Molecular Ecology (1998) **7**, 381–397

Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history

ALAN R. TEMPLETON
Department of Biology, Washington University, St Louis, Missouri 63130-4899, USA

Abstract

Since the 1920s, population geneticists have had measures that describe how genetic variation is distributed spatially within a species' geographical range. Modern genetic survey techniques frequently yield information on the evolutionary relationships among the alleles or haplotypes as well as information on allele frequencies and their spatial distributions. This evolutionary information is often expressed in the form of an estimated haplotype or allele tree. Traditional statistics of population structure, such as *F* statistics, do not make use of evolutionary genealogical information, so it is necessary to develop new statistical estimators and tests that explicitly incorporate information from the haplotype tree. One such technique is to use the haplotype tree to define a nested series of branches (clades), thereby allowing an evolutionary nested analysis of the spatial distribution of genetic variation. Such a nested analysis can be performed regarding the geographical sampling locations either as categorical or continuous variables (i.e. some measure of spatial distance). It is shown that such nested phylogeographical analyses have more power to detect geographical associations than traditional, nonhistorical analyses and, as a consequence, allow a broader range of gene-flow parameters to be estimated in a precise fashion. More importantly, such nested analyses can discriminate between phylogeographical associations due to recurrent but restricted gene flow vs. historical events operating at the population level (e.g. past fragmentation, colonization, or range expansion events). Restricted gene flow and historical events can be intertwined, and the cladistic analyses can reconstruct their temporal juxtapositions, thereby yielding great insight into both the evolutionary history and population structure of the species. Examples are given that illustrate these properties, concentrating on the detection of range expansion events.

Keywords: coalescence, colonization, fragmentation, gene flow, haplotype tree, population structure, range expansion

Introduction

Starting with Wahlund (1928), population geneticists have realized that genetic survey data can reveal information about population subdivision. Wright (1931, 1943) introduced *F* statistics as a way of utilizing allele frequency data gathered in different geographical locations to quantify population subdivision and estimate the amount of gene flow. However, modern genetic surveys using restriction site or DNA sequence data also provide information on the evolutionary relationships of the genetic variation being scored, which is often portrayed as an allele or haplotype tree. Consequently, information is now available about the alleles' existence through evolutionary time as well as geographical space. Can this new temporal information be used to shed more light upon the spatial distribution of current allelic variation? The purpose of this article is to answer this question by worked examples that make use of the evolutionary time dimension provided by haplotype trees.

Traditional *F* statistics do not use temporal information on allelic variation, but several new statistical procedures can make use of haplotype trees (Hudson *et al*. 1992; Slatkin 1989, 1993; Slatkin & Maddison 1989, 1990; Templeton 1993; Templeton *et al*. 1995; Templeton & Georgiadis 1996). This article provides worked examples demonstrating that these new approaches have enhanced

Correspondence: Tel.: +1-314-935-6868; Fax: +1-314-935-4432; E-mail: temple_a@biodec.wustl.edu

power over F statistics, allow greater precision of gene-flow estimation, and can separate population structure (recurrent forces such as gene flow) from historical events (fragmentation and range expansion events). Because the detection of range expansion events has proven to be particularly controversial (Templeton 1993, 1994, 1996a; Ayala 1995), the validity of the criteria used to detect range expansions given in Templeton *et al*. (1995) is examined by applying their methodology to several data sets for which strong a priori evidence exists of range expansions. Finally, it is shown that the method of Templeton *et al*. (1995) does not merely identify and geographically localize the various factors influencing the spatial distribution of genetic variation but it also estimates the dynamic structure and temporal juxtaposition of these evolutionary factors.

Detecting and estimating restricted gene flow

Wright (1931) showed that there is a nonlinear relationship between the amount of gene flow and the degree of genetic differentiation among subpopulations as measured by F statistics. For example, under the island model in which a population is subdivided into a large number of local populations of size N with a proportion m of each population dispersing at random over all local populations, the expected F_{ST} value (the ratio of the observed variance of allele frequencies across local populations to the theoretical maximum variance) is:

$$F_{ST} = \frac{1}{4Nm + 1} \quad (1)$$

From eqn 1, F_{ST} quickly approaches 0 as Nm (the effective number of migrants) increases. This means that once Nm exceeds a value of 4 or 5, there is little effect on F_{ST} values even with large changes in Nm. Worse, if the estimated F_{ST} is not significantly different from 0, the nonlinearity from eqn 1 ensures that not even the order of magnitude of Nm could be estimated. As an example, consider a restriction site genetic survey of the alcohol dehydrogenase locus (Adh) of 39 lines of *Drosophila melanogaster* sampled at four localities in the eastern half of the USA (Aquadro *et al*. 1986). The algorithm of Davis *et al*. (1990) was used to estimate F_{ST} as 0.030, which was not significantly different from zero. The only statistical conclusion one could make about Nm is that it was greater than 2 with a probability of 95%.

Note that in the Adh example, Nm could be 2, 20, 200, or 200 000 000, and there is no way the traditional F_{ST} measurement can distinguish among these alternatives. However, a haplotype tree can be estimated from the restriction site variation found at the Adh locus (Aquadro *et al*. 1986), and this haplotype tree can be converted into a nested series of clades (branches) by using the nesting rules given in Templeton *et al*. (1987) and Templeton & Sing (1993). Basically, these nesting rules start at the tips of the haplotype network and move one mutational step into the interior, uniting all haplotypes that are connecting by this procedure into a '1-step clade.' After pruning off the initial 1-step clades from the tips, this procedure is then repeated on the more interior portions of the haplotype network if needed until all haplotypes have been placed into 1-step clades. The next level of nesting uses the 1-step clades as its units, rather than individual haplotypes. The nesting rules are the same, but result in '2-step clades'. This nesting procedure is repeated until a nesting level is reached such that the next higher nesting level would result in only a single category spanning the entire original haplotype network. The resulting nested clades are designated by '*C*–*N*' where '*C*' is the nesting level of the clade and '*N*' is the number of a particular clade at a given nesting level. Some special nesting rules are needed to deal with symmetries and ambiguities in the estimated haplotype network (Templeton & Sing 1993).

The resulting nested set of clades for the Adh haplotype tree is shown in Fig. 1, along with the geographical distributions of the various haplotypes found in the survey. These nested series of branches constitute an evolutionary based statistical design that was originally used for investigating the relationship between genotype and phenotype (Adh activity in this case; Templeton *et al*. 1987). This nested design can also be used to look for geographical associations in two ways (Templeton *et al*. 1995). Only the more simple of the two approaches is considered in this section. The simple procedure is a nested contingency analysis (Templeton & Sing 1993) in which each geographical location is regarded as a categorical variable. Because geographical distance is ignored, this approach is somewhat of an analogue to the island model. To see how the nested contigency analysis is implemented, consider nesting clade 2–1. Clade 2–1 contains two nested clades within it: 1–1 and 1–2 (Fig. 1). Clade 1–1 consists of one haplotype (haplotype 1, Fig. 1) found in only one line sampled from Kansas and a second haplotype (haplotype 2, Fig. 1) found in two lines collected in Wisconsin. Clade 1–2 includes six lines from Rhode Island and one from Wisconsin. An exact 2×3 permutational contingency test (2 clades vs. 3 geographical locations) is performed to test the null hypothesis of no association of clades with geographical location. In this case the null hypothesis is rejected with an exact probability level of 0.033. Similarly, the nested contingency analysis of nesting clade 3–1 (Fig. 1) which contains clades 2–1, 2–2, and 2–3 is significant with a null probability of 0.028. These contingency tests are repeated on all nesting clades containing more than one nested clade and that were found at more than one location. No other nested contingency tests were significant at the 5% level. Hence, a geographical association is detectable when the haplotype

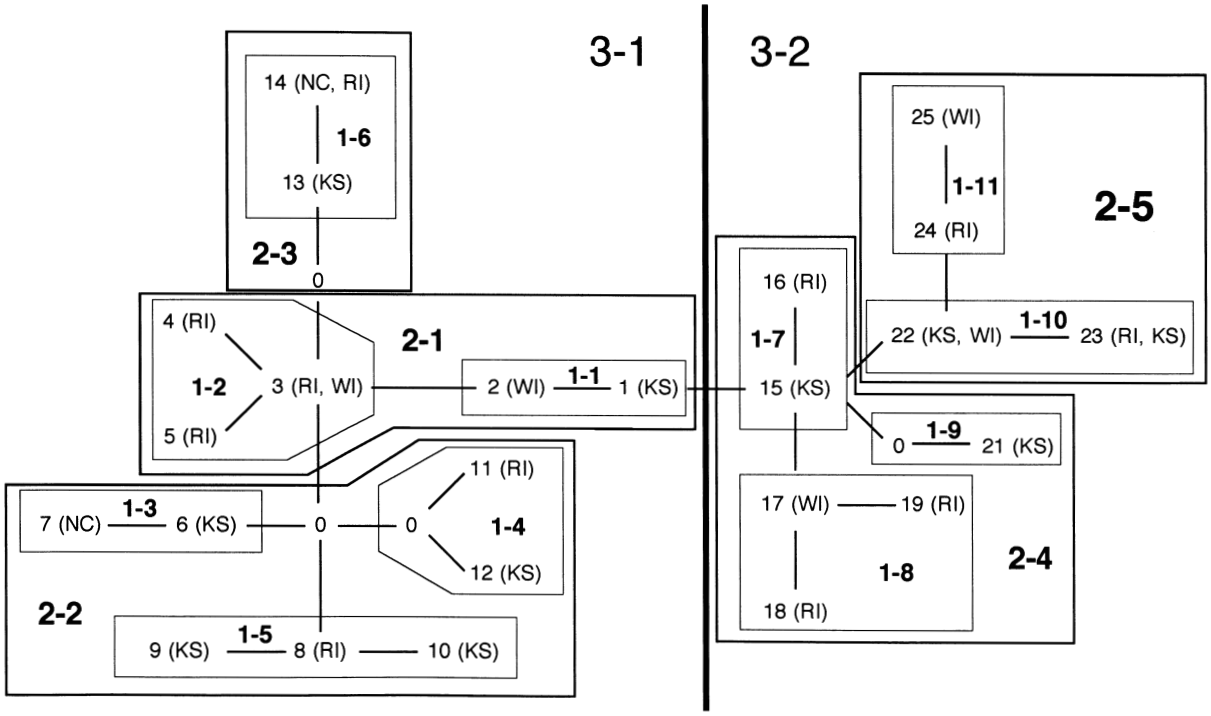

Fig. 1 The haplotype network for haplotypes at the alcohol dehyrogenase locus of *Drosophila melanogaster* from Aquadro *et al*. (1986) with the nesting design of Templeton *et al*. (1987). Each line in the network represents a single mutational change. 0 indicates an interior node in the network that was not present in the sample; that is, these are inferred intermediate haplotypes between two nearest neighbour haplotypes in the network that differed by two or more mutations. Haplotype numbers are those given in Templeton *et al*. (1987), although haplotype 20 in that study is excluded in this analysis. Haplotype 20 came from a single line from Japan. All other haplotypes came from the eastern USA, and the states in which they were collected are indicated after the haplotype number. Thin-lined polygons indicate the haplotypes grouped together into 1-step clades, medium-lined polygons indicate the 1-step clades nested together into 2-step clades, and the thick line in the middle indicates the 2-step clades nested together into 3-step clades.

tree is used to generate a nested statistical design even though no geographical associations are detectable at all with an F_{ST} statistic.

Given that a significant geographical association has been detected, and assuming for the moment that the association is due to restricted gene flow, it is now meaningful to estimate *Nm*, the effective number of migrants per generation among the geographical locations. An estimation procedure that makes use of the gene tree is given by Slatkin & Maddison (1989). Their procedure does not test whether there is a statistically significant geographical association. Consequently, the Slatkin & Maddison (1989) procedure should only be used when a significant geographical association has been detected; otherwise, the resulting estimator is of dubious statistical validity (Templeton *et al*. 1995). Moreover, the Slatkin & Maddison (1989) estimation procedure should only be used when the cause of the geographical association is inferred to be due to restricted gene flow; otherwise, the resulting estimator is biologically misleading. More on this point will be given in the next section. For now the assumption is made that the significant associations detected by the nested clade analyses are due to restricted gene flow. Applying the Slatkin & Maddison (1989) algorithm to the tree given in Fig. 1 yields an estimate of *Nm* of 5.4, with a 95% confidence interval of 2.0–19.6. Recall that the insensitivity of the traditional F_{ST} statistic to large values of *Nm* precluded any inference on the possible upper bound of the *Nm* value. Hence, the phylogenetic approach has resulted in far greater precision in estimating *Nm* than is possible from a traditional F_{ST} analysis.

Any estimate of *Nm* based upon a single locus should be regarded as preliminary because much evolutionary stochasticity is associated with the variation at any given locus (Ewens 1983), and because locus-specific forces (such as selection) can distort the apparent F_{ST} or *Nm* value (Lewontin & Krakauer 1973). Langley *et al*. (1988) studied the same lines of *D. melanogaster* analysed above for a different locus, the duplicated amylase (Amy) locus. As with Adh, there is no significant F_{ST} for the Amy locus. The Amy region has been subject to a nested contingency analysis for isozyme associations (Templeton & Sing 1993), and the same nested design is now used for geographical associations. Unlike the Adh locus, however,

the duplicated Amy locus shows much internal recombination. This necessitated subdividing the Amy region into subregions that had much recombination between, but little or no recombination within (Templeton & Sing 1993). Hudson *et al*. (1992) have shown that in such cases it is better to estimate *Nm* for each recombinationally separated block, so the geographical contingency analysis is applied to the left and right subregions separately (the nested designs are given in Figs 3 and 5, respectively, in Templeton & Sing 1993). Unfortunately, by subdividing this DNA region, both of the resulting subregions have few distinct haplotypes and result in low-resolution haplotype trees. The nested contingency analyses in both subregions failed to detect any significant geographical associations. The different (but compatible) results obtained with the Adh vs. the Amy loci may be due to either the variation across loci mentioned above or an erosion of statistical power caused by the low-resolution haplotype trees found at the Amy locus vs. the much higher resolution tree for Adh (Fig. 1). The roles of tree resolution and interlocus heterogeneity in determining statistical power in these analyses obviously needs to be explored further in future studies.

Discriminating between recurrent gene flow and historical events

The major limitation of the above analysis was that it assumed that the significant association between haplotype variation and geographical location was due to restricted gene flow. However, suppose a species has been fragmented into two or more subpopulations that experience no gene flow at all. If they had a recent shared ancestry, the populations could still display some genetic similarity that would yield a traditional estimate of $F_{ST} < 1$, thereby erroneously implying nonzero gene flow. Alternatively, suppose the species recently expanded its range over a large area from some smaller subpopulation within the ancestral range. Then there would be much genetic similarity over this expanded range, leading to an overestimate of gene flow. Using *F* statistics or an algorithm that assumes that all geographical associations are due to gene flow (e.g. Slatkin & Maddison 1989) can therefore yield an estimator of *Nm* that is biologically misleading.

Fortunately, this potential confoundment of population structure with population history can be investigated by using haplotype trees. Indeed, the primary advantage of using the haplotype tree information is not the quantitative advantage of enhanced power and precision; rather, it is the qualitative advantage of discriminating among various biological explanations for any detected geographical association. To show that haplotype trees can discriminate among cases that appear identical to the nonhistorical *F*-statistic analyses, consider the study of Templeton & Georgiadis (1996) on mitochondrial DNA (mtDNA) restriction site variation in Eastern African populations of buffalo (*Syncerus caffeer*) and impala (*Aepyceros melampus*). The *F*-statistic estimator of Davis *et al*. (1990) yields an F_{ST} of 0.08 for the buffalo and 0.10 for the impala. Both of these F_{ST} values are significantly different from zero, but they are not significantly different from each other. Moreover, in both species most of the geographical sites surveyed are relatively close together in Kenya and Tanzania, but one site (Chobe) is far to the south. In both species, the Chobe samples had many haplotypes not found in the other locations, and it was the Chobe samples that were primarily responsible for the significant F_{ST} values in both cases. Hence, this *F*-statistic analysis implies that both species are equally subdivided, have comparable rates of gene flow, and display restricted gene flow primarily between the Chobe vs. Kenya/Tanzania localities.

Templeton & Georgiadis (1996) also estimated haplotype trees for the mtDNA, as shown here in Fig. 2. Figure 2 also indicates the haplotypes found only in Chobe in both species. Even a cursory glance at Fig. 2 reveals that the pattern of distribution of the Chobe-only haplotypes in the haplotype trees are completely different in these two species which appear indistinguishable by *F*-statistic analysis. In the buffalo, the Chobe haplotypes are scattered throughout the haplotype network; in the impala, the Chobe haplotypes are tightly clustered in the haplotype network. Although both species show the same degree of spatial subdivision as measured by F_{ST}, they have obviously achieved this degree of subdivision in very different fashions through time. Clearly, the use of haplotype networks allows a finer discrimination of biological pattern than an *F*-statistic analysis. The reason for this is straightforward; by using a haplotype network, one is examining a spatial/temporal pattern of genetic variation whereas with the *F*-statistic and other nonhistorical analysis one can only examine the current spatial pattern.

The scattered spatial/temporal pattern found in the buffalo (Fig. 2) indicates recurrent genetic interchange between Chobe and the more northerly populations throughout the time period from the coalescence of mtDNA to the present. The impala pattern is more difficult to interpret. Such a strong evolutionary clustering of haplotypes in a geographical region, particularly when the haplotype clusters are separated by a long branch length with missing intermediates, is often interpreted as evidence of a past fragmentation event (see Avise 1994 and references therein for examples). However, because impala are found in intermediate geographical locations that were not sampled, it is possible that this pattern arose from isolation-by-distance (i.e. a restricted gene flow model rather than an historical event) such that geographically intermediate populations would fill in the

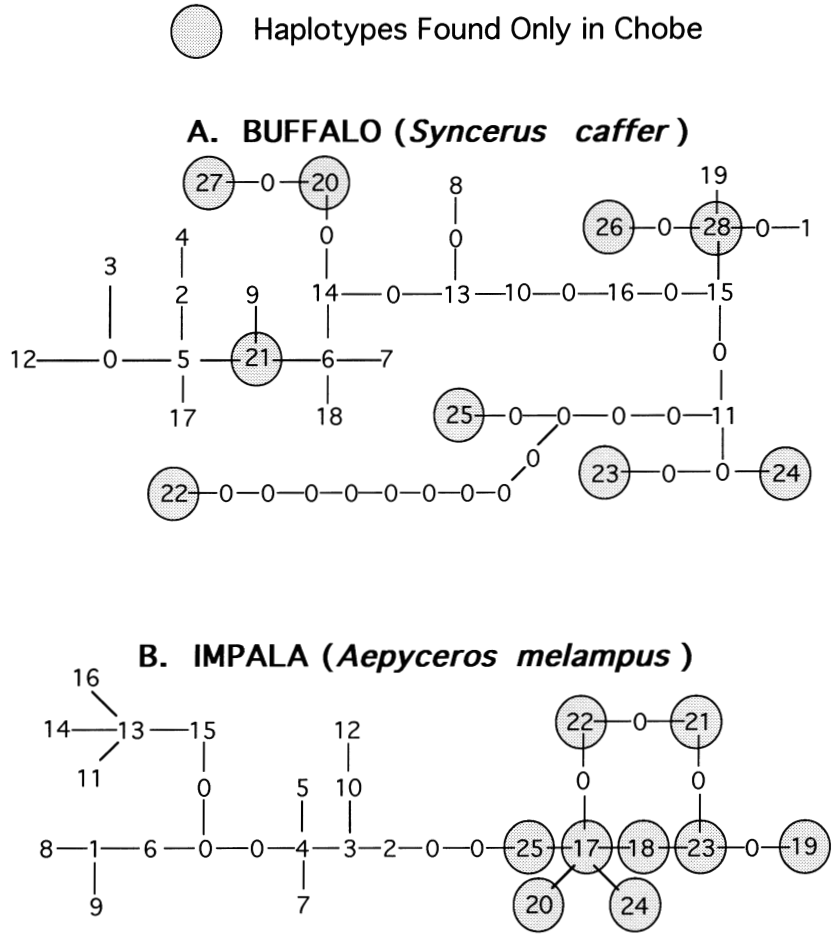

Fig. 2 The haplotype networks for mitochondrial DNA as estimated by Templeton & Georgiadis (1996) from two species of African bovids: A, buffalo (*Syncerus caffer*); and B, impala (*Aepyceros melampus*). Each line in the network represents a single mutational change. 0 indicates an interior node in the network that was not present in the sample; that is, these are inferred intermediate haplotypes between two nearest neighbour haplotypes in the network that differed by two or more mutations. Haplotype numbers are those given in Templeton & Georgiadis (1996). The haplotypes in each species that are found only in the Chobe sample location are indicated by stipled circles.

missing haplotype nodes and show a gradual shift from one cluster of haplotypes to the other. Indeed, a rigorous quantitative analysis of these data (of the type to be discussed below) reveals that the sparseness of sampling prevents one from distinguishing between isolation-by-distance vs. fragmentation of the Chobe population from the Kenyan/Tanzanian populations (Templeton & Georgiadis 1996).

The analysis of Templeton & Georgiadis (1996) on the impala illustrates the dangers of making biological inferences simply by a visual inspection of how geography overlays upon a haplotype tree. Such visual inferences are commonplace in the phylogeographic literature (Avise 1994 and references therein), but they make no assessment of adequate sample sizes for statistical significance nor adequate sampling of geographical locations for distinguishing among potential causes of geographical associations. What is needed is an objective statistical analysis that first rejects the null hypothesis of no association between haplotype variation and geography, and then interprets the statistically significant patterns using explicit criteria that include an assessment of sampling adequacy. To accomplish this task, Templeton *et al*. (1995) have proposed a quantitative analysis of geographical data using the same nested design generated by the haplotype network that was used in the contingency analyses. As this technique and its inference criteria are discussed at length along with a detailed worked example in Templeton *et al*. (1995), only a brief summary will be given here.

The geographical data are quantified in two main fashions: the clade distance, D_c, which measures the geographical range of a particular clade; and the nested clade distance, D_n, which measures how a particular clade is geographically distributed relative to its closest evolutionary sister clades (i.e. clades in the same higher-level nesting category). In particular, the clade distance measures the average distance that an individual bearing a haplotype from the clade of interest lies from the geographical centre of all individuals bearing haplotypes from the same clade. The nested clade distance measures the average distance that an individual bearing a haplotype from the clade of interest lies from the geographical centre of all individuals bearing haplotypes from the next higher-level nesting clade that contains the clade of interest. Contrasts

in these distance measures between tip clades (clades that are not interior nodes in the haplotype tree) and the clades immediately interior to them in the cladogram are important in discriminating the potential causes of geographical structuring of the genetic variation (Templeton *et al*. 1995), as will be discussed later. The statistical significance of the different distance measures and the interior-tip contrasts are determined by random permutation testing which simulates the null hypothesis of a random geographical distribution for all clades within a nesting category given the marginal clade frequencies and sample sizes per locality.

If statistically significant patterns are detected, they then need to be interpreted biologically. Templeton *et al*. (1995) consider three major biological factors that can cause a significant spatial/temporal association of haplotype variation. The first factor is restricted gene flow, particularly gene flow restricted by isolation-by-distance (Wright 1943). Because restricted gene flow implies only limited movement by individuals during any given generation, it takes time for a newly arisen haplotype to spread geographically. Obviously, when a mutation first occurs, the resulting new haplotype is found only in its area of origin. With each passing generation, a haplotype lineage that persists has a greater chance of spreading to additional locations via restricted gene flow. Hence, the clade distances should increase with time under a model of restricted gene flow. If an outgroup can be successfully used to root the haplotype tree, any series of nested clades can be polarized temporally in an unambiguous fashion. However, often intraspecific haplotype trees cannot be rooted reliably by the outgroup method or other standard rooting procedures (Templeton 1993; Castelloe & Templeton 1994). Fortunately, in a nested series of clades, a nesting clade has to be as old or older than all the lower level clades nested within it. Hence, as nesting level increases, there is a nondecreasing age series even when the root is not known. Accordingly, the clade distances are expected to increase with increasing nesting level. This expected increase will continue until either the highest nesting level is reached or, if the gene flow is sufficiently high relative to the coalescent time of the haplotype tree, a nesting level will be reached in which the clades are uniformly distributed over the entire sampled geographical range, and all higher nesting levels will replicate that pattern. Another aspect of the expected patterns under restricted gene flow is that when a mutation occurs to create a new haplotype, that new haplotype obviously resides initially within the range of its ancestral haplotype. As the ancestral haplotype is older than its mutational offshoot, it should have a wider geographical distribution. Therefore, when the new haplotype starts spreading via gene flow, it will often remain within the geographical range of its ancestor for many generations, particularly under an isolation-by-distance model. Because there is a strong tendency for the ancestral haplotypes to be immediately interior to the derived haplotypes in terms of the topology of the haplotype network (Castelloe & Templeton 1994), this means that there will be a strong tendency under restricted gene flow for tip clades to have a geographical range smaller and often nested within the range of the clades that are immediately interior to them. Moreover, because the ancestral haplotype is expected to be most frequent near its site of geographical origin, most mutational derivatives of the ancestral haplotype will also occur near the ancestral site of geographical origin. This means that the geographical centres of all the clades nested together should be close; hence, the clade distances and nested clade distances should show similar patterns under restricted gene flow.

A second factor is past fragmentation events. When the nesting level reaches the temporal period at which the fragmentation event occurred, the clade distance cannot increase beyond the geographical ranges of the fragmented subpopulations, but the nested clade distances will generally show a marked increase when the fragmented clades are allopatric, as is typically the case. If the fragmentation event is an old one relative to the rate at which mutations accumulate, the branch lengths between the clades displaying large nested clade distances but plateaued clade distances will tend to be longer than the average branch length in the tree (due to the accumulation of mutations that differentiate the fragmented subpopulations).

Range expansion (including colonization) is the third factor that can create a geographical association with the haplotype network. When range expansion occurs, those haplotypes found in the ancestral population(s) that were the source of the range expansion will become geographically widespread (large clade distances), and the distinction between the relative geographical ranges of tip vs. interior clades expected under restricted gene flow breaks down or can even be reversed. Moreover, some of the haplotypes found in the expanding populations can become quite distant from some of the older haplotypes that are confined to the ancestral, pre-expansion area (large nested clade distances), particularly when long-distance colonization is involved. As mutations first start to accumulate in the colonizing population, they will be tips with large nested clade distances because the interior haplotypes from which they mutated will also be found in the ancestral range.

The procedure of Templeton *et al*. (1995) first limits inference to those clades showing statistically significant geographical associations. Next, the patterns displayed by these significant associations are evaluated relative to the above expectations. In order to make this pattern evaluation explicit and consistent, an inference key is provided as an appendix to Templeton *et al*. (1995) (reproduced here as

an appendix and hereafter referred to as the inference key). This quantitative geographical nested analysis and inference key should be used before using an estimation algorithm such as that of Slatkin & Maddison (1989) in order to ensure that the biological situation being examined corresponds to the assumptions of the estimation algorithm. Subjecting the *D. melanogaster* Adh data to such an analysis reveals a significantly small tip clade distance nested within clade 3–1, and application of the inference key led to the conclusion of restricted gene flow. Hence, it is meaningful to estimate *Nm* in this case, and the estimated $Nm = 5.4$ that was given earlier indicates that these eastern USA populations of *D. melanogaster* experience relatively high levels of recurrent gene flow. The two Amy subregions were also subjected to this quantitative distance analysis even though neither subregion displayed significant geographical associations in the categorical contingency analysis. No significant effects were detected in the right subregion, but significant effects were detected within nesting clade 1–1 and among clades 2–1 and 2–2 nested within the total haplotype network (see Fig. 3 in Templeton & Sing 1993). Application of the inference key given in the appendix led to the conclusion of restricted gene flow within 1–1, and an inconclusive outcome at the total network level. Hence, the same statistically significant qualitative conclusions were reached for both loci. However, for the low-resolution Amy haplotype tree, all the interior haplotypes are so widely distributed geographically that it is impossible to use the estimation procedure of Slatkin & Maddison (1989). All that one can conclude from the Amy locus is that the level of recurrent gene flow among these populations must be high even though somewhat restricted by isolation by distance, a conclusion also compatible with the Adh results.

Although only one cause of nonrandom spatial distributions of clades was inferred for the Adh and Amy examples, the nested analysis of Templeton *et al*. (1995) searches for multiple, overlaying patterns within the same data set. For example, in the analysis of mtDNA restriction site variation in the salamander *Ambystoma tigrinum* given in Templeton *et al*. (1995), an historical fragmentation event is inferred between two named subspecies followed by independent range expansion within each subspecies, overlaid upon a pattern of isolation by distance occurring within each subspecies. There is nothing about the evolutionary factors of restricted gene flow, fragmentation events, or range expansion events that make them mutually exclusive alternatives. One of the great strengths of this inference procedure is that it explicitly searches for the combination of factors that best explains the current distribution of genetic variation and does not make a priori assumptions that certain factors should be excluded or be regarded as unlikely. Moreover, by using the temporal polarity inherent in a nested design (or by outgroups when available), the various factors influencing current distributions of genetic variation are reconstructed as a dynamic process through time. For example, the analysis of the *A. tigrinum* mtDNA data reveals that the fragmentation event occurred prior to the expansion events (the expansion events were inferred in clades nested within the clade detecting the fragmentation event, as shown in Table 3 of Templeton *et al*. 1995). Moreover, the inference of restricted gene flow via isolation-by-distance is found in clades that nest and are nested within the clade leading to an inference of an expansion event, thereby implying that isolation-by-distance characterized the salamanders' population structure both before and after the expansion event. Hence, this procedure does not merely identify and geographically localize the various factors influencing the spatial distribution of genetic variation, but rather it brings out the dynamic structure and temporal juxtaposition of these evolutionary factors.

The inference key also incorporates the types of pattern artifacts that can arise from inadequate sampling, thereby leading to no definitive biological inference. The ability of the key to yield an inconclusive outcome is a strength, not a weakness, because the deficiencies of the current sample for making unambiguous biological inference are identified. For example, the application of this TRP key to the buffalo mtDNA data set (Fig. 2A) yields a conclusion of gene flow between Chobe and Tanzania, but the absence of samples between these localities leave it ambiguous as to whether this gene flow is characterized by isolation-by-distance or by occasional but recurrent long-distance dispersal (Templeton & Georgiadis 1996; Table 13.2). For the impala, there is no discrimination between restricted gene flow, fragmentation, or range expansion as possible explanations for the Chobe/Tanzanian pattern because of the absence of samples between Chobe and Tanzania (Table 13.4 in Templeton & Georgiadis 1996). Obviously, in both cases, future sampling should be directed towards the geographical gap between Tanzania and Chobe. Thus, the inference key gives specific and detailed guidance for future sampling activities.

The buffalo/impala example also illustrates the difficulty of knowing what is an adequate sampling design a priori. Suppose that an investigator only wanted to know if the Tanzanian and Chobe populations were interconnected by recurrent gene flow (for example, in designing a conservation program). For the buffalo, the inference of recurrent gene flow has already been established with the samples given in Templeton & Georgiadis (1996), although the details of the nature of the gene flow remain hidden. On the other hand, the impala sample is inconclusive on this issue even though sampling occurred in the same areas and with comparable sample sizes. Each species has a potentially unique population structure and

history that shaped its evolution, so it is difficult to design an optimal sampling scheme when sampling resources are limited. It is better to use only a portion of the sampling resources available to perform an initial analysis, and then use the inference key as a guide in allocating the remaining resources to obtain the most critical samples needed for strong inference.

Validity of the criteria used to infer range expansion

The basic patterns associated with restricted gene flow that were incorporated into the inference key are well justified by recent work in coalescent theory and computer simulations (Hudson *et al*. 1992; Slatkin & Maddison 1990; Slatkin 1991, 1993; Nei & Takahata 1993; Neigel *et al*. 1991; Neigel & Avise 1993; Takahata & Slatkin 1990; Takahata 1991). Similarly, the predictions under fragmentation are straightforward and represent a quantitative rendering of the patterns commonly used to infer fragmentation events (Avise 1994). The least theoretically justified pattern is that associated with range expansion. The range expansion expectations of widespread tip clades with some ancestral haplotypes restricted to the ancestral range were first described by Cann *et al*. (1987). Although these expectations seem reasonable, they have not been confirmed analytically or through extensive computer simulations. Part of the problem is that range expansion can arise in many different situations and can interact with many different patterns of gene flow and/or fragmentation events. Hence, the range of possible assumptions that could be incorporated into an analytical or simulation model is daunting, and it is not clear which assumptions are most biologically realistic.

The best method of insuring biological realism is to examine actual examples of range expansion. Fortunately, there are many cases in which range expansion can be inferred with much certainty without the use of genetic data. Table 1 presents 13 data sets that have strong prior evidence for range expansion, a genetic survey using restriction site mapping or DNA sequencing (in all cases involving mtDNA), and well-documented geographical sampling with spatial frequency information on all haplotypes. The first seven cases involve organisms with current ranges that include areas that were not inhabitable during the Pleistocene; hence, post-Pleistocene range expansion must have occurred. The remaining six cases involve organisms whose ranges have been expanded by human activities or range expansions by humans themselves. These natural examples of range expansion provide an excellent vehicle for validating the expectations of Cann *et al*. (1987) and the inference key.

The most common statistically significant inference in these data sets is restricted gene flow, which is inferred in

Table 1 Summary of the nested geographical analyses of 13 mitochondrial DNA data sets with a priori evidence for range expansion. N_L refers to the number of localities in the study, N_I the total number of individuals in the study, and N_H the number of haplotypes detected in the genetic survey

Organism	N_L	N_I	N_H	*A priori* reason for range expansion	Statistically significant historical events	Reference*
Ambystoma tigrinum tigrinum	13	139	6	Current range includes areas climatically uninhabitable in the Pleistocene	Northwest range expansion into the northern Great Plains	Templeton *et al*. (1995)*
A. t. mavortium	40	381	16	Current range includes areas climatically uninhabitable in the Pleistocene	Northeast contiguous range expansion into the northern Great Plains	Templeton *et al*. (1995)*
Etheostoma blennioides blennioides	11	34	4	Current range includes areas under glaciers in the Pleistocene	Range expansion from southern Ohio drainage rivers to rivers in formerly glaciated areas Fragmentation between populations separated by the Kanawha Falls	Wilson (1997)*
E. b. pholidotum	15	93	5	Current range includes areas under glaciers in the Pleistocene	Range expansion from northern Ozark rivers to rivers in formerly glaciated areas Fragmentation between Ozark rivers draining into the Missouri vs. Mississippi Rivers	Wilson (1997)*

Table 1 *Continued*

Organism	N_L	N_I	N_H	A priori reason for range expansion	Statistically significant historical events	Reference*
Trimerotropis saxatalis	62	613	41	Ozarks climatically uninhabitable prior to the xerothermic maximum	Range expansion into the Ozarks from the southwest followed by fragmentation	Gerber (1994)*
Geomys bursarius	13	159	19	Current range includes areas under glaciers in the Pleistocene	Northwest range expansion from Texas to New Mexico Range expansion from Texas to Minnesota Colonization event east of Mississippi River	Davis (1986)
Galaxias truttaceus	16	211	58	Current range includes lakes created by melting Pleistocene glaciers in central Tasmania	Range expansion to central lakes from coast followed by fragmentation Range expansion to north coast	Ovenden & White (1990
Drosophila melanogaster	18	144	23	Current global distribution due to dispersal by human activities; native to Africa	Both contiguous and long-distance range expansion on a global basis	Hale & Singh (1991)
Drosophila buzzatii	15	283	26	Introduction to Europe from South America by human transport of host plant	No population-level historical events detected	Rossi *et al.* (1996)
Canis latrans	25	327	32	Historical range expansion since 1900	Range expansion from southern and western North America to north and east coast	Lehman & Wayne (1991)
Macaca fascicularis	3	52	17	Introduced to the Island of Mauritius in the 1500s by Portuguese sailors	Colonization event to Mauritius from the Phillipines and/or Indonesia	Lawler *et al.* (1995)
Homo sapiens	14	345	127	Human settlement of remote Pacific Islands	Contiguous range expansion to nearby islands Colonization events to distant islands	Sykes *et al.* (1995)
Homo sapiens	21	532	126	Human settlement of Siberia and the Americas	Contiguous range expansion in Siberia 1 or 2 colonization events in America followed by fragmentation between America and Siberia Contiguous range expansion in the Americas	Torroni *et al.* (1993a,b)

*Reference includes the nested cladistic analysis.

all data sets except the darter *Etheostoma blennioides pholidotum*, the macaque monkey, and the fish *Galaxias truttaceus*. For the darter, the failure to detect restricted gene flow may be due to low genetic resolution (only five haplotypes, Table 1), and in the macaque, too few localities (only three, Table 1). In *G. truttaceus* there is genetic variation but homogeneity across most sampling locations at all clade levels, thereby indicating much gene flow among most of the sampled localities in this species. This inference is consistent with the an isozyme survey (Ovenden & White 1990). In the 10 examples showing restricted gene flow, it is most commonly observed through isolation-by-distance when the sampling was sufficient to discriminate the type of restricted gene flow. Only the two human examples showed some recurrent long-distance exchange, but even in humans gene flow is restricted primarily through isolation-by-distance.

Because the focus in this section is on historical events that influenced the spatial distribution of haplotype variation rather than gene flow, Table 1 only gives the events that were inferred using the inference key. If the geographical sampling is adequate, the inference key discriminates between contiguous range expansion (a gradual, moving front of range expansion) vs. colonization (an abrupt establishment of a population in a new geographical region), and these inferred discriminations are indicated in Table 1. However, in many cases the inference key yields an inference of range expansion, but the sample is inadequate to discriminate the details of the nature of the expansion. These events are simply referred to as 'range expansion' in Table 1.

The first five cases in Table 1 (two subspecies of tiger salamanders, two subspecies of a darter, and the lichen grasshopper) have already been analysed with a nested geographical analysis and the inference key. The remaining eight cases all represent new nested analyses of previously published data. The analysis of the first of these, the gopher *Geomys bursarius* (Davis 1986), is given below, but space limitations preclude giving the details of the other analyses. However, the details of any or all of these analyses are available upon request to the author.

Davis (1986) surveyed mtDNA restriction site variation in the gopher, *G. bursarius* (note, this is part of a complex of gopher taxa whose status is debated: this analysis excludes many of the controversial taxa and restricts the analysis only to the widespread populations of this species in the central part of its range). These gophers are found in a mid-continental belt in North America that extends from southern locations that were never glaciated, such as Texas, to northern locations such as Minnesota that were under Pleistocene ice sheets. Figure 3 represents the mtDNA haplotype network estimated from the data of Davis (1986) using the algorithm of Templeton *et al*. (1992), along with the resulting nested design. In this case, outgroup data can root the tree, and the only effect the outgroup has on the nested design is to designate clade 4–1 as an interior clade at the highest level of nesting. Figure 4 presents the results of the nested analysis of clade and nested clade distances, along with the inferences reached by using the inference key. As can be seen from Fig. 4, there is much evidence for range expansion in this species. The inferences of range expansion found within clades 1–1, 2–1 and 4–1 all involve populations in Texas and New Mexico, indicating a north-west range expansion from Texas. These three clades define a continuous nested series of clades (Fig. 4, note that clade 3–1 is the same as clade 2–1; such 'degeneracy' arises when there are internal nodes in the haplotype tree that are not represented by haplotypes actually present in the sample). This indicates that this range expansion occurred gradually on the time scale marked by mtDNA coalescent events. The range expansion within clade 3–4 is specifically inferred to be due to a colonization event, and involves the only population of this species on the east bank of the Mississippi River (in Illinois), indicating a transfer event of gophers to Illinois from the west. The range expansion detected within clade 4–2 involves populations from Kansas to Minnesota, and this event in turn is nested within a range expansion at the highest clade level involving populations from Texas and New Mexico expanding into the northern states from Kansas through Minnesota. The nesting of these two range expansion events both geographically and within the clade structure of the haplotype tree indicates that this was also a range expansion that occurred gradually on the timescale marked by mtDNA coalescent events. All of these inferences are consistent with prior expectations (Davis 1986).

Discussion

The results given and summarized in this study clearly demonstrate that nested analyses of haplotype trees with geographical data provide greater statistical power and precision than traditional *F*-statistic analysis for detecting genetic/geographical associations. More importantly, the nested haplotype tree approach can reveal extremely different patterns of association even in those cases that appear indistinguishable to nonhistorical analyses. The fact that different spatial/temporal patterns can be detected with the nested analysis opens up the potential for discriminating among various evolutionary causes for associations arising between genetic and spatial variation. Restricted gene flow, fragmentation events, and range expansion events (including both contiguous range expansion and long-distance colonization events) can all create genetic/spatial associations. Of these, the patterns associated with range expansions are the most controversial and

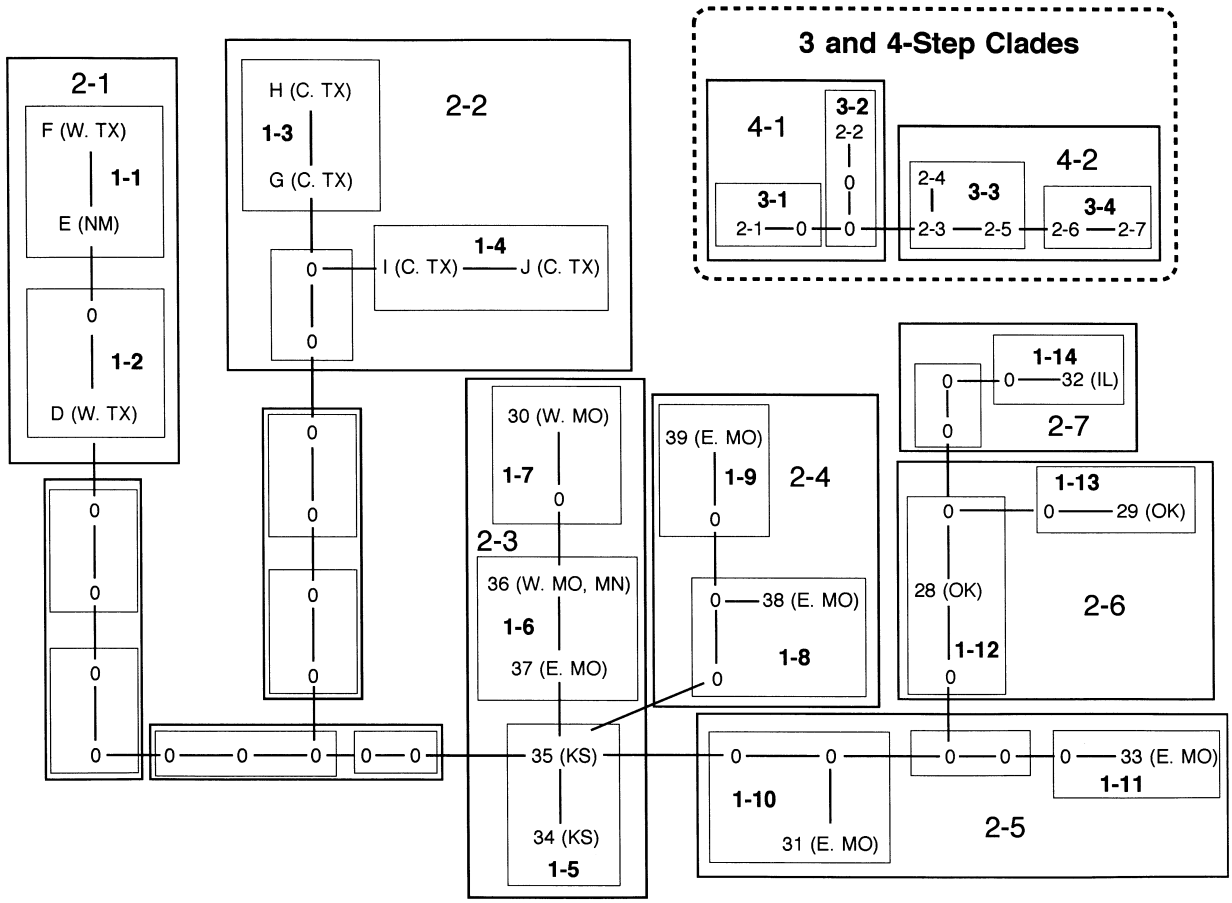

Fig. 3 The haplotype network for mitochondrial DNA haplotypes from the gopher *Geomys bursarius*. Each line in the network represents a single mutational change. 0 indicates an interior node in the network that was not present in the sample; that is, these are inferred intermediate haplotypes between two nearest neighbour haplotypes in the network that differed by two or more mutations. Haplotype numbers are those given in Davis (1986). The states in which they were collected are indicated after the haplotype number. Thin-lined polygons indicate the haplotypes grouped together into 1-step clades, and medium-lined polygons indicate the 1-step clades nested together into 2-step clades. The higher nesting categories are shown in the network of 2-step clades given in the curved-cornered rectangle in the top right of the Figure. In the 2-step clade network, thin-lined polygons indicate the 2-step clades grouped together into 3-step clades, and medium-lined polygons indicate the 3-step clades nested together into 4-step clades.

least studied. This study therefore examined the ability of the nested analysis to infer range expansion events. Accordingly, 13 data sets with strong prior evidence of range expansions were analysed, with the results summarized in Table 1. As can be seen, range expanions were inferred in 12 of the 13 cases.

One potential explanation for this high success rate is that the criteria for range expansion given in the inference key are so broad that range expansion events will be commonly inferred. If this were true, a large number of false positives would be expected as well. Fortunately, this does not seem to be the case. A total of 99 nesting clades had significant deviations from the null hypothesis in the 13 data sets analysed, but only a subset of 35 led to the inference of range expansion. The most common inference was restricted gene flow, and a few fragmentation events were also inferred (Table 1). Moreover, of the 35 nesting clades associated with a significant range expansion pattern, 34 were consistent with prior knowledge (many range expansion events influenced mutiple clades, as shown by the gopher example). Only one inferred range expansion event was not expected a priori, and that is the expansion of *Galaxias truttaceus* from the south-eastern coastal rivers to the north in Tasmania. Even this inferred range expansion event is not necessarily a false positive because a land bridge existed between Tasmania and Australia 10 000–20 000 years ago (Ovenden & White 1990) which may have prevented the south-eastern fish from reaching the northern streams until relatively recently. However, even if this case is regarded as a false positive, the fact that 34 of the 35 clades inferring range expansion were compatible with prior knowledge out of a

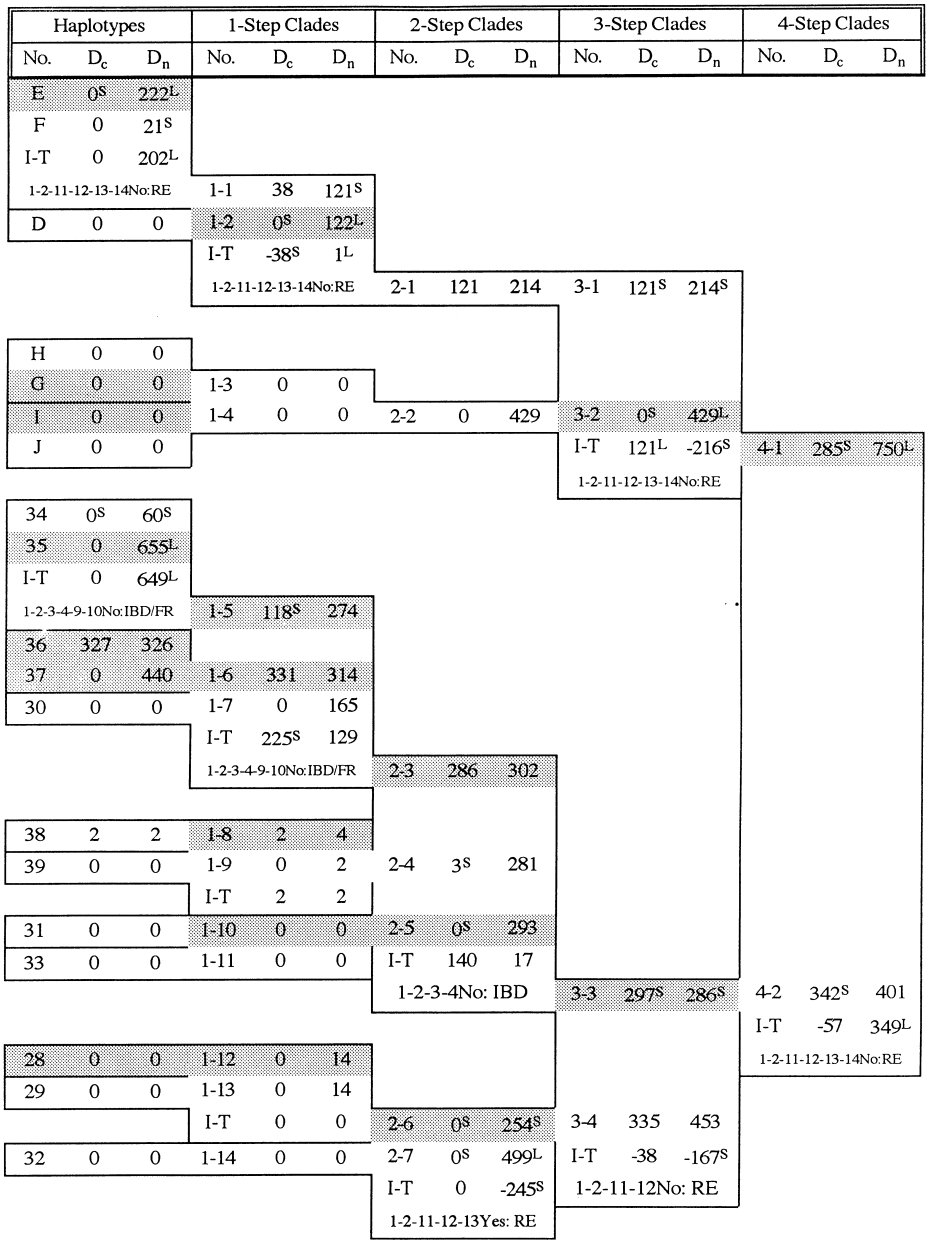

Fig. 4 Results of the nested geographical analysis of the *Geomys bursarius* mtDNA haplotypes. The nested design is given in Fig. 3, as are the haplotype and clade designations. Following the name or number of any given clade are the clade and nested clade distances. Also, in those nesting clades containing both tip and interior nested clades, the average difference between interior vs. tip clades for both distance measures is given in the row labelled I-T. A superscript S means that the distance measure was significantly small at the 5% level, and a superscript L means that the distance measure was significantly large. At the bottom of the boxes that indicate a nested set of clades in which one or more of the distance measures was significantly large or small is a line indicating the biological inference. The numbers refer to the sequence of questions in the key that the pattern generated, followed by the answer to the final question in the inference key. Following this answer is the biological inference generated by use of the inference key, where RE is range expansion, FR is fragmentation, and IBD is recurrent gene flow restricted by isolation by distance. If two or more of these symbols are after the colon, the inference key could not distinquish among the indicated alternatives.

total of 99 significant clades implies that the inference key does not lead to frequent false positives.

This conclusion is reinforced by contrasting these analyses to other nested analyses done on species or sets of populations where there was no prior expectation of range expansion. There was no prior expectation of range expansion in the *Drosophila melanogaster* populations from the eastern half of the USA, and none was inferred for either the Adh or Amy loci even though both loci detected significant geographical associations due to gene flow. There was also no prior expectation of range expansions in the three African bovid species analysed by Templeton & Georgiadis (1996), and only one was inferred for the impala (which in hindsight is biologically reasonable, Templeton & Georgiadis 1996). Another recently published example of an organism with no prior expectation of range expansion is provided by the work of Williams & Benzie (1997) on the high-dispersal starfish *Linckia laevigata* in the Indo-West Pacific. The mtDNA haplotype tree given in that paper was subjected to the nested distance analysis. Out of 23 nested clades, six led to the rejection of the null hypotheses of no geographical associations. Using the inference key, four of these associations were inferred to be due to gene flow constrained by isolation by distance

and two were inconclusive. These conclusions are consistent with the interpretations given in Williams & Benzie (1997). If the inference key is an accurate guide, there should be an excess of inferences of range expansion in the data sets with prior knowledge of range expansion. This is indeed the case: 12 of the data sets with prior knowledge of range expansion led to a statistically significant inference of range expansion (s) and one did not; of the six data sets without prior knowledge of range expansion discussed above, one led to a significant inference of range expansion and five did not. This difference in the frequency of range expansion inference is significant (the two-tailed Fisher's exact test P value is 0.003). Even if the *D. melanogaster* Adh and Amy results are regarded as a single test rather than two, the difference is still highly significant (the two-tailed Fisher's exact test P value is 0.008). The difference between these two data sets is also seen when the Fisher's exact test is applied to nesting clades as the unit of analysis rather than species of gene region: for the 13 data sets with prior knowledge of range expansion, 35 significant nesting clades led to inferences of range expansion while 64 did not; for the six data sets with no prior knowledge of range expansion, one significant nesting clade led to an inference of range expansion and 23 did not. This difference has a two-tailed P value of 0.002. These results reinforce the conclusion that the inference key is not prone to false positives for range expansion.

However, as the *Drosophila buzzatii* example reveals, the inference key is not infallible (Table 1). The patterns described by Cann *et al*. (1987) which were the basis of the criteria incorporated into the inference key require that the expanding populations carry along with them only a subset of the haplotype variation found in the ancestral geographical range, and places great importance upon tip clades found in the expanded area. There are two ways in which tip clades can be found in the expanded area. First, one or more tip clades could be carried over from the ancestral population into the expanding population. Second, after expansion occurs, restricted gene flow (or fragmentation) between the colonized region and the ancestral region would allow the mutational process to create new tip clades that are found primarily or exclusively in the colonized region. Hence, the criteria for range expansion in the inference key will not be satisfied if the range expansion took place by a colonization event associated with an extreme founder effect such that no tip haplotypes were included in the original colonizing population (or at least failed to survive to the present time in the descendants of the colonists), and if the colonization event was sufficiently recent such that no new mutations have arisen in the colonized area.

This is apparently what happened in the case of *D. buzzatii*. Rossi *et al*. (1996) studied mtDNA restriction site variation in Argentinean and Iberian Peninsula populations. Although Rossi *et al*. (1996) had reasonable sample sizes in the Iberian Peninsula, all Iberian flies had only one mitochondrial haplotype (haplotype I in Rossi *et al*. 1996). Moreover, the haplotype that was fixed in the Iberian populations is located at an interior node in the mtDNA haplotype tree (Fig. 3 in Rossi *et al*. 1996) and is the most common haplotype in Argentina. The other data set analysed in the present study that was affected by a long-distance colonization event with an extreme founder effect was that of the macaque monkeys (*Macaca fascicularis*) (Lawler *et al*. 1995). As with *D. buzzatii*, only a single mtDNA haplotype lineage connects the Mauritian colony with its Indonesian/Filipino ancestral population (Lawler *et al*. 1995). This result supports previous work using protein electrophoresis (Kondo *et al*. 1993) of an extreme founder effect even though the current Mauritian population numbers between 25 000 and 35 000 macaques. However, unlike the case of *D. buzzatii* in which no Iberian specific mutations exist, three mutations are found in the macaque haplotype tree (Fig. 2 of Lawler *et al*. 1995) which are restricted to Mauritius, two of which represent fixed differences between Mauritian and ancestral populations. This indicates that mutations have occurred in the Mauritian population after the original colonization event. Hence, even though the Mauritian sample has fewer locations than the *D. buzzatii* sample (three vs. 15), fewer numbers of individuals surveyed (52 vs. 283) and less genetic resolution (17 haplotypes with loops in the haplotype tree vs. 26 haplotypes with no loops), a highly significant colonization event is detected by the nested analysis of the macaque data but not the *D. buzzatii* data. As illustrated by these contrasting data sets, the inference key may not detect all range expansions due to a small founder colony being established so recently that no new mutations are detectable in the sample.

Despite the failure to detect the *D. buzzatii* colonization event, the success in detecting almost all range expansion events consistent with prior knowledge and with only one potential false positive indicates that the criteria for range expansion incorporated into the inference key are a valid and accurate means of identifying range expansion events due either to contiguous range expansion or due to a colonization event. The range expansion criteria in the inference key also have a broad range of generality, given the fact that the detected range expansion events occurred in a diverse set organisms upon many different geographical scales and over a broad range of time scales (with coalescent phenomenon, the relevant time scale is in generations not years). As shown in Table 1, these results also indicate the statistical robustness of the inference key, as the data sets differed greatly in the number of localities sampled, the overall sample size, and the degree of genetic resolution (i.e. the number of haplotypes and depth of the nesting design).

At the beginning of this article, the question was raised 'can this new temporal information be used to shed more light upon the spatial distribution of current allelic variation?' In light of the results presented and summarized in this article, the answer is clearly yes. Haplotype trees allow more power in testing for genetic/geographical associations and more precision in estimating gene flow parameters. Nested analyses of haplotype trees solve one of the major problems of interpreting spatial patterns of genetic variation, i.e. separating the effects of population structure from population history. Moreover, unlike drawing inferences from pictorial overlays of haplotype networks upon geography, the nested geographical analysis coupled with the inference key provides an assessment of statistical significance, explicit inference criteria, and guidance to the researcher for how to collect future samples to make sound biological inference. Finally, the nested analyses allow a dynamic, temporal reconstruction of how population structure and historical events have been interwoven to shape the present-day composition of the population under study. For these reasons, haplotype trees represent a powerful tool that quantitatively and qualitatively enhances the ability to study population structure and recent evolutionary history.

One major limitation of this approach is that it is basically a single-locus analysis. As a result, both evolutionary stochasticity and locus-specific evolutionary forces such as natural selection may either erode power or even mislead the investigator. One way to circumvent this problem would be to perform separate haplotype tree analyses on multiple loci surveyed in the same individuals. This approach is exemplified by the nested geographical distance analyses of the Adh and Amy loci surveyed in the same stocks of *D. melanogaster* in the eastern USA. Both analyses detected significant associations, and in both cases the inference key led to the inference of restricted (but still high) gene flow as the cause. The compatibility of the results across loci is reassuring, but in the future it would be desirable to go beyond an assessment of compatibility. Just as many loci can be pooled together to yield a single estimate of F_{ST} in the traditional, nonhistorical analyses of population structure, methods need to be developed for pooling the results across loci into an integrated analysis of population structure and history. However, even an assessment of compatibility across loci could represent a powerful tool for investigating evolutionary forces. Lewontin & Krakauer (1973) suggested that loci that yield F_{ST} values which are discrepant with the F_{ST} values estimated from the majority of loci are good candidates for having been influenced by natural selection. Given that the results of this study indicate that nested haplotype tree analyses are more powerful and detailed than nonhistorical F_{ST} analyses, it may prove that haplotype tree analyses will be much more powerful in detecting discrepant DNA regions that have been subjected to locus-specific evolutionary forces. Thus, a second need for future development would be to integrate these haplotype tree analyses of geographical distribution with haplotype tree analyses of natural selection (e.g. Templeton 1996b). With such an integrated geographical/selectional analyses, it would be possible to test directly the hypothesis that outlier DNA regions have been subjected to natural selection. This would result not only in cleaner and harder inferences about population structure and history, but would also provide a potentially powerful tool for studying natural selection. This is indeed an exciting prospect.

Acknowledgements

I would like to thank Keri Williamson, Christina Gardenal, Reinaldo Alves de Brito, and Rebecca Reichardt for their many valuable comments and suggestions on an earlier version of this article. I would also like to thank Robert Ricklefs, Eldredge Bermingham, and an anonymous reviewer for their excellent reviews which have resulted in a stronger article.

References

Aquadro CF, Desse SF, Bland MM, Langley CH, Laurie-Ahlberg CC (1986) Molecular population genetics of the alcohol dehydrogenase gene region of *Drosophila melanogaster*. *Genetics*, **114**, 1165–1190.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.

Ayala FJ (1995) The myth of Eve: molecular biology and human origins. *Science*, **270**, 19301936.

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature*, **325**, 31–36.

Castelloe J, Templeton AR (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. *Molecular Phylogenetics and Evolution*, **3**, 102–113.

Davis SK (1986) Population structure and patterns of speciation in *Geomys* (Rodentia: Geomyidae): an analysis using mitochondrial and ribosomal DNA. PhD thesis, Division of Biomedical and Biological Sciences. Washington University, St. Louis.

Davis SK, Strassmann JE, Hughes C, Pletscher LS, Templeton AR (1990) Population structure and kinship in *Polistes* (Hymenoptera, Vespidae): an analysis using ribosomal DNA and protein electrophoresis. *Evolution*, **44**, 1242–1253.

Ewens WJ (1983) The role of models in the analysis of molecular genetic data, with particular reference to restriction fragment data. In: *Statistical Analysis of DNA Sequence Data* (ed. Weir BS), pp. 45–73. Marcel Dekker, New York.

Gerber A (1994) The semiotics of subdivision: an empirical study of the population structure of *Trimerotropis saxatilis* (Acrididae). PhD thesis, Division of Biomedical and Biological Sciences. Washington University, St. Louis.

Hale LR, Singh RS (1991) A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. IV. mitochondrial DNA variation and the role of history vs. selection in the genetic structure of geographic populations. *Genetics*, **129**, 103–117.

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics*, **132**, 583–589.

Kondo M, Kawamoto Y, Nozawa K, Matsubayashi K, Watanabe T, Griffiths O, Stangley MA (1993) Population genetics of the crab-eating macaques (*Macaca fascicularis*) on the island of Mauritius. *American Journal of Primatology*, **29**, 167–182.

Langley CH, Ito K, Voelker RA (1988) Naturally occurring variation in the restriction map of the *Amy* region of *Drosophila melanogaster*. *Genetics*, **119**, 447–454.

Lawler SH, Sussman RW, Taylor LL (1995) Mitochondrial DNA of the Mauritian Macaques (*Macaca fascicularis*): an example of the founder effect. *American Journal of Physical Anthropology*, **96**, 133–141.

Lehman N, Wayne RK (1991) Analysis of coyote mitochondrial-DNA genotype frequencies: estimation of the effective number of alleles. *Genetics*, **128**, 405–416.

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, **74**, 175–195.

Nei M, Takahata N (1993) Effective population size, genetic diversity, and coalescence time in subdivided populations. *Journal of Molecular Evolution*, **37**, 240–244.

Neigel JE, Avise JC (1993) Application of a random-walk model to geographic distributions of animal mitochondrial DNA variation. *Genetics*, **135**, 1209–1220.

Neigel JE, Ball RM, Avise JC (1991) Estimation of single generation migration distances from geographic variation in animal mitochondrial DNA. *Evolution*, **45**, 423–432.

Ovenden JR, White RWG (1990) Mitochondrial and allozyme genetics of incipient speciation in a landlocked population of *Galaxias truttaceus* (Pisces, Galaxiidae). *Genetics*, **124**, 701–716.

Rossi MS, Barrio E, Latorre A, Quezada-Diaz JE, Hasson E, Moya A, Fontdevila A (1996) The evolutionary history of *Drosophila buzzatii*. XXX. Mitochondrial DNA polymorphism in original and colonizing populations. *Molecular Biology and Evolution*, **13**, 314–323.

Slatkin M (1989) Detecting small amounts of gene flow from phylogenies of alleles. *Genetics*, **121**, 609–612.

Slatkin M (1991) Inbreeding coefficients and coalescence times. *Genetical Research*, **58**, 167–175.

Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. *Evolution*, **47**, 264–279.

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics*, **123**, 603–613.

Slatkin M, Maddison WP (1990) Detecting isolation by distance using phylogenies of genes. *Genetics*, **126**, 249–260.

Sykes B, Leiboff A, Low-Beer J, Tetzner S, Richards M (1995) The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. *American Journal of Human Genetics*, **57**, 1463–1475.

Takahata N (1991) Genealogy of neutral genes and spreading of selected mutations in a geographically structured population. *Genetics*, **129**, 585–595.

Takahata N, Slatkin M (1990) Genealogy of neutral genes in two partially isolated populations. *Theoretical Population Biology*, **38**, 331–350.

Templeton AR (1993) The 'Eve' hypothesis: a genetic critique and reanalysis. *American Anthropology*, **95**, 51–72.

Templeton AR (1994) 'Eve': hypothesis compatibility vs. hypothesis testing. *American Anthropology*, **96**, 141–147.

Templeton AR (1996a) Gene lineages and human evolution. *Science*, **272**, 1363.

Templeton AR (1996b) Contingency tests of neutrality using intra/interspecific gene trees: the rejection of neutrality for the evolution of the mitochondrial cytochrome oxidase II gene in the hominoid primates. *Genetics*, **144**, 1263–1270.

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343–351.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

Templeton AR, Georgiadis NJ (1996) A landscape approach to conservation genetics: conserving evolutionary processes in the African Bovidae. In: *Conservation Genetics: Case Histories From Nature* (eds Avise JC, Hamrick JL), pp. 398–430. Chapman & Hall, New York.

Templeton AR, Routman E, Phillips C (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767–782.

Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659–669.

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo Wallace DC (1993a) Asian affinities and continental radiation of the four founding native American mtDNAs. *American Journal of Human Genetics*, **53**, 563–590.

Torroni A, Sukernik RI, Schurr TG, Starikovskaya YB, Cabell MF, Crawford MH, Comuzzie AG, Wallace DC (1993b) mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *American Journal of Human Genetics*, **53**, 591–608.

Wahlund S (1928) Zuzammensetzung von Poulationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, **11**, 65–106.

Williams ST, Benzie JAH (1997) Indo-West Pacific patterns of genetic differentiation in the high-dispersal starfish *Linckia laevigata*. *Molecular Ecology*, **6**, 559–573.

Wilson P (1997) Mitochondrial DNA variation and biogeography among some etheostomid darters of the Central Highlands. PhD thesis, Division of Biomedical and Biological Sciences. Washington University, St. Louis.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

The work presented here is a continuation of the author's research on data analysis through statistical overlays upon gene trees. The gene tree approach was first developed to study the association of genetic variation at candidate trees with clinical phenotypes related to coronary artery disease. The author has since extended this approach to a wide variety of applied and basic problems.

Appendix I: Inference key for the nested haplotype tree analysis of geographical distances

Start with haplotypes nested within a 1-step clade:

1. Are there any significant values for D_c, D_n, or I-T within the clade?
- NO: the null hypothesis of no geographical association of haplotypes cannot be rejected (either panmixia in sexual populations, extensive dispersal in nonsexual populations, small sample size, or inadequate geographical sampling). Move on to another clade at the same or higher level.
- YES: go to step 2.

2. Are the D_c values for tip or some (but not all) interior clades significantly small or is the I–T D_c distance significantly large?
- NO: go to step 11.
- YES: go to step 3.
- Tip/interior status cannot be determined: inconclusive outcome.

3. Are any D_n and/or I-T D_n values significantly reversed from the D_c values, and/or do one or more tip clades show significantly large D_n values or interior clades significantly small D_n values or I-T significantly small D_n with the corresponding D_c values being nonsignificant?
- NO: go to step 4.
- YES: go to step 5.

4. Do the clades (or two or more subsets of them) with restricted geographical distributions have ranges that are completely or mostly nonoverlapping with the other clades in the nested group (particularly interiors), and does the pattern of restricted ranges represent a break or reversal from lower level trends within the nested series (applicable to higher-level clades only)?
- NO: restricted gene flow with isolation-by-distance (restricted dispersal by distance in non-sexual species). This inference is strengthened if the clades with restricted distributions are found in diverse locations, if the union of their ranges roughly corresponds to the range of one or more clades (usually interiors) within the same nested group (applicable only to nesting clades with many clade members or to the highest level clades regardless of number), and if the D_c values increase and become more geographically widespread with increasing clade level within a nested series (applicable to lower level clades only).
- YES: go to step 9.

5. Do the clades (or two or more subsets of them) with restricted geographical distributions have ranges that are completely or mostly nonoverlapping with the other clades in the nested group (particularly interiors), and does the pattern of restricted ranges represent a break or reversal from lower level trends within the nested series (applicable to higher-level clades only)?
- NO: go to step 6.
- YES: go to step 15.

6. Do clades (or haplotypes within them) with significant reversals or significant D_n values without significant D_c values define geographically concordant subsets, or are they geographically concordant with other haplotypes/clades showing similar distance patterns?
- No: go to step 7.
- YES: go to step 13.
- too few clades (< 2) to determine concordance: insufficient genetic resolution to discriminate between range expansion/colinization and restricted dispersal/gene flow. Proceed to step 7 to determine if the geographical sampling is sufficient to discriminate between short vs. long distance movement.

7. Are the clades with significantly large D_n values (or tip clades in general when D_n for I-T is significantly small) separated from the other clades by intermediate geographical areas that were sampled?
- NO: go to step 8.
- YES: restricted gene flow/dispersal but with some long-distance dispersal.

8. Is the species absent in the nonsampled areas?
- NO: sampling design inadequate to discriminate between isolation by distance (short distance movements) vs. long-distance dispersal
- YES: restricted gene flow/dispersal but with some long-distance dispersal over intermediate areas not occupied by the species.

9. Are the different geographically concordant clade ranges separated by areas that have not been sampled?
- NO: past fragmentation. (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by, at least partially, nonoverlapping distributions are mutationally connected to one another by a larger than average number of steps.)
- YES: go to step 10.

10. Is the species absent in the nonsampled areas?
- NO: geographical sampling scheme inadequate to discriminate between fragmentation and isolation by distance.
- YES: allopatric fragmentation. (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by at least partially nonoverlapping distributions are mutationally connected to one another by a larger than average number of steps.)

11. Are the D_c values for some tip clades significantly large, and/or the D_c values for all interiors significantly small, and/or the I-T D_c significantly small?

- NO: go to step 17.
- YES: range expansion, go to step 12.

12. Are the D_n and/or I-T D_n values significantly reversed from the D_c values?
- NO: contiguous range expansion.
- YES: go to step 13.

13. Are the clades with significantly large D_n values (or tip clades in general when D_n for I-T is significantly small) separated from the other clades by intermediate geographical areas that were sampled?
- NO: go to step 14.
- YES: long-distance colonization.

14. Is the species absent in the nonsampled areas?
- between contiguous range expansion and long-distance colonization.
- YES: long-distance colonization.

15. Are the different geographically concordant areas separated by areas that have not been sampled?
- NO: past fragmentation. (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by, at least partially, nonoverlapping distributions are mutationally connected to one another by a larger than average number of steps.)
- YES: go to step 16.

16. Is the species absent in the nonsampled areas?
- NO: go to step 18.
- YES: allopatric fragmentation. (If inferred at a high clade level, additional confirmation occurs if the clades displaying restricted by, at least partially, nonoverlapping distributions are mutationally connected to one another by a larger than average number of steps.)

17. Are the D_n values for tip or some (but not all) interior clades significantly small, or the D_n for one or more interior clades significantly large, or is the I-T D_n value significantly large?
- NO: inconclusive outcome.
- YES: go to step 4.

18. Are the clades found in the different geographical locations separated by a branch length with a larger than average number of mutational steps?
- NO: geographical sampling scheme inadequate to discriminate between fragmentation, range expansion, and isolation-by-distance.
- YES: geographical sampling scheme inadequate to discriminate between fragmentation and isolation-by-distance.