/a	n/a	–	114
	31	155–157	9.3	6–13	34	116–117, 119
	32	158	n/a	n/a	–	102
	33	159	n/a	n/a	–	122
	34	160	n/a	n/a	–	115
	35	161–162	8.0	8	84	118, 121
	36	163	n/a	n/a	–	120
	37	164–166	2.7	1–4	99	93–95
	38	167	n/a	n/a	–	96
	39	168–178	5.6	0–13	28	97–101, 103, 109

quence differences between all 11,925 pairwise comparisons between haplotypes of different species complexes are shown in [Fig. 2C](#). [Fig. 3](#) shows the amount of sequence variation within each of the haplotype groups with sample sizes of five or more individuals.

Taking geography into consideration, the haplotype groups have ranges that are almost all parapatric and allopatric to each other. There is one isolated locality (population 123) east of the previously known range of coastal *B. attenuatus* in the central valley of California, where there are two individuals of group 39, far outside the range of the remaining members of that group (see below). There are also four localities that have members of two different haplotype groups.

[Figs. 4–7](#) show the geographic ranges of the five species complexes and the 39 haplotype groups. Because of the few populations in some of the groups, the geographic limits of their ranges may not be represented accurately, especially those thirteen groups represented by only a single individual. All but one of the contiguous haplotype groups that occur in geographic units are spatially allopatric or parapatric to all other groups. Such a good match between haplotype groups and geographically contiguous units would not be expected if all 123 populations of the study belonged to a single widespread species of interbreeding populations. In that case, considerable evidence of gene flow at contacts would be expected, resulting in genetic admixture and far more cases of

geographic overlap of different haplotype groups. For mitochondrial DNA data, frequent occurrence of haplotypes of other groups would be expected at group boundaries where intergradation was occurring, but in spite of at least 35 geographic contacts between haplotype groups, there are only four in which haplotypes of two parapatric groups were found at the same locality. None of these are at the five contacts between species complexes.

The four cases in which samples with two different haplotype groups occur at a single locality are all in the vicinity of contacts between parapatric groups. As discussed above, each of these might represent a polymorphism within a single population, different subspecies, intergrades between subspecies, different species, or hybrids between species. The four populations, all in California, which include two different haplotype groups are:

- (1) Two individuals from [Martínez-Solano et al. \(2007\)](#) locality 21 of haplotype groups 15 and 16 of the Northern Species Complex ([Fig. 4](#)) have very different haplotypes, differing by 17–19 bp. MVZ 151997 is the only individual of group 15, so the extent of its range is not yet known, while MVZ 152018 has a haplotype that is very similar to those of haplotype group 16 with bp differences ranging from 0 to 8 bp to that group. Other individuals of Group 16 were taken in southern Humboldt and northern Mendocino counties (their localities 14–16), while their locality 21 is in southern Men-

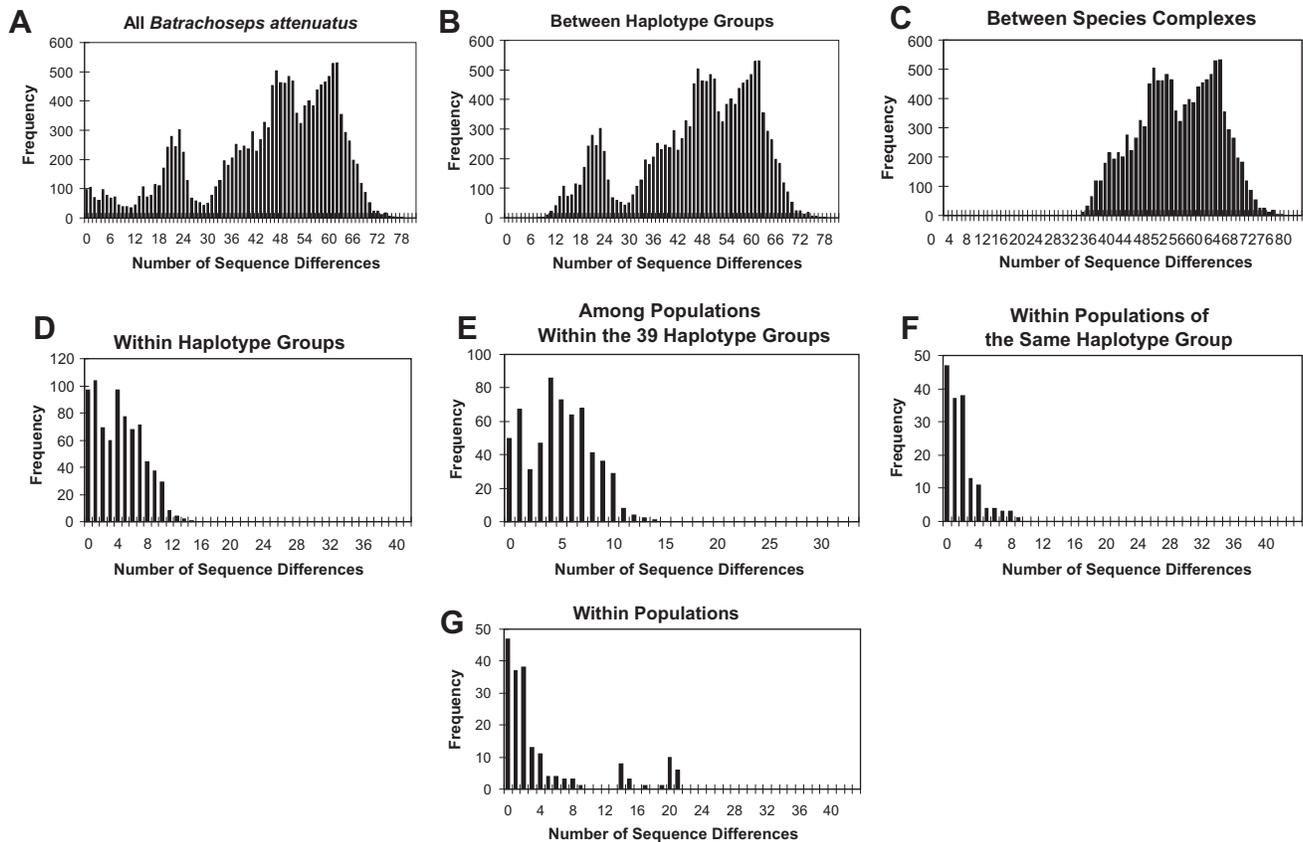


Fig. 2. Histograms of pairwise comparisons of bp differences within and between the various levels of divergence in *B. attenuatus*. A. All 15,753 pairs. B. 14,976 comparisons between pairs of individuals of different haplotype groups. C. 11,925 comparisons between species complexes. D. Between individuals within all haplotype groups. E. Among individuals between populations within haplotype groups. F. Among individuals within populations of the same haplotype group. G. All within-population comparisons.

docino County. These two areas appear separated by the range of haplotype group 18 with three localities in central Mendocino County (their localities 17–19). Haplotype group 16 is the only one in the entire study that has a range separated into two segments by another haplotype group (Group 18). However, both segments of Group 16 might be connected, either to the east or west of the range of group 18. To the west, they might even have been connected in an area that is now under the Pacific Ocean when sea level was lower during the glacial periods of the Pleistocene Epoch.

- (2) Eleven individuals from their locality 28 of the Northern Species Complex (Fig. 4) were sequenced. Ten belong to haplotype group 11, also known from two other localities (25, 26) in Sonoma County. All have similar sequences with bp differences of 0–4. One (DBW 6358) has a sequence that differs by 14–15 bp from the other ten salamanders, but is similar to those of six individuals from two nearby Sonoma County localities (24, 27) of haplotype group 12 with only 0–6 bp differences to that group. The distance between localities 27 and 28 is 0.31 km, across the Russian River from one another. A man-made bridge crosses the river only 0.2 and 0.3 km from two sites respectively (Martínez-Solano et al., 2007). The two groups may be separated by the river, as rivers often separate species of completely terrestrial salamander species (e.g., Highton, 1972, 1989, 1999, 2009). Individuals of population 27 might have crossed the bridge since they or the descendants of females that did were collected only 0.2 km away from the bridge into the range of population 28 of group 11, or they may have crossed the river by natural means.

- (3) A series of 10 salamanders from Napa County in the Eastern Species Complex (Fig. 5) from locality 59 includes eight individuals of haplotype group 3 (DBW 6375–79, 6382–84) and two individuals from haplotype group 6 (DBW 6380–81). Haplotype group 3 is known only from this locality. Haplotype group 6 includes seven additional individuals from three nearby sites (54, 56, 57). The differences between these two parapatrically distributed haplotype groups range from 20 to 21 bp.
- (4) At their locality 88, three individuals were sequenced, two of which (MVZ 224437, 224446) belong to haplotype group 19 and one (MVZ 224438) belongs to haplotype group 21, parapatric groups belonging to the Southern (North) Species Complex (Fig. 6). Group 19 occurs at two other nearby sites (their localities 87 and 89) in Marin County, and group 21 also occurs in Sonoma and Marin counties (populations 74, 85–86). The bp sequence differences between these two groups at locality 88 range from 15 to 17.

At their locality 123 in Stanislaus County, there are two individuals of haplotype group 39, a group in the Southern (South) Species Complex (Fig. 7), which otherwise occurs in and south of the San Francisco Bay Peninsula. Locality 123 is >125 km to the east of the range of group 39, across a region occupied by several other groups (30–31, 33–36) and in an area otherwise outside the known range of *B. attenuatus*. Martínez-Solano et al. (2007) suggest that the population there may be the result of a human introduction. Since salamanders are often used as fish bait and are also sometimes kept as pets, it is possible that this population is derived from individuals collected within the range of haplotype group

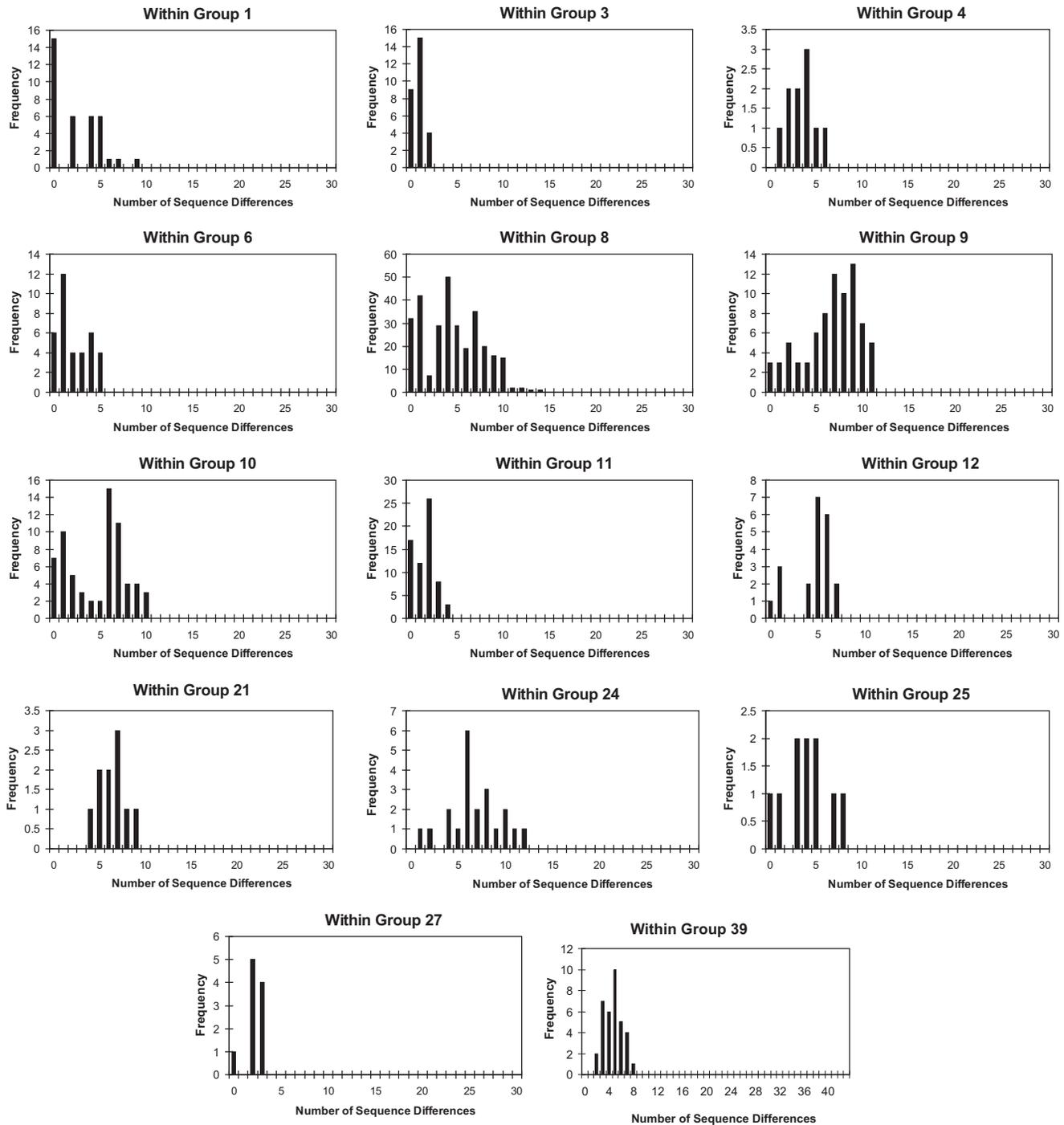


Fig. 3. Histograms of pairwise comparisons within the 14 haplotype groups with sample sizes ≥ 5 individuals.

39. This population is not included in my analysis, but its location is shown in Fig. 7.

There are 741 pairwise comparisons within and between the 39 haplotype groups. Histogram comparisons between those with small samples (1 or 2 individuals) are not informative. Selected for illustration (Fig. 8) are 16 examples of paired comparisons that include groups with the most individuals, often from nearby populations, both within and between species complexes. All have bimodal distributions and clearly show that the haplotype groups represent different species. Fig. 9 provides histogram comparisons within the five species complexes, indicating that four (except for Bodega Bay) are composed of multiple species. These are the same

groups present in Martínez-Solano et al. (2007: Fig. 3). Fig. 10 shows the 10 pairwise histogram comparisons among the five species complexes. All histograms are completely separated from all distributions within haplotype groups (see Figs. 2 and 3). The hypothesis that any of the 557 pairs of populations in haplotype groups of different species complexes have interbred or amalgamated with each other has no support.

3.3. Detecting species by the Good–Wake method

Fig. 11 shows the 10 paired comparisons among all five species complexes using the Good–Wake method. None of the

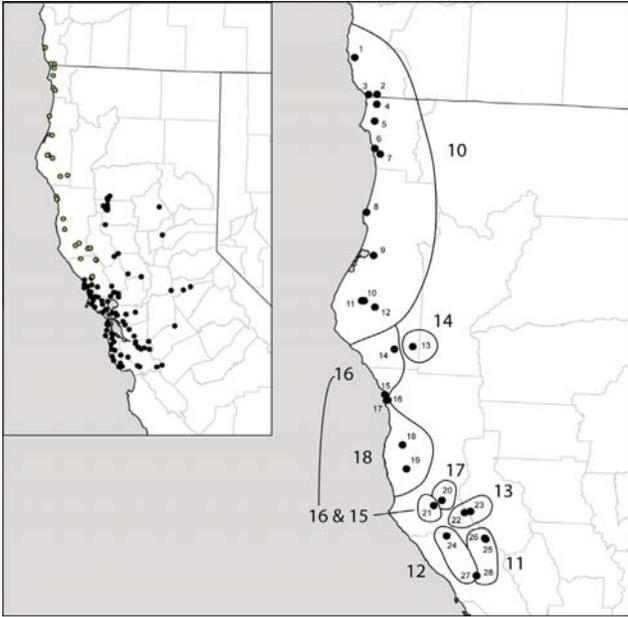


Fig. 4. Geographic ranges of the haplotype groups within the Northern Species Complex. The inset shows the distribution of the Northern Species Complex (green) and the entire data set of *B. attenuatus*. Population numbers are those of Martínez-Solano et al. (2007) and larger-sized numbers are those of the haplotype groups. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

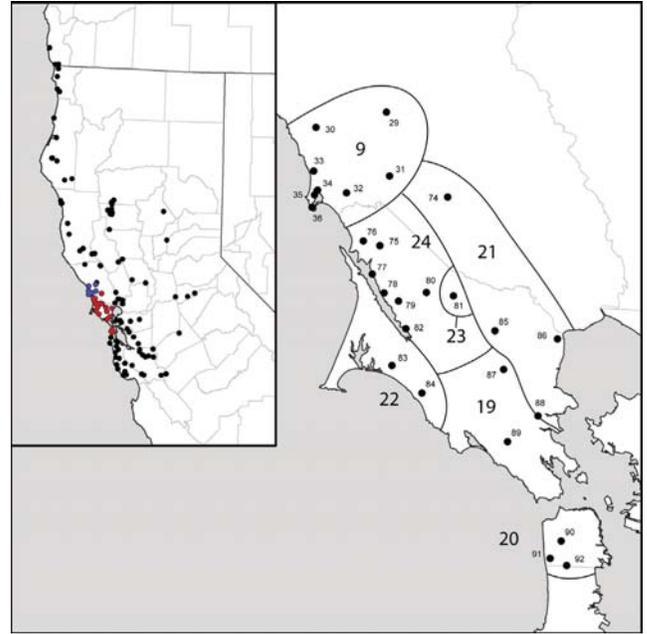


Fig. 6. Geographic ranges of the haplotype groups within the Bodega Bay (dark blue) and Southern (North) (red) Species Complexes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

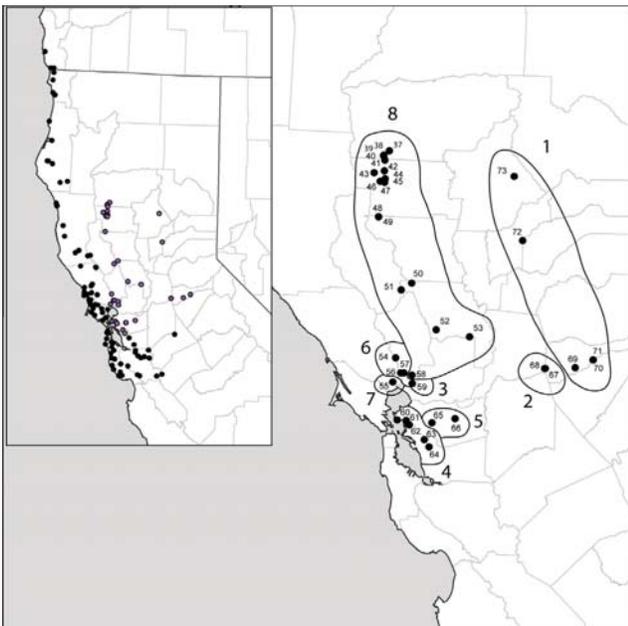


Fig. 5. Geographic ranges of the haplotype groups within the Eastern Species Complex (purple). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

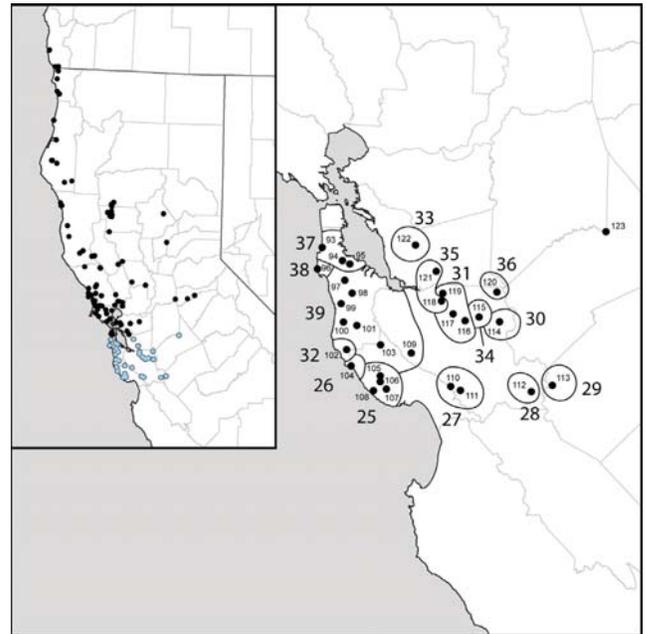


Fig. 7. Geographic ranges of the haplotype groups within the Southern (South) (light blue) Species Complex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

regression lines approach the origin, and there is no correlation between genetic divergence and geographic distance in any of the comparisons. This indicates that there is no evidence of gene exchange among the different species complexes since they were isolated from one another before the end of the Miocene Epoch, even though some populations in five of the pairs of species groups now occur within 10 km of each other. Thus the hypothesis that any of the 557 between-species complex pairs belong

to the same species is again not supported. All pairs of probable interspecies comparisons within species complexes are also supported as species by Good–Wake tests. Of the 184 paired comparisons within-species complexes, 24 were selected for illustration (Fig. 12). These are mostly of different haplotype groups within species complexes with the largest sample sizes. They show little or no correlation of genetic divergence and geographic distance, thereby indicating lack of support for the hypothesis that any one of these 24 pairs of different haplotype

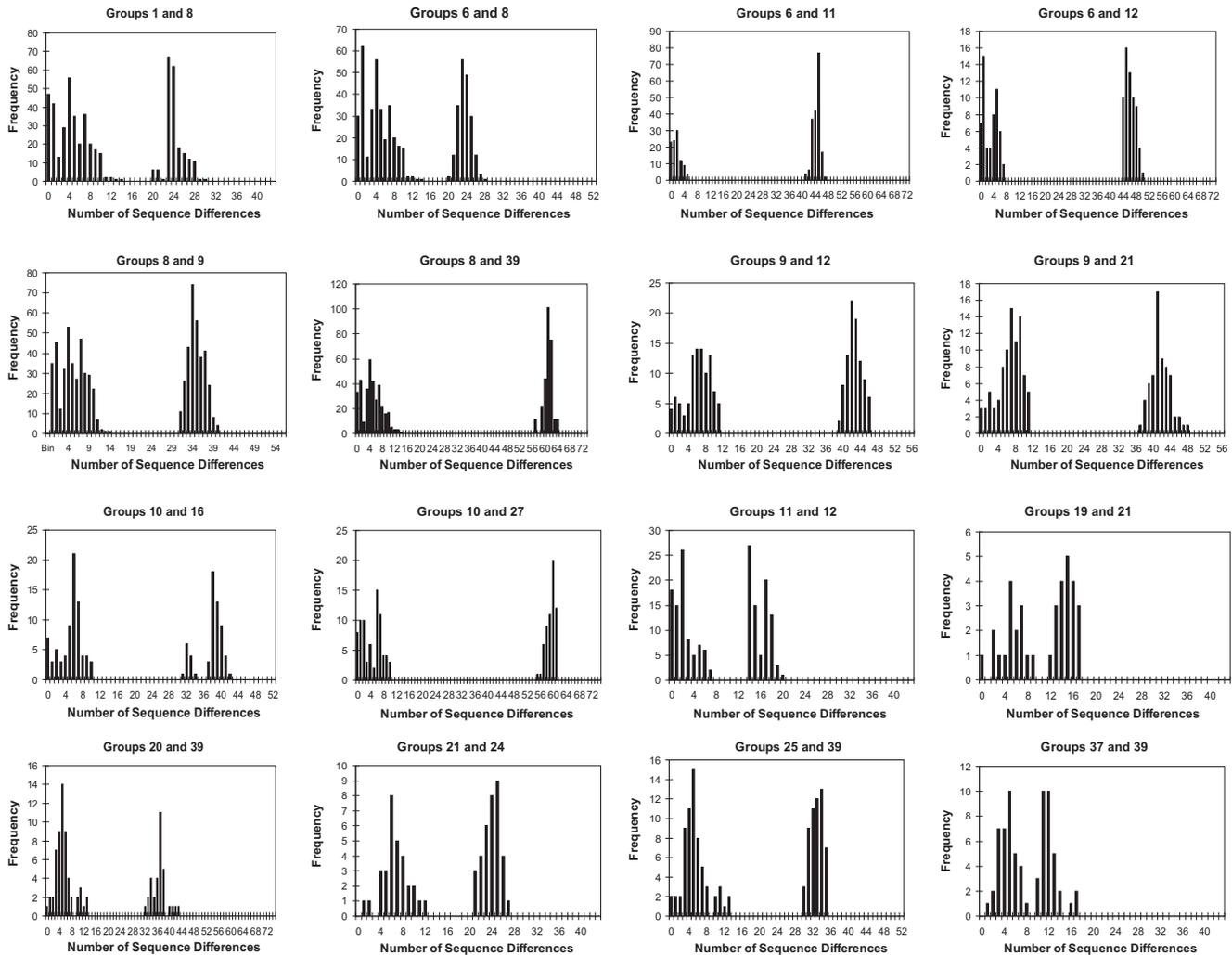


Fig. 8. Histogram comparisons of within and between 16 pairs of haplotype groups, selected on the basis of their geographic proximity and/or large samples sizes. All are bimodal and there is no overlap in the range of the within- and between-group distributions.

groups represents a single species. Six of the 24 between-haplotype groups in Fig. 12 include pairs of haplotype groups that were taken less than 10 km from each other. Because of the small sample sizes among so many of the pairwise comparisons within haplotype groups, considerable additional information is needed to use this method to support the hypothesis that all groups represent different species. It is expected that they too will be confirmed as species because of the patterns of variation indicated above in the histogram comparisons, but the correct decision in each case remains to be determined by further molecular or morphological analysis, or reproductive-isolation studies to determine the interaction of each pair in their parapatric contact and/or overlap zones. Especially interesting would be an analysis of populations 57 (Group 6), 58 (Group 8) and 59 (Group 3) of the Eastern Species Complex, which occur within 6–11 km of one another (Fig. 5), and populations 20 and 21 of Groups 15, 16 and 17 of the Northern Species Complex, which occur within 8 km of each other (Fig. 4).

Fig. 13 shows Good–Wake comparisons within the 15 haplotype groups with the most individuals. Most have a positive correlation between genetic distance and geographic distance as expected in within-species comparisons. The low genetic distances show the close relationships among individuals within all groups, as expected within species.

Fig. 14 combines all the Good–Wake tests for 763 paired-comparisons among individuals within each of the haplotype groups with >1 individual, showing a positive correlation between genetic and geographic distances, which is expected in within-species comparisons. This was not found in any comparisons between haplotype groups or between species complexes using the Good–Wake method. Since there are no within-group comparisons when only one individual occurs in a haplotype group, there are no data on within-group comparisons for the 13 haplotype groups with only one individual.

4. Discussion

Both methods for detecting species from genetic distance data reach the same conclusion: there are probably about 39 species within the currently recognized single species *Batrachoseps attenuatus*. Most arose in the late Pliocene and early Pleistocene Epoch and apparently have been isolated from each other for at least 1.8 my, with little or no evidence in the mtDNA data of gene exchange among the haplotype groups since their isolation. These are arranged in five species complexes, which were isolated much earlier, probably in the late Miocene Epoch, with no evidence of gene exchange since their original isolation. Further studies using

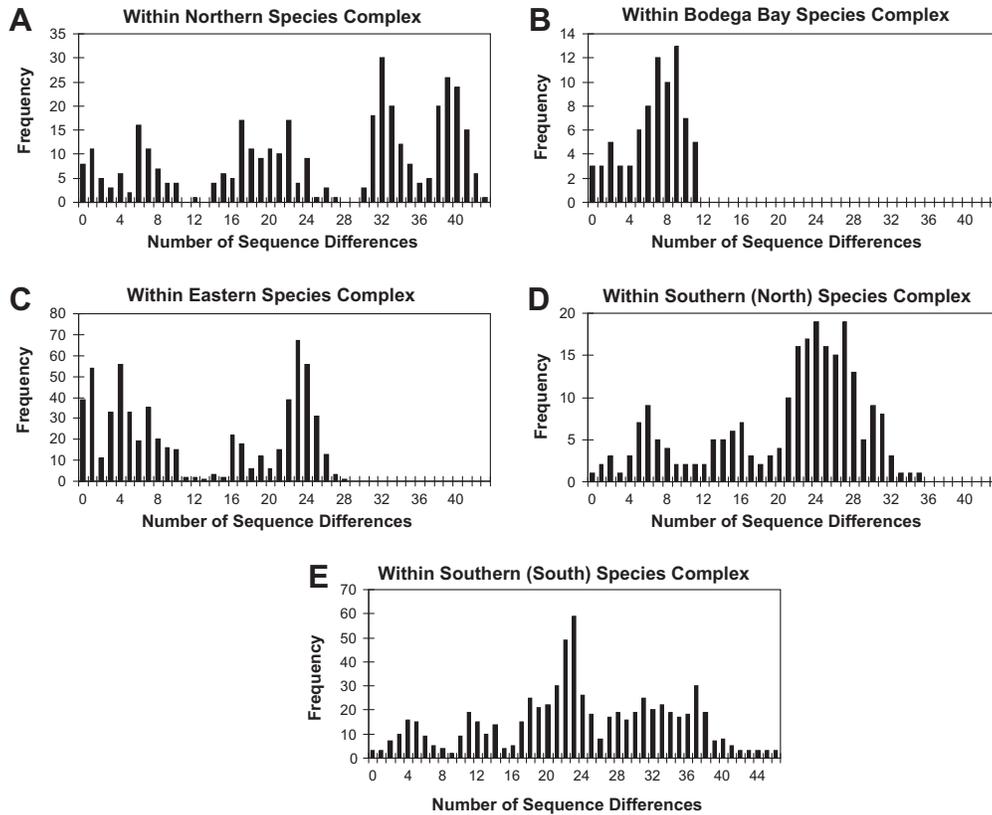


Fig. 9. Histogram comparisons between all pairwise haplotype groups within each of the five species complexes. All but the Bodega Bay form are polymodal.

nuclear genes from a larger number of populations and/or morphological studies are required to establish the validity of each of the probable undescribed species represented in the [Martínez-Solano et al. \(2007\)](#) data set.

The five species complexes in the ML tree ([Fig. 1](#)) are estimated to have diverged from each other at various times between 9.2 and 5.5 mya. During the late Pliocene and early Pleistocene Epochs, each of four of these species complexes underwent numerous speciation events between 4.8 and 1.8 mya, producing the present 39 identified haplotype groups. Many of these experienced few or no splitting events for a considerable period during the first half of the Pleistocene Epoch, suggesting that their ranges were subdivided into small pockets isolated from each other, resulting in speciation. More recently, they probably have expanded their ranges until they contacted other nearby groups so that the present ranges of most haplotype groups now have a parapatric pattern of geographic relationships with other groups, and more recently local populations within the groups are beginning to diverge from one another. A long period of isolation in the first half of the Pleistocene Epoch would explain the large amount of geographic differentiation among all of the haplotype groups. Genetic mixing would not produce intermediate haplotypes but it would produce disparate haplotypes in and near contact zones. The almost complete absence of any present genetic overlap between these groups may indicate that few or no haplotype transfers among any of the groups took place during much of the Pleistocene Epoch, yielding the sharp cutoff between the within- and between-group distributions at 12 bp.

There are many cases in which parapatric pairs within species groups are sister groups on the tree ([Fig. 1](#)), but there are also cases in which parapatric populations are among the most divergent populations within a species group. Examples are groups 10 and 16 in [Fig. 4](#), groups 3, 6, and 8 in [Fig. 5](#), groups 19, 22, and 24 in [Fig. 6](#), and groups 26, 32, and 39 in [Fig. 7](#). So many such divergent

parapatric groups would not be expected if all were interbreeding freely with adjacent populations throughout the Pleistocene Epoch. The same is true for the six pairs of parapatric species-group contacts.

The four cases of sympatry of pairs of different haplotype groups need further study to discriminate between the hypotheses of intraspecific polymorphism, hybridization between different lineages, and sympatry of distinct species. Sequences of fast-evolving nuclear genes or allozyme studies of fast-evolving genetic loci might reject two of these alternatives in each case.

Fourteen populations from the San Francisco Bay area were studied by [Martínez-Solano and Lawson \(2009\)](#), where three of the different species groups are present on islands in the bay and nearby mainland sites. They sequenced *cyt-b* DNA from salamanders of each population in order to assign each to its species complex. At two sites on Yerba Buena Island, they found haplotypes of two different species complexes: Southern (North) and Southern (South). They report 13 and 9 individuals from these two sites but do not indicate how many of each species complex they identified at the two localities. They discuss allozyme data from 111 additional specimens from Yerba Buena Island, but do not provide the results. Dr. Martínez-Solano kindly sent me the allozyme data on the individuals they studied, but the species complex for each of these individuals is unknown since they are not the same individuals they identified to species complex by sequence data. The 21 allozyme genetic loci that they analyzed provide insufficient information to distinguish reliably either the species or the species complex of individuals.

There are major differences in allele frequencies between the three species complexes (Eastern, Southern [North], and Southern [South]) at only two or three of the 21 loci they analyzed, but there are no fixed or complete differences between species or species complexes at any of the 21 loci in populations in the San Francisco Bay area. It would be especially important to study the genetic

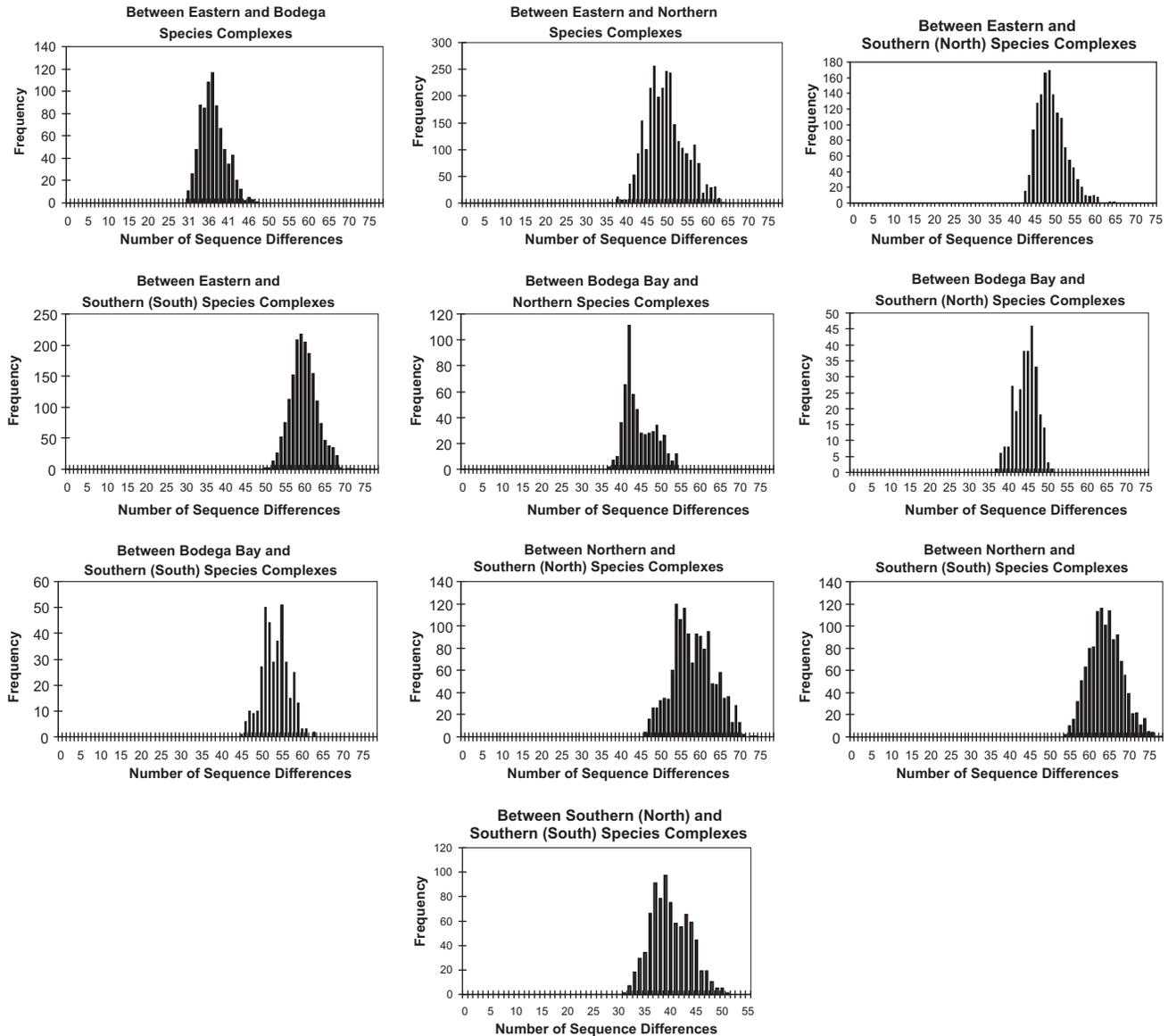


Fig. 10. Histogram of the 10 between-species complex pairwise comparisons between all individuals of each of the five species complexes.

interaction of the two populations of different species complexes on Yerba Buena Island.

In both of their papers on *B. attenuatus* (Martínez-Solano et al., 2007; Martínez-Solano and Lawson, 2009), the possibility that it is a complex of cryptic species was considered. They suggest that the low genetic distances in allozyme comparisons among the major groups and the rarity of fixed differences are within the range found in intraspecific studies in the genus *Batrachoseps* (Yanev, 1978). However, it is now known that the number of species in the genus was then underestimated. They now recognize 21 species compared to the 10 in 1978 (Jockusch and Wake, 2002; Jockusch et al., 2012; Martínez-Solano et al., 2012), so that some intraspecific differences in 1978 are now known to be interspecific distances.

One reason for the low genetic distances in the Martínez-Solano and Lawson (2009) allozyme data set may be the selection of genetic loci used for allozyme studies of western plethodontid salamanders by these workers. I previously noted (Highton, 1998) that Berkeley researchers usually did not include fast evolving loci in their studies, e.g., the blood proteins and esterases that I always

use in my allozyme studies referenced above. Their genetic distances are thus not comparable to many of the allozyme studies of other salamander studies, probably making the published Nei *D*-values of their species significantly lower compared to those within species of many other plethodontid studies. I calculated that some good reproductively-isolated sympatric species of eastern *Plethodon* would have very low *D*-values if the four fast-evolving loci had not been included in my studies; e.g. the sympatric eastern species *P. cinereus* and *P. shenandoah* (Highton, 1999) would have only a mean Nei *D* = 0.02 if esterases and blood proteins had not been included. This is likely the case in the allozyme data set for the 14 populations of *B. attenuatus* in the San Francisco Bay area (Martínez-Solano and Lawson, 2009: Table 5). These populations included three different species complexes, yet none of the Nei genetic distances are >0.18. Those between two populations of the Eastern Species Complex compared to 12 populations of the two Southern (North) Species Complex range from 0.08 to 0.18 (mean = 0.12), while those among the first six populations in their Table 5, all within the Southern (North) Species Complex range from 0.01 to 0.04 (mean = 0.02). Most of the remaining *D*-

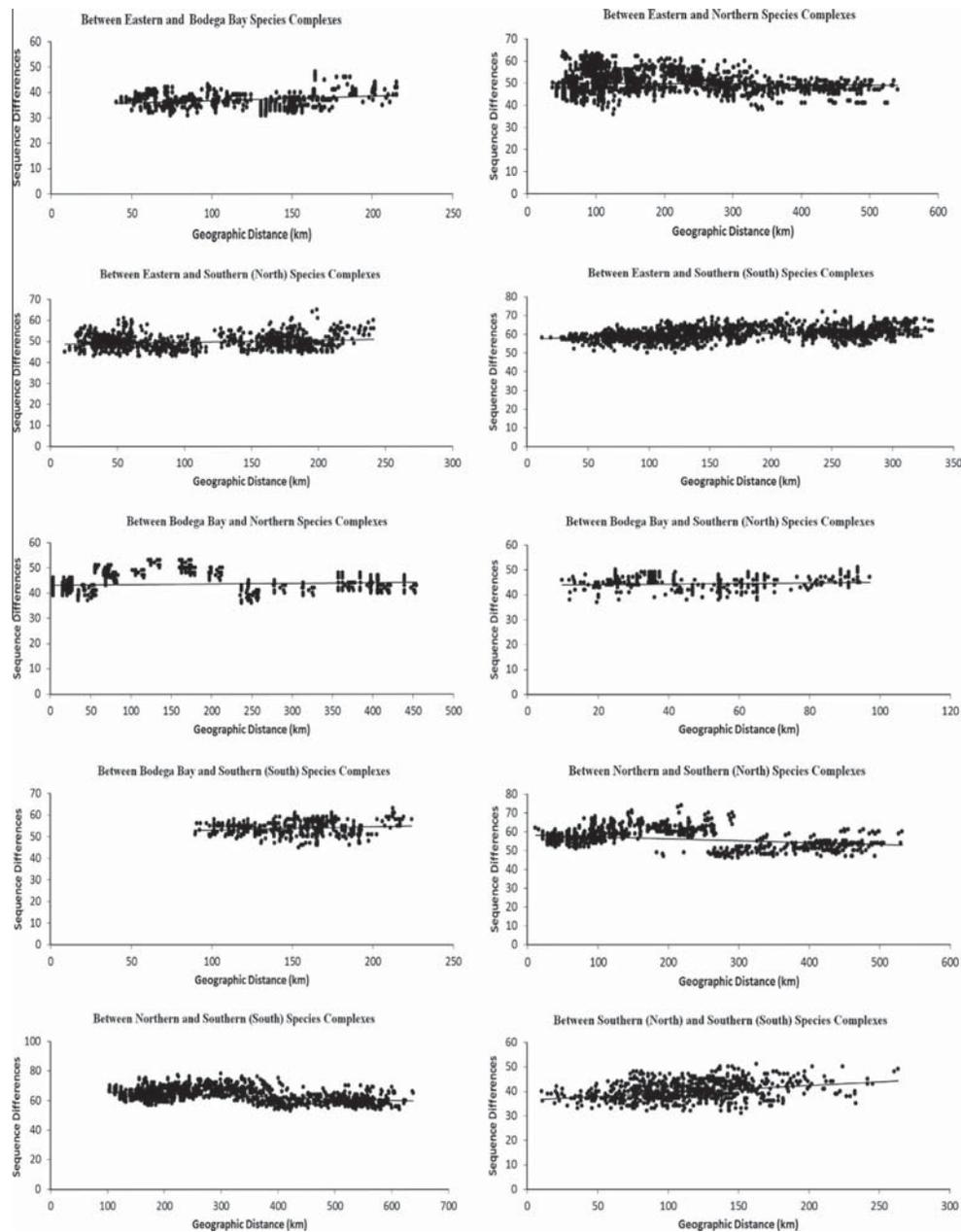


Fig. 11. Good-Wake tests of all 10 between-species group pairwise comparisons including all individuals in each of the five species groups. All regression lines are parallel to the X-axis.

values in their table are intermediate between these two groups, as expected because they probably include comparisons between the Southern [North] and Southern [South] species complexes, as well as within- and between-different species within those complexes. All of these are more closely related to each other than any are to the Eastern Species Complex. Moreover, they may include comparisons involving populations of members of both the Southern (North) and Southern (South) species complexes from Yerba Buena Island that they identified by their mt-DNA sequence data. Whether their allozyme data include both of the species complexes from Yerba Buena Island is not known because the same individuals were not used in the two studies. Thus they missed an opportunity to discover, by analyzing variation in nuclear genes, whether or not there is interbreeding between these two very different sympatric species of different species complexes on that island.

Studies on other species of *Batrachoseps* have steadily increased the number of recognized species in the genus in the last 60 years from three (Schmidt, 1953) to the present 21 (in two subgenera and probably at least seven species groups) including studies (Wake, 1996; Jockusch et al., 1998; Wake and Jockusch, 2000; Jockusch et al., 2001; Martínez-Solano et al., 2012) that recognize numerous species with ranges of small size. Most of these have diagnostic morphological characters as well as concordant molecular differences. It is therefore not surprising that the *B. attenuatus* species group is also polytypic.

Anderson (1960) compared variation in several morphological characters of *B. attenuatus* from several islands in San Francisco Bay and nearby mainland populations. He found a number of differences between populations. These morphological differences may be useful in the taxonomic analysis of these populations and might be helpful in diagnosing some of the 10 probable taxa in

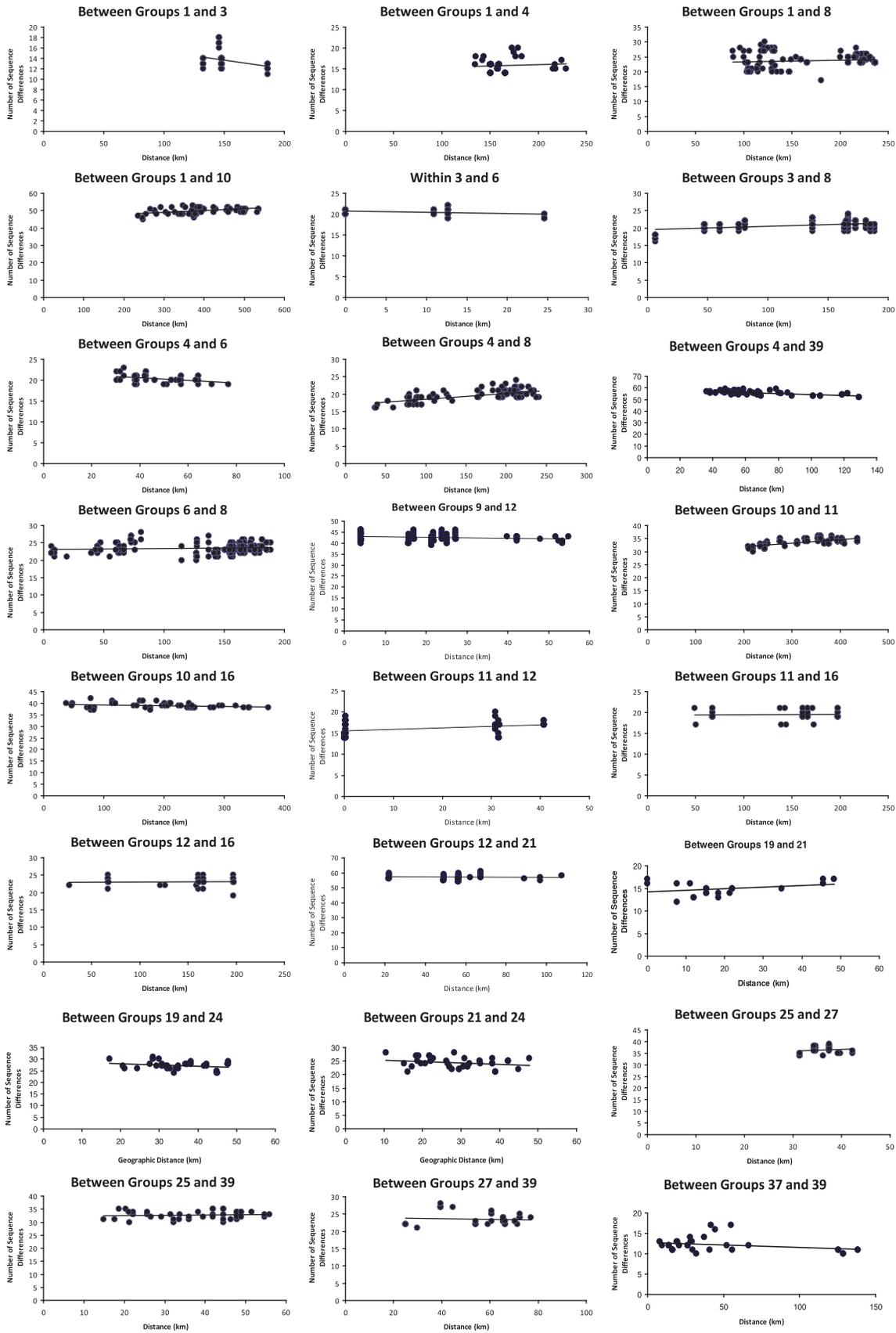


Fig. 12. Good–Wake pairwise comparisons between 24 selected pairs ($n \geq 5$ in each group) of probable species in different haplotype groups. All but two were selected on the basis of large sample sizes of both groups.

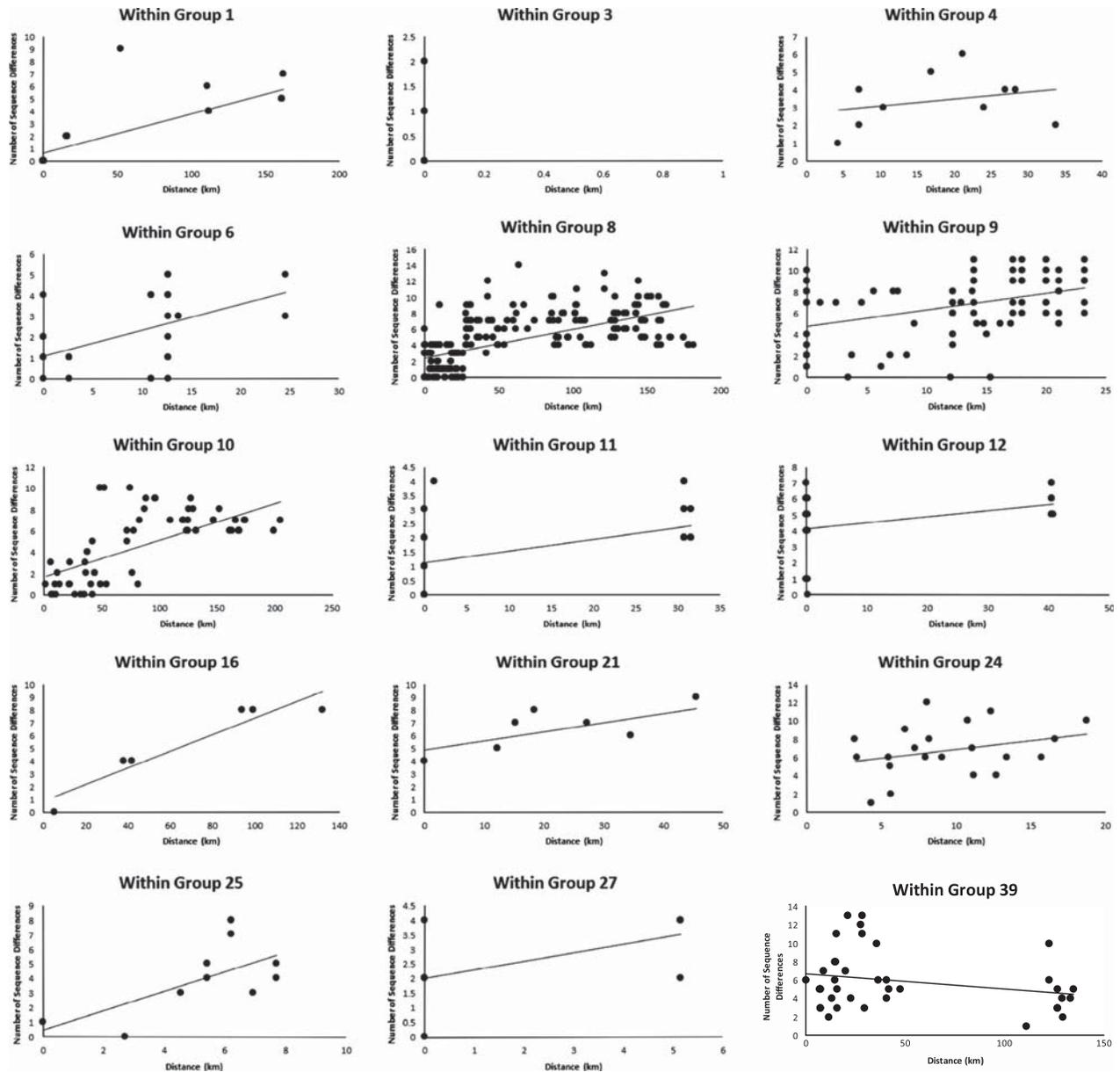


Fig. 13. Good-Wake pairwise comparisons within each of the 15 haplotype groups with sample sizes ≥ 4 individuals. There is a positive correlation between genetic distance and geographic distance in all but group 3 (all individuals are from the same population) and 39.

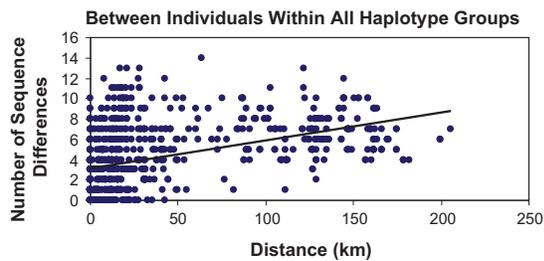


Fig. 14. Combined Good-Wake pairwise comparisons within all 25 haplotype groups with sample sizes ≥ 2 individuals.

three species groups that occur in the San Francisco area (haplotype groups 3, 4, 7, 19, 20, 21, 33, 35, 37, 39).

The type locality of *B. attenuatus* is “the vicinity of the Bay of San Francisco” (Eschscholtz, 1833). A holotype was not designated

in the original description. Unless morphological differences between the various species can be matched with Eschscholtz’s description or members of the original series, a neotype should be selected for one of the species from the San Francisco Bay area in order to designate which species should retain the name *Batrachoseps attenuatus*.

There appears to be little or no geographical separation between most of the adjacent pairs of species at the present time; their ranges seem to be mostly parapatric with neighboring species. Exceptions are the species isolated from each other in the San Francisco Bay area by the bay, and the two probable species (haplotype groups 1, 2) that may be isolated from all the others in the Sierra Nevada Mountains (Fig. 5). Thus analysis of contact zones should not be difficult for many present contacts.

Speciation within the Northern Species Complex has produced an estimated nine haplotype groups whose divergence occurred primarily in three different periods, about 2–3, 5, and 6 mya (Fig. 9A). The Bodega Species Complex has only one known species

with within-species distances ranging from 0 to 11 bp (Fig. 9B). Speciation within the Eastern Species Complex has produced an estimated eight haplotype groups that diverged between 2.5 and 3 mya (Fig. 9C). Speciation within the Southern (North) Species Complex has produced an estimated six haplotype groups that diverged mainly in two different periods, about 2 and 3–5 mya (Fig. 9D). Speciation within the Southern (South) Species Complex has produced an estimated 15 haplotype groups that diverged in at least three different periods, between 2 and 6 mya (Fig. 9E).

Highton (1995) suggested that there were only five clades of eastern *Plethodon* in the Miocene Epoch of eastern North America that have survived to the present, and that speciation during the late Miocene and Pliocene Epochs produced the large number of species (45) that exist there today. Allozyme analysis of 14 species of the *P. glutinosus* group from the southern Appalachian Mountains provided pairwise comparisons of 91 mean genetic distances for 14 species of eastern *Plethodon* between 34 and 38 degrees north latitude (Highton and Peabody, 2000: Table 2). The range of their genetic distances (0.13–0.53) provides an estimate that speciation events in this group occurred between 1.8 and 7.4 mya. The range of the approximately 39 species of the *B. attenuatus* complex is largely in mountainous areas between 37 and 42 degrees north latitude in Oregon and California, and the speciation events in this group are estimated above to have occurred between 1.8 and 9.6 mya. Interactions of climate with local topography within the ranges of both groups may have provided isolated refuges of suitable habitat surrounded by unsuitable grasslands. A major difference between these two plethodontid groups is that in the east, there are 23 known cases of wide sympatry between the species of the *P. glutinosus* group, while in the *B. attenuatus* species group there are none. This might be due to fewer ecological niches available in the west compared to the eastern United States.

Most modern ecological, behavioral, systematic, and evolutionary studies require basic background taxonomic information: the number of species included. Yet some studies, such as the one by Martínez-Solano et al. (2007), have failed to recognize many species. When attempting to investigate important evolutionary questions, such as whether the ranges of species are contracting, stable, or expanding, a good knowledge of the taxonomy is required. Martínez-Solano et al. (2007) used a method (F_s Test of Fu, 1997), to test for expansion or contraction of geographic ranges. This method was not designed for testing range expansion or contraction simultaneously in a data set containing multiple species. Martínez-Solano et al. (2007) indicated that the results were “surprising” in that they did not detect recent expansion of ranges, which, as they noted, are likely to have recently occurred in northern populations of the *B. attenuatus* species group. The large ranges of northernmost populations of the three northernmost haplotype groups (1, 8, 10 in Figs. 4 and 5) suggest that their ranges have recently expanded northward. The genetic similarity of the most northerly populations within several eastern species of *Plethodon* in the northern portion of the range of the genus has been observed in *P. cinereus* (Highton and Webster, 1976), *P. glutinosus*, *P. cylindraceus*, *P. chlorobryonis*, *P. albagula* (Highton, 1989), and *P. hoffmani* (Highton, 1999, 2009). This is probably the result of post-Pleistocene northern expansion of each species from a limited source area in the northernmost part of each species' range at the height of the Wisconsin glacial period. The same pattern appears in two of the three most northern species of the *B. attenuatus* complex. These are populations 1–6 of haplotype group 10 (all have the same haplotype), see Figs. 1 and 4, and the northern populations 37–49 of haplotype group 8 (number of bp differences 0–4, mean = 1.8), see Figs. 1 and 5. They have the largest geographic ranges of all the haplotype groups, and within at least two of the three groups, northern populations are very similar to each other.

5. Conclusions

Martínez-Solano et al. (2007) provided a valuable data set on sequence variation in the mitochondrial DNA cytochrome-b gene of 178 individuals in 123 populations for the highly variable salamander *Batrachoseps attenuatus*. A reanalysis of their data indicates that there are probably at least 39 species within the complex. Speciation in the common ancestor during the late Miocene Epoch separated the current five species complexes. A second burst of speciation within four of these complexes occurred in the late Pliocene and early Pleistocene Epochs, producing the estimated 39 known living species.

Acknowledgments

I am grateful to E. Jockusch and I. Martínez-Solano for sending me their data and other helpful information. Sincere thanks go to C. Cornblatt and A. Rizk for help with the computer analysis, and to Addison Wynn for preparing the four maps. Allan Larson, Shawn R. Kuchta, and Stephen G. Tilley kindly provided very helpful comments on the manuscript, although they may not agree with some of my conclusions.

References

- Anderson, P.W., 1960. Ecology and evolution in island populations of salamanders in the San Francisco Bay region. *Ecological Monographs* 30, 359–385.
- Baker, R.J., Bradley, R.D., 2006. Speciation in mammals and the genetic species concept. *Journal of Mammalogy* 87, 643–662.
- Baverstock, P.R., Watts, C.H.S., Cole, S.R., 1977. Electrophoretic comparisons between allopatric populations of five Australian pseudomyine rodents (Muridae). *Australian Journal of Biological Science* 30, 471–485.
- Eschscholtz, F., 1833. *Zoologischer Atlas* 5, 1–14.
- Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17, 368–376.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Fu, Y.X., 1997. Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.
- Good, D.A., Wake, D.B., 1992. Geographic variation and speciation in the Torrent Salamanders of the genus *Rhyacotriton* (Caudata: Rhyacotritonidae). University of California Publications in Zoology 126, 1–91.
- Highton, R., 1972. Distributional interactions among eastern North American salamanders of the genus *Plethodon*. In: Holt, P.C. (Ed.), *The Distributional History of the Biota of the Southern Appalachians*. Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA, pp. 139–188.
- Highton, R., 1989. Biochemical evolution in the slimy salamanders of the *Plethodon glutinosus* complex in the eastern United States. Part I. Geographic protein variation. *Illinois Biological Monographs* 57, 1–78.
- Highton, R., 1995. Speciation in eastern North American salamanders of the genus *Plethodon*. *Annual Review of Ecology and Systematics* 25, 579–600.
- Highton, R., 1997. Geographic protein variation and speciation in the *Plethodon dorsalis* complex. *Herpetologica* 53, 345–356.
- Highton, R., 1998. Is *Ensatina eschscholtzii* a ring species? *Herpetologica* 54, 254–278.
- Highton, R., 1999. Geographic protein variation and speciation in the salamanders of the *Plethodon cinereus* group with the description of two new species. *Herpetologica* 55, 43–90.
- Highton, R., 2000. Detecting cryptic species using allozyme data. In: Bruce, R.C., Jaeger, R.G., Houck, L.D. (Eds.), *The Biology of Plethodontid Salamanders*. Kluwer Academic/Plenum Publishers, New York, New York, USA, pp. 215–241.
- Highton, R., 2009. Microgeographic and morphological variation in the woodland salamanders *Plethodon hoffmani* and *Plethodon virginia*, and hybridization between the two species. In: Roble, S.M., Mitchell, J.C. (Eds.), *A Lifetime of Contributions to Myriapodology and the Natural History of Virginia: A Festschrift in Honor of Richard Hoffman's 80th Birthday*. Virginia Museum of Natural History Special Publication No. 16. Martinsville, Virginia, U.S.A., pp. 59–100.
- Highton, R., Peabody, R.B., 2000. Geographic protein variation and speciation in salamanders of the *Plethodon jordani* and *Plethodon glutinosus* complexes in the southern Appalachian Mountains with the description of four new species. In: Bruce, R.C., Jaeger, R.G., Houck, L.D. (Eds.), *The Biology of Plethodontid Salamanders*. Kluwer Academic/Plenum Publishers, New York, New York, USA, pp. 31–93.
- Highton, R., Webster, T.P., 1976. Geographic protein variation and divergence in populations of the salamander *Plethodon cinereus*. *Evolution* 30, 35–45.
- Jockusch, E.L., Martínez-Solano, I., Hansen, R.W., Wake, D.B., 2012. Morphological and molecular diversification of Slender Salamanders (Caudata:

- Plethodontidae: *Batrachoseps*) in the southern Sierra Nevada of California with descriptions of two new species. *Zootaxa* 3190, 1–30.
- Jockusch, E.L., Wake, D.B., 2002. Falling apart and merging: diversification of Slender Salamanders (Plethodontidae: *Batrachoseps*) in the American West. *Biological Journal of the Linnean Society* 76, 361–391.
- Jockusch, E.L., Wake, D.B., Yanev, K.P., 1998. New species of Slender Salamanders, *Batrachoseps* (Amphibia: Plethodontidae) from the Sierra Nevada of California. *Contributions in Science, Natural History Museum of Los Angeles County* 472, 1–17.
- Jockusch, E.L., Yanev, K.P., Wake, D.B., 2001. Molecular phylogenetic analysis of Slender Salamanders, genus *Batrachoseps* (Amphibia: Plethodontidae), from central coastal California with descriptions of four new species. *Herpetological Monographs* 15, 54–99.
- Johns, G.C., Avise, J.C., 1998. A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Molecular Biology and Evolution* 15, 1481–1490.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27, 209–220.
- Marshall, J.C.E., Arévalo, E., Benavides, E., Sites Jr., J.W., 2006. Delimiting species: comparing methods for Mendelian characters using lizards of the *Sceloporus grammicus* (Squamata: Phrynosomatidae) complex. *Evolution* 60, 1050–1065.
- Martínez-Solano, I., Jockusch, E.L., Wake, D.B., 2007. Extreme population subdivision throughout a continuous range: phylogeography of *Batrachoseps attenuatus* (Caudata: Plethodontidae) in western North America. *Molecular Ecology* 16, 4335–4355.
- Martínez-Solano, I., Lawson, R., 2009. Escape to Alcatraz: evolutionary history of Slender Salamanders (*Batrachoseps*) on the islands of San Francisco Bay. *BMC Evolutionary Biology* 9 (38), 1–14.
- Martínez-Solano, I., Peralta-García, A., Jockusch, E.L., Wake, D.B., Vázquez-Domínguez, E., Parra-Olea, G., 2012. Molecular systematics of *Batrachoseps* (Caudata, Plethodontidae) in southern California and Baja California: mitochondrial-nuclear DNA discordance and the evolutionary history of *B. major*. *Molecular Phylogenetics and Evolution* 63, 131–149.
- Maxson, L.R., Maxson, R.D., 1979. Comparative albumin and biochemical evolution in plethodontid salamanders. *Evolution* 33, 1057–1062.
- Moritz, C., Schneider, C.J., Wake, D.B., 1992. Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. *Systematic Biology* 41, 273–291.
- Nei, M., 1972. Genetic distance between populations. *American Naturalist* 106, 283–292.
- Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* 21, 1864–1867.
- Rubinoff, D., Cameron, S., Will, K., 2006. A genomic perspective on the shortcomings of mitochondrial DNA for “barcoding” identification. *Journal of Heredity* 97, 581–594.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4, 406–425.
- Schmidt, K.P., 1953. *A Check List of North American Amphibians and Reptiles*. University of Chicago Press, 280 p.
- Sites Jr., J.W., Marshall, J.C., 2004. Operational criteria for delimiting species. *Annual Review of Ecology and Evolution* 35, 199–227.
- Slatkin, M., Hudson, R.R., 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–562.
- Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary genetic analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24, 1596–1599.
- Thorpe, J.P., 1982. The molecular clock hypothesis: biochemical evolution, genetic differentiation and systematics. *Annual Review of Ecology and Systematics* 13, 139–168.
- Wake, D.B., 1996. A new species of *Batrachoseps* (Amphibia, Plethodontidae) from the San Gabriel Mountains, southern California. *Natural History Museum of Los Angeles County Contributions in Science* 463, 1–12.
- Wake, D.B., Jockusch, E.L., 2000. Detecting species borders using diverse data sets. In: Bruce, R.C., Jaeger, R.G., Houck, L.D. (Eds.), *The Biology of Plethodontid Salamanders*. Kluwer Academic/Plenum Publishers, New York, New York, USA, pp. 95–119.
- Yanev, K.P., 1978. *Evolutionary Studies of the Plethodontid Salamander Genus *Batrachoseps**. Ph.D. Dissertation, University of California, Berkeley, California USA.