UC Irvine
UC Irvine Previously Published Works

Title

SAMPLING PROPERTIES OF GENEALOGICAL PATHWAYS UNDERLYING POPULATION PEDIGREES

Permalink

https://escholarship.org/uc/item/3xp698p5

Journal

Evolution, 52(4)

ISSN

0014-3820

Authors

Wollenberg, Kurt
Avise, John C

Publication Date

1998-08-01

DOI

10.1111/j.1558-5646.1998.tb01825.x

Copyright Information

Peer reviewed

SAMPLING PROPERTIES OF GENEALOGICAL PATHWAYS UNDERLYING POPULATION PEDIGREES

KURT WOLLENBERG[1] AND JOHN C. AVISE
Department of Genetics, University of Georgia, Athens, Georgia 30602

Abstract.—In sexual species, autosomal alleles are transmitted through multigeneration organismal pedigrees via pathways of descent involving both genders. Here, models assess the sampling properties of these gender-described transmission pathways. An isolation-by-distance model of mating was used to construct a series of computer population pedigrees by systematically varying neighborhood size and the timing of isolation events in sundered populations. For each known pedigree, a matrix of true coancestry coefficients between all individuals in the final generation was calculated and compared (using cophenetic correlations) to mean pairwise times to common ancestry as estimated by sampling varying numbers of gender-defined lineage routes available to individual alleles through that pedigree. When few lineage routes were sampled, agreement between the estimated and the true pedigree was poor and showed a large variance. Agreement improved as more lineage routes were incorporated and asymptotically approached plateau levels predictably relatable to the magnitude of population structure. Results underscore a distinction between the composite genealogical information in a population pedigree and the subsets of that information registered in allelic lineage pathways.

Key words.—Coancestry coefficient, evolution, genealogy, molecular phylogeny, pedigree.

Coalescent theory deals with demographic and other evolutionary factors governing the genealogical histories of alleles (ideally, nonrecombining pieces of DNA at each locus) within and among populations (Hudson 1990). However, the process of sexual reproduction dictates that different alleles follow different transmission routes through an organismal pedigree. Pathways traversed by alleles at unlinked loci are unlikely to be identical to one another or to the composite history of populations (Tajima 1983; Neigel and Avise 1986; Pamilo and Nei 1988; Takahata 1989; Wu 1991; Maddison 1995). Thus, a current challenge is to move toward a coalescent theory that encompasses information at multiple loci (Donnelly 1983; Wu 1991; Slade et al. 1994; Bickebôller and Thompson 1996a,b). Toward that end, we use computer simulations to examine sampling properties of multiple pathways for alleles at unlinked loci through specified organismal pedigrees. We build upon a recent model in which organismal pedigrees are analyzed as collections of extended gender-defined pathways of descent (Avise and Wollenberg 1997).

Gender-defined perspectives on coalescence (Avise 1998) were motivated by a consideration of the properties of mitochondrial DNA (mtDNA) and Y-chromosome gene trees. The gender-defined transmission route for all mitochondrial alleles in a population of higher animals is F → F → F → F . . . (F denotes female), whereas that for mammalian Y chromosomes is M → M → M → M . . . (M denotes male). At an autosomal locus, the situation is more complicated because any allele may have descended from an individual's mother or father, any of four grandparents, eight great-grandparents, and so on. Although the extended transmission route of a particular allele in principle could be specified by gender (e.g., F → M → F → F → M → F . . .), there are many such routes. No single gender-defined transmission history can be specified for a collection of autosomal alleles because of their

different inheritance routes through past generations. The net effect is that multitudinous gender-defined pathways for nuclear alleles coexist within and across loci in organismal pedigrees. The total number of such pathways possible is 2^G, where G is the number of organismal generations in the pedigree, including that of the founders (Avise 1995).

Nonetheless, the historical pathways of collections of autosomal alleles can be considered heuristically in a fashion analogous to that for mtDNA or the Y-chromosome. In other words, in principle it is possible to consider the coalescent process for a theoretical ensemble of alleles that by definition have had a specified history of transmission by gender. For example, a cohort of nuclear alleles that had been transmitted through alternating genders (F → M → F → M → F . . .) in a 20-generation pedigree would constitute one gender-defined pathway among the exact total of 1,048,576 such pathways possible in a pedigree of this length. Each pathway would have its own idiosyncratic coalescent, and each would represent a tiny fraction of the hereditary history of the species (Avise and Wollenberg 1997).

In this study, we investigate the sampling properties of gender-defined inheritance pathways through computer-specified pedigrees. The parameters varied in the model's gene flow regimes under isolation by distance and the timing of vicariant events in sundered populations. Questions addressed include: How many independent transmission pathways sampled from a known pedigree are required to adequately summarize the overall genealogical information in that pedigree? How does spatial population structure affect the outcome?

METHODS

A computer program (PEDIGREE) was written to generate organismal pedigrees with discrete, nonoverlapping generations and a constant number of diploid, dioecious individuals (arranged in a linear array) per generation. Gender (1:1 expected ratio) was randomly assigned to individuals in each

[1] Present address: Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695; E-mail: zooboy@coltrane.gnets.ncsu.edu.
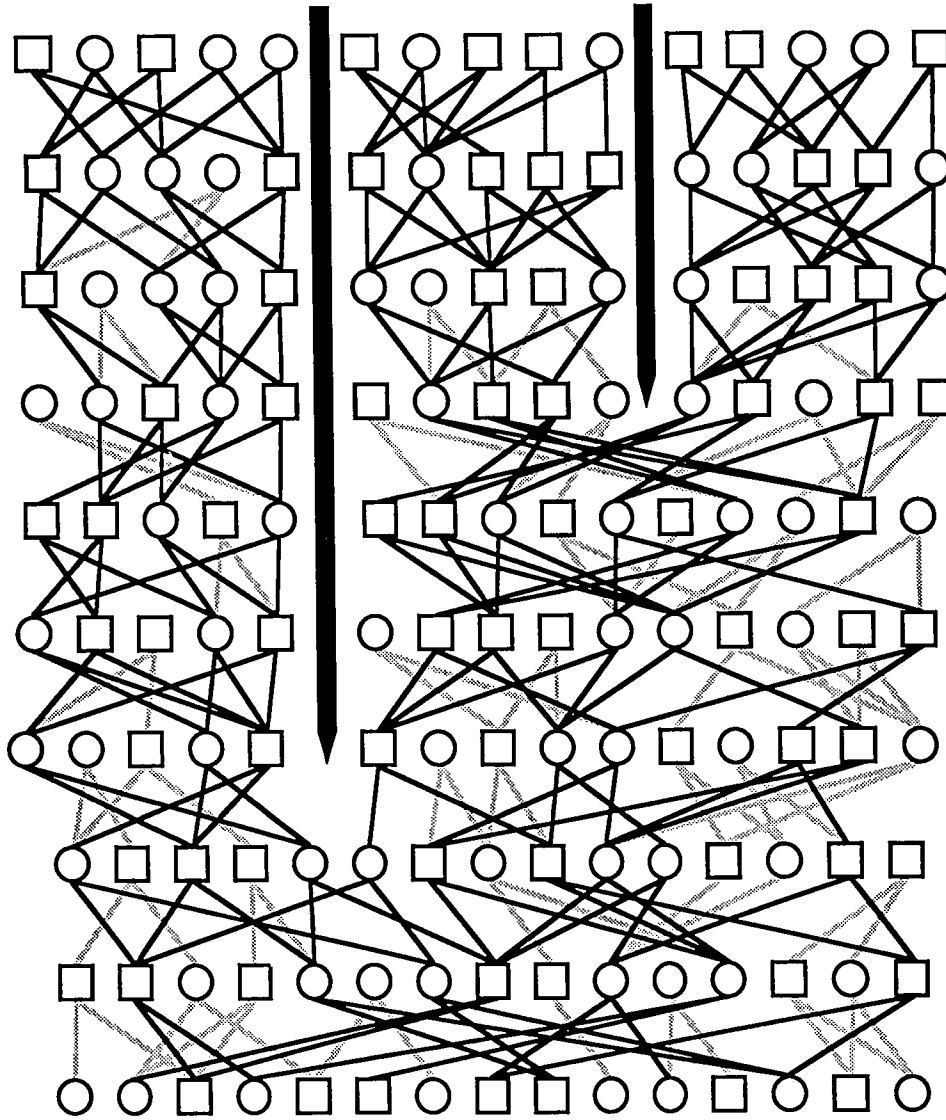
FIG. 1. Illustration of the type of information generated by PEDIGREE, consisting in this case of genealogical relationships of 15 individuals per generation over 10 generations with hard population sundering events at generations 4 and 7. The pedigree begins at the bottom of the figure and grows upward. The genealogical paths in bold are those that lead to individuals in the final generation (top).

generation. New generations were initiated by the assignment of offspring to mothers. Offspring assignment was performed by randomly selecting a female from the current generation and assigning to her a Poisson-distributed random number of offspring (with a mean of two). Because this model assumed an isolation-by-distance mating model, if no males were within one-half of the neighborhood size on either side of the prospective mother, the program randomly chose another female for maternity. This process was repeated until the specified population size was achieved. Mothers were chosen with replacement from the pool of females, so it was possible that some females could be assigned multiple broods and others could be passed over for maternity. Individuals were in a linear array, so females near population edges would have fewer potential mates. This should have a noticeable effect only when neighborhood size is small. However, when the neighborhood size is small, the number of females ex-

periencing an edge effect also is reduced. For this reason, we assumed edge effects on pedigree structure to be negligible.

Next, fathers were assigned to the offspring. The program assumed an isolation-by-distance model in the assignment of paternity, meaning that only males within one-half the neighborhood size on either side of a given mother could sire her offspring. Within this restriction, fathers were chosen randomly on an offspring-by-offspring basis. This process of assigning offspring to mothers and then to fathers was repeated for a specified number of generations. A small example illustrating the nature of hereditary information generated by PEDIGREE is presented in Figure 1.

All analyses were performed using pedigrees of 150 individuals per generation over 100 generations. In the first phase of the analysis, the effect of mating structure was examined for pedigrees that displayed varying degrees of isolation by distance. These pedigrees were generated under
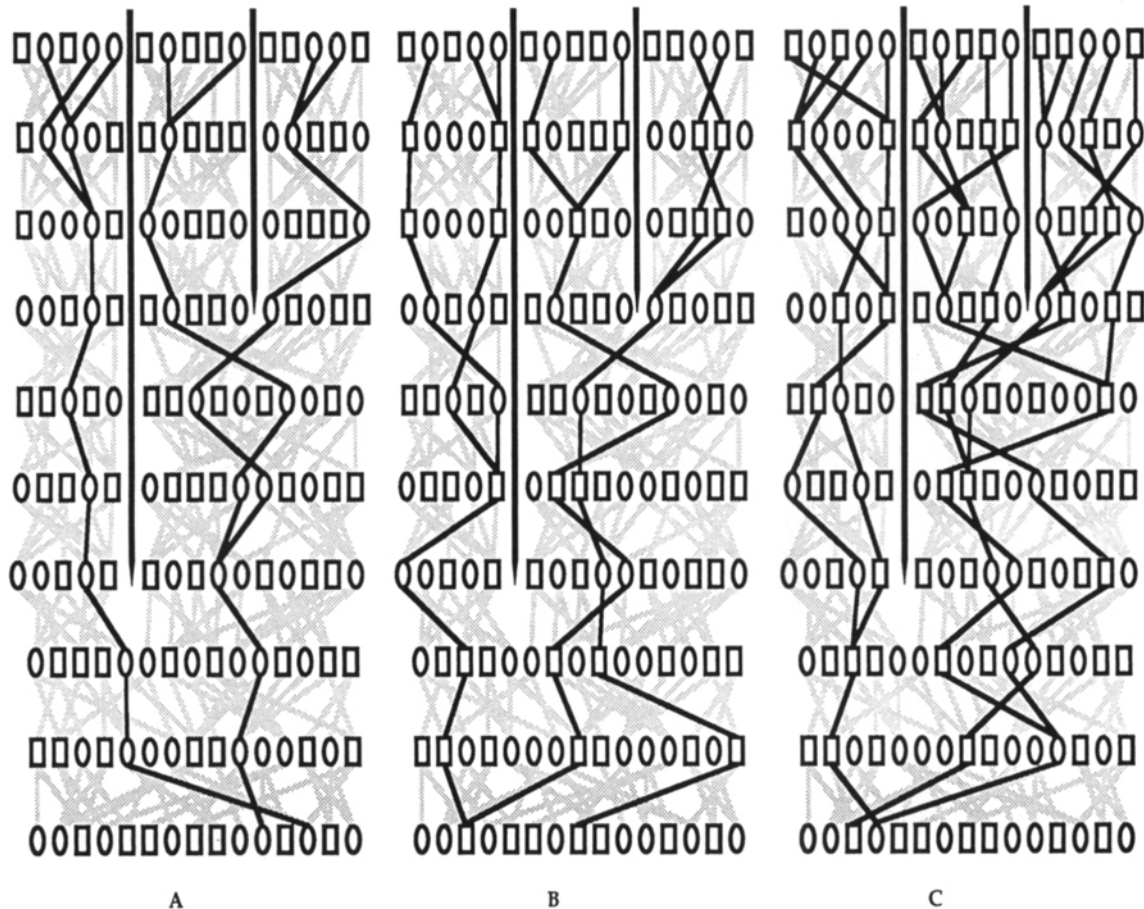
FIG. 2. Examples of transmission pathways through the sample pedigree of Figure 1. Note that the pathway-defined clades in (B) and (C) are not congruent with the true population relationships. (A) The matrilineal (F → F → F → F . . .) pathway traversed, for example, by mtDNA. (B) Example of another gender-defined pathway (M → M → M → F → M → F → F → M → M → F), with the gender of transmission held constant for all individuals within a given generation but randomly changed between generations. (C) A gender-stochastic pathway, with the gender of transmission chosen at random for each individual in each generation.

assumptions that ranged from complete panmixia (fathers chosen at random from the entire linear array) to situations in which fathers were chosen from within two array locations on either side of the mother.

In the second phase of the analysis, further structure was imposed on the pedigree through the inclusion of two vicariant events that firmly sundered the populations (Fig. 1). These were "hard" isolation events: after the specified generation, no interbreeding between the populations was permitted. Three demographic scenarios were employed: (1) two isolation events were 10 generations apart and moved through the pedigree from generations 10 and 20 to generations 80 and 90; (2) one isolation event was held constant at generation 10 and a second varied from generation 20 to 90 in steps of 10 generations; and (3) a first isolation event was varied from generation 10 to 80 and a second was held constant at generation 90. In each scenario, the pedigree ultimately was split into three subpopulations of 50 individuals each. The neighborhood size used in this phase of the analysis was 10 (five locations on either side of the mother).

From each pedigree, true genetic relationships of individuals in the most recent generation were calculated as coefficients of coancestry using the INBREED procedure of the

SAS software package (SAS Institute 1992). A coancestry (or kinship) coefficient is the probability that an allele randomly drawn from one individual is identical by descent (autozygous) within the pedigree to an allele drawn from another individual and is equivalent in value to the inbreeding coefficient for these individuals' prospective offspring (Hartl and Clark 1989). Alleles not identical by descent within the 100-generation pedigree arbitrarily were assumed (for sake of standardization) to have coalesced to a single ancestral allele one generation prior to the initial founding generation. Matrices of pairwise coancestry coefficients for all individuals in the final generation were calculated for each pedigree.

Samples of the pathways for alleles within each pedigree then were employed to estimate genetic relationships between all individuals in the most recent (extant) generation. Two procedures were employed. In the first approach (the "gender-defined" pathway), a random gender for transmission was assigned to each generation in a pedigree. All alleles in that generation of the pedigree then were constrained to pass through that specified gender. Thus, across multiple generations of the pedigree, each extended pathway could be characterized by a multigeneration gender-based description, such as M → M → M → F → M → F → F → M → M → F (Fig.
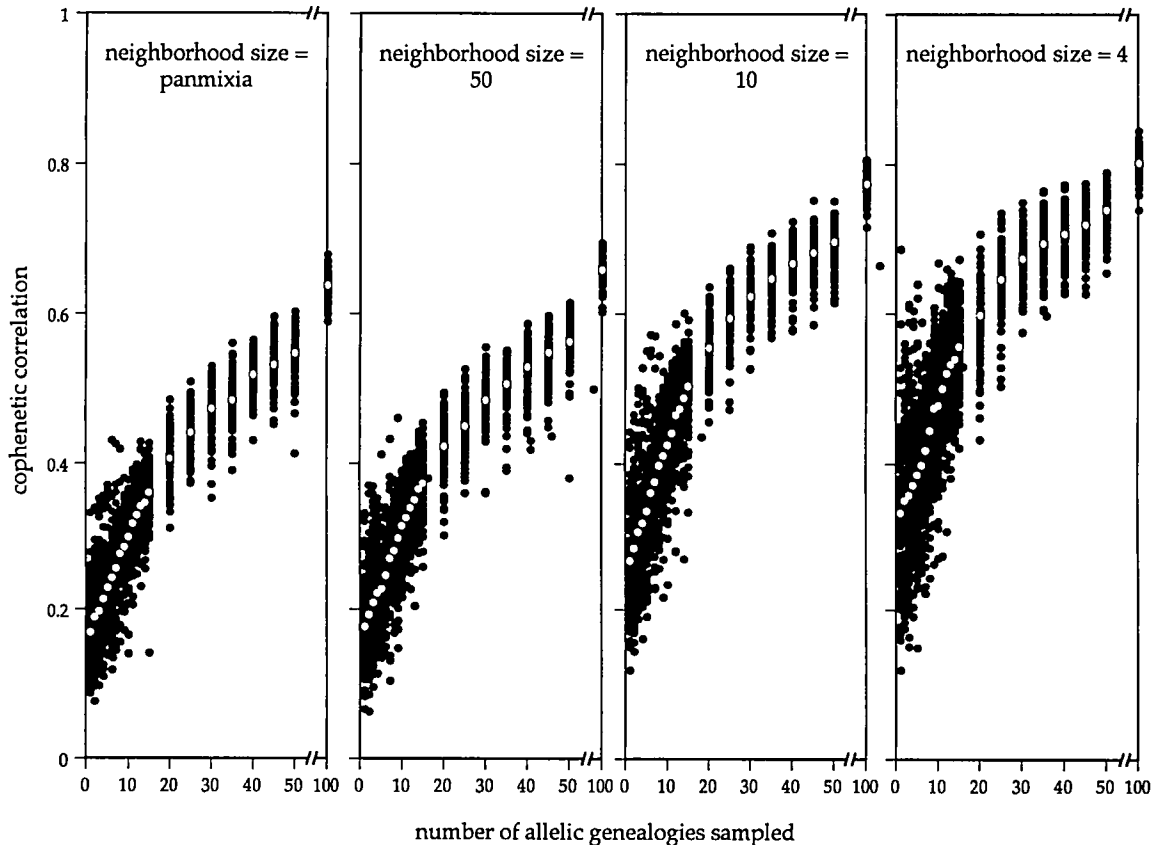
FIG. 3. Cophenetic correlations from the mating structure analyses of gender-defined transmission pathways in panmictic populations and in populations of varying neighborhood size. An open circle is the mean cophenetic correlation (r_m) for the indicated number of allelic pathways sampled (g).

2B). Then, varying numbers of such randomly chosen pathways were sampled from each pedigree and combined to calculate mean times to common ancestry for all pairs of extant individuals.

In a second approach (the "gender-stochastic" pathway), the transmitting gender was varied at random from individual to individual within each generation of the pedigree (Fig. 2C). This heuristic approach was meant to more closely approximate transmission pathways available to multiple alleles at real nuclear loci (at least those that have not recombined intragenically). For example, an allele in any given individual could have stemmed with equal probability from either its mother or its father (rather than only one of them, as under the gender-defined model). As in the earlier approach, mean times to common ancestry across the sampled pathways then were calculated for all pairs of extant individuals. In both the gender-defined and gender-stochastic approaches, the numbers of pathways (g) used to estimate these mean times to common ancestry varied in each pedigree from 1 to 100 (in increments of one from 1 to 15, in increments of five from 15 to 50, and at $g = 100$).

Variability exists in the estimated pairwise relationships between extant individuals as sampled from the multiple transmission pathways through a pedigree. To assess this variation, 100 independent matrices of mean pairwise divergence times were calculated for each numerical class of path-

ways sampled. Then a cophenetic correlation coefficient (Sokal and Rohlf 1962) was calculated for each matrix to measure agreement between the estimated mean times to common ancestry and the true composite coefficient of coancestry. A mean cophenetic correlation (r_m) then was calculated from these 100 replicates. Finally, r_m was regressed against g to produce an equation describing the dynamics of agreement between coancestry coefficients and mean pairwise divergence times of pathways. The first derivative of this equation was evaluated at specific points to describe the slope of the relationship and thereby to quantify any improvement in the cophenetic correlations as more pathways were added to the analyses.

RESULTS

Nonsundered Populations

For populations in the isolation-by-distance models, the mean cophenetic correlations increased and their variances generally decreased as increasing numbers of transmission pathways (g) were averaged into estimates of mean times to common ancestry between pairs of extant individuals (Figs. 3, 4). In other words, inclusion of data from larger numbers of transmission pathways improved the estimates of the composite genealogical information within the known pedigrees. This improvement was rapid for small numbers of transmis-
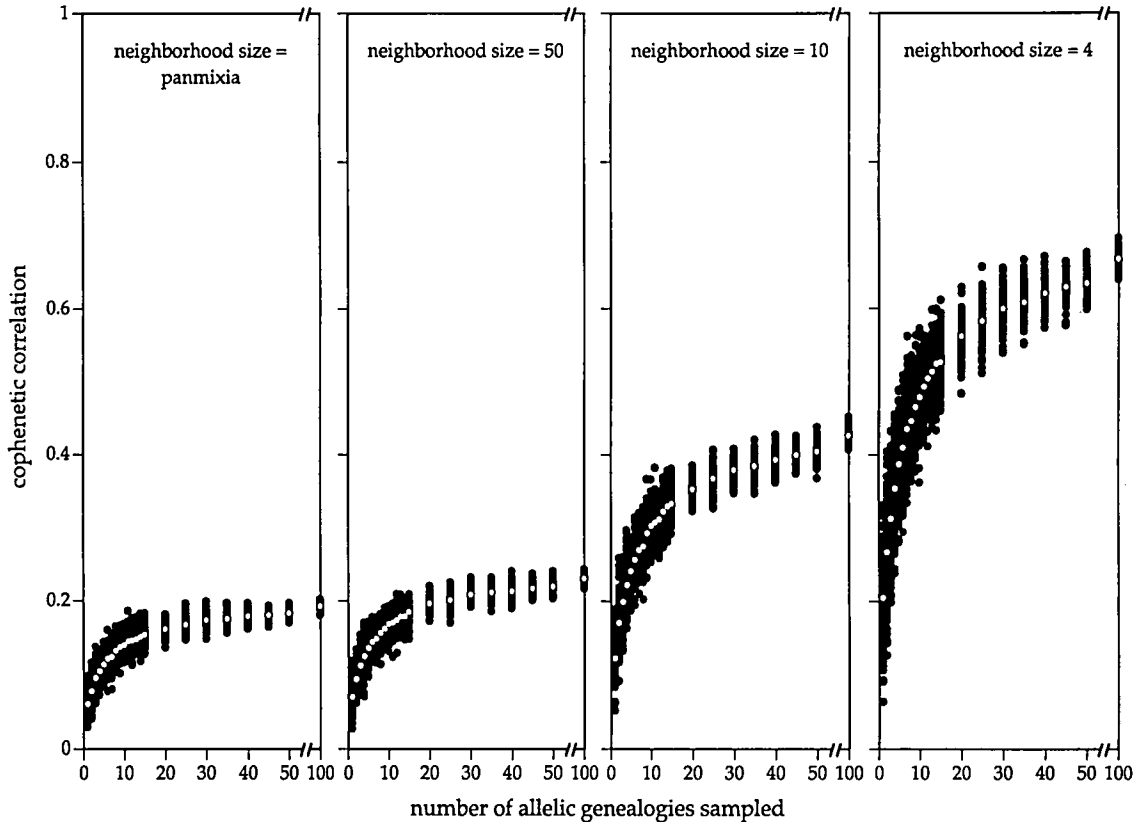
FIG. 4. Cophenetic correlations from the mating structure analyses of gender-stochastic transmission pathways in panmictic populations and in populations of varying neighborhood size (see legend to Fig. 3).

sion pathways (as g increased from about 1 to 20), but thereafter typically slowed considerably as plateau levels of agreement were approached.

For both the gender-defined (Fig. 3) and gender-stochastic (Fig. 4) transmission pathways, the mean cophenetic correlations also increased as neighborhood size decreased. Thus, the imposition of mild spatial structure on populations produced genealogical structures that are better recovered in appraisals of mean pairwise divergence times in the transmission pathways. As shown below, this trend becomes even more pronounced in populations that have been strongly isolated. Finally, the cophenetic correlations for a given set of population conditions and pathways sampled were consistently higher for the gender-defined (Fig. 3) than for the gender-stochastic (Fig. 4) pathways of allelic transmission.

Sundered Populations

Figures 5 and 6 show the trends in mean cophenetic correlations as firm sundering events varied in time among the three scenarios for both the gender-defined and gender-stochastic inheritance pathways. Again, mean cophenetic correlations increased as g increased. In all cases, the maintenance of one vicariant event in the distant past (generation 10) as opposed to more recent generations produced better agreements of reconstructed relationships with the true ones (as seen in the curves labeled T1 = 10 in Figs. 5, 6). Maintenance of the sundering events 10 generations apart (Figs.

5A, 6A) led to roughly the same behavior in the correlation coefficients as did the maintenance of one sundering event at generation 90 and varying the time of the earlier separations (Figs. 5C, 6C). Mean cophenetic correlations again were better for the analyses involving gender-defined as opposed to gender-stochastic pathways of transmission (cf. Figs. 5A vs. 6A, 5B vs. 6B, and 5C vs. 6C).

The initial increase in mean cophenetic correlations with g was rapid and asymptotic under nearly all demographic conditions examined. Also, as g increased and/or as the population sundering events occurred further in the past, the variances across cophenetic correlations typically decreased, often dramatically. Under similar demographic scenarios, the variances in the data for the sundered populations (not shown) displayed similar behaviors to those for the unsundered populations.

To characterize the behavior of the mean cophenetic correlations as a function of the numbers of transmission pathways monitored, slopes of the regressions between these variables were calculated. The regression equation giving the best fit to the data was

$$r_m = (a + bg)/(c + dg),$$

where a, b, c, and d are arbitrary constants defined for each demographic scenario. The fit of the regressions was uniformly tight, with R^2-values (coefficients of determination) ranging from 0.9775 to 0.9997. Slopes of the regressions were calculated as the first derivatives of this equation.
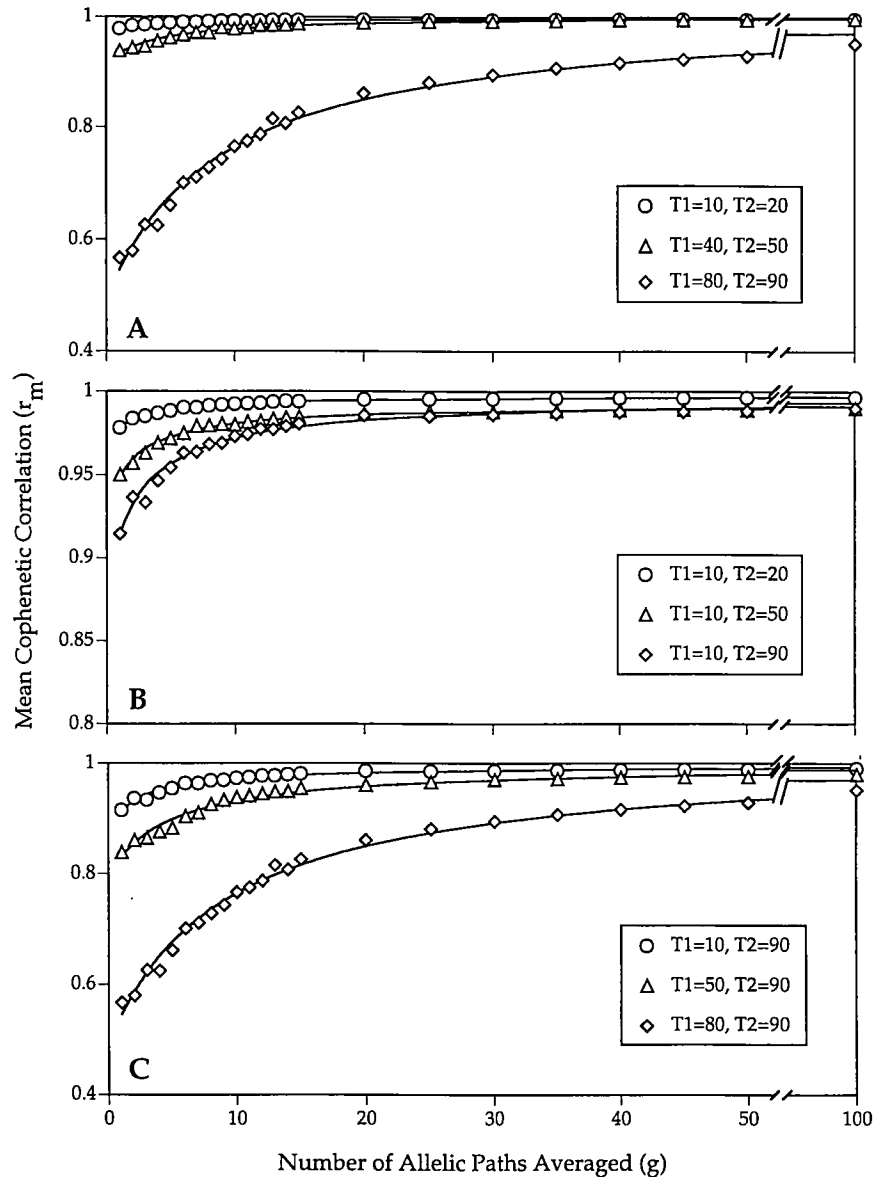
FIG. 5. Mean cophenetic correlations (r_m) from analyses of population sundering events using gender-defined transmission pathways. Plots show r_m versus g (and regressions of g on r_m)for the indicated times that populations were sundered. (A) T1 and T2 separated by 10 generations; (B) T1 = 10 and T2 varied; (C) T2 = 90 and T1 varied.

A slope of 0.01 or less is interpreted here as indicative that the amount of information to be gained by increasing g incrementally is negligible (i.e., a gain in the mean cophenetic correlation of $\leq$ 0.01 per transmission pathway added). For each demographic scenario, the values of g at which the slope became $\leq$ 0.01 are given in Table 1.

Independence of Transmission Pathways

Within each of our 100-generation pedigrees, a total of $2^{100} \cong 1.27 \times 10^{30}$ different gender-defined transmission pathways exist. At most, only four of these are completely independent from one another (i.e., never share for even a single generation an identical gender-based route of transmission from parent to offspring; Avise and Wollenberg

1997): the mtDNA pathway F → F → F → F ... (Fig. 2A); the Y-chromosome pathway M → M → M → M ...; and the reciprocal alternating pathways F → M → F → M ... and M → F → M → F ... .. Do these four independent transmission pathways, when averaged to yield mean pairwise coancestry times for extant individuals, produce substantially better agreements with true coancestry coefficients than do comparable appraisals based on sets of four randomly drawn transmission pathways from the same pedigrees? The empirical answer is no (Fig. 7). Thus, any lack of independence (occasional overlays in gender-defined transmission routes) present in the randomly sampled pathways has no appreciable affect on the recovery of composite genealogical structure from the pedigree.
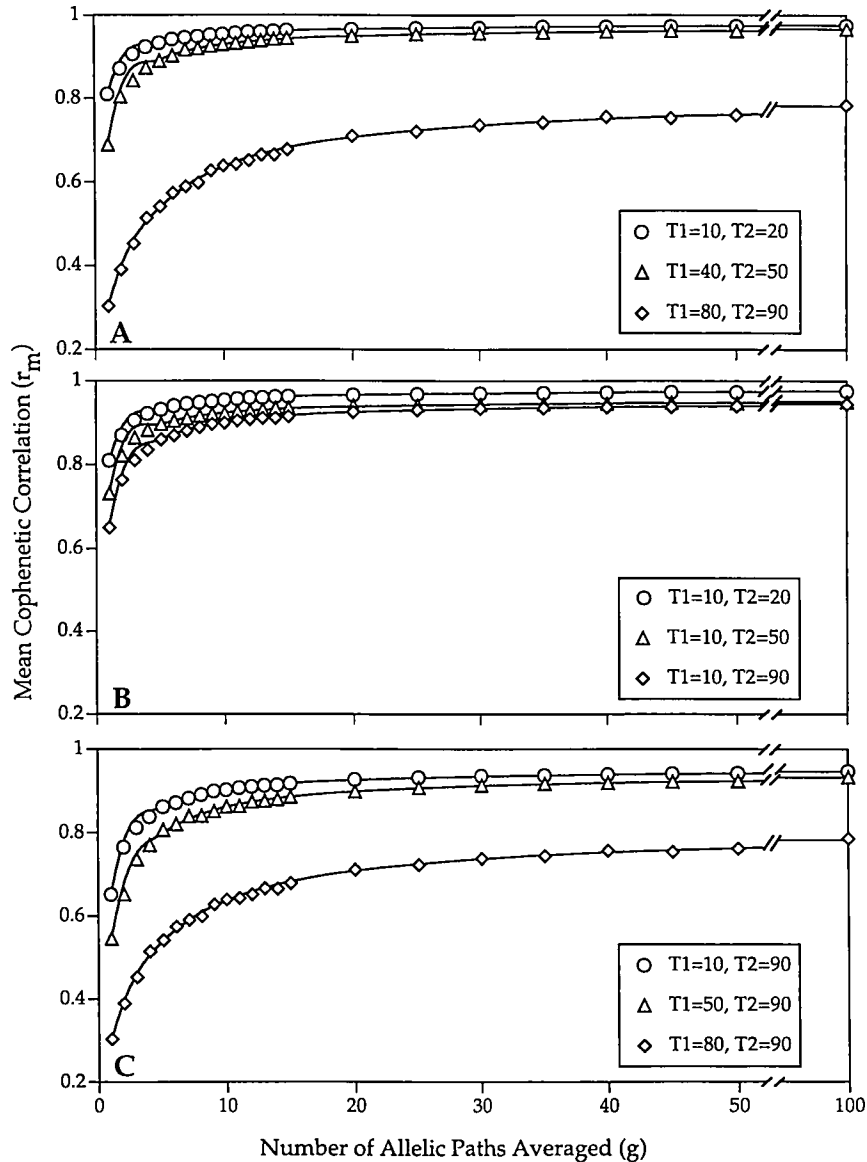
FIG. 6.   Mean cophenetic correlations (r_m) from analyses of population sundering events using gender-stochastic transmission pathways (see legend to Fig. 5).

DISCUSSION

Our approach is inherently genealogical and, hence, bears some similarity to traditional coalescent theory (Griffiths 1980; Kingman 1982a,b; Hudson 1983; Tavaré 1984; Watterson 1984; Griffiths and Tavaré 1997). However, it also introduces several novel modifications. These include a focus on (1) gender-defined transmission pathways (allowing nuclear and cytoplasmic genomes to be viewed in analogous fashion and also serving to equalize mean effective population sizes and expected allelic coalescence times across these genomes), (2) genealogical structures in populations isolated by distance as well as vicariantly sundered over microevolutionary times, and (3) the statistical sampling properties of multiple genealogical pathways.

We have used computer simulations to investigate how population demographic parameters can affect statistical re-

lationships between the true organismal relationships and estimates of those relationships based on finite random samples of transmission pathways. As expected, some conclusions of the current study merely reiterate previous insights from coalescent theory. For example, it was known that increased percentages of gene genealogies come into topological agreement with species phylogeny when internodal distances in the species tree are large relative to the effective sizes of populations traversing those nodes (Tajima 1983; Nei 1987; Pamilo and Nei 1988; Takahata 1989) and when population subdivisions in extant species took place longer in the past such that lineage sorting has proceeded further (Neigel and Avise 1986). Using a different measure of genealogical fit, our current study likewise documents a tendency for better agreement between allelic genealogies and composite organismal pedigrees when nodes between vicariantly separated

TABLE 1. The number of allelic genealogies sampled (g) beyond which the regression slope between g and r_m became ≤ 0.01 for various population scenarios. For the sundered populations, T1 and T2 are the generation numbers of the oldest and most recent isolation events, respectively.

Isolation-by-distance analysis Neighborhood size	g for slope ≤ 0.01
Gender-defined allelic transmission pathway	
panmixia	4
50	5
10	9
4	13
Gender-stochastic allelic transmission pathway	
panmixia	15
50	15
10	17
4	17

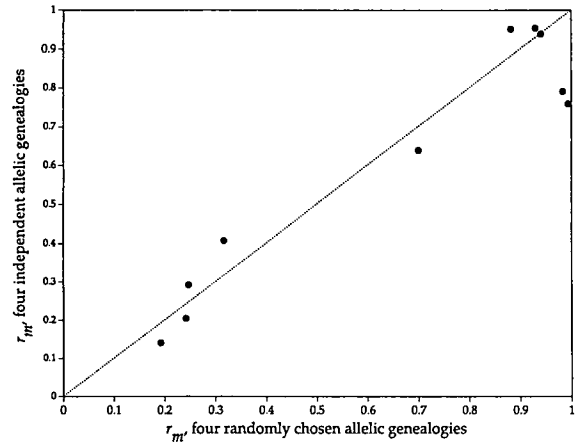| Sundered-population analysis T1, T2	g for slope ≤ 0.01
Gender-defined allelic transmission pathway	
10, 20	1
40, 50	2
80, 90	12
10, 20	1
10, 50	1
10, 90	3
10, 90	3
50, 90	6
80, 90	12
Gender-stochastic allelic transmission pathway	
10, 20	5
40, 50	7
80, 90	12
10, 20	5
10, 50	6
10, 90	7
10, 90	7
50, 90	9
80, 90	12

FIG. 7. Relative performance of sets of four completely independent transmission pathways (vertical axis; see text) versus sets of four randomly chosen transmission pathways (horizontal axis) in generating agreements (r_m-values) with true coancestry coefficients for each of 10 organismal pedigrees. The appraisal for each pedigree is summarized by a closed circle. The dotted line indicates the outcome expected if both sets of transmission pathways performed equally well.

Demographic Influences

Because mean times to common ancestry in transmission pathways and coefficients of coancestry both describe genealogical features of pedigrees, albeit in different ways, we expect the former to approximate more closely the latter as increasing numbers of transmission pathways are averaged. Of interest was the empirical nature of this developing relationship as a function of the number of transmission paths sampled (g). The curves summarizing the cophenetic correlations indicate that estimates of organismal coancestry from gender-defined transmission pathways can improve dramatically with even a small or modest increase in the initial numbers of pathways analyzed. They also demonstrate that different population demographic histories can have appreciable impact on the outcomes. A quantitative measure of such effects was estimated by the slopes of the regressions of r_m on g. These analyses demonstrate that limits exist to the amount of genealogical information retained in finite samples of transmission pathways from a species pedigree and that these limits are approached rapidly by sampling even modest numbers of such pathways (typically 17 or less; Table 1).

Although the simulations examined only a small fraction of the demographic parameter spaces that might be envisioned, they do point to demographic trends likely to be generalizable. For genealogical parameters influenced by population size (namely, coancestry coefficients and pairwise divergence times in transmission pathways), changes in demography affect the magnitude and sampling dynamics of the agreement metric (r_m). For example, the cophenetic correlations for the nonsundered populations were relatively low. Because these nonsundered pedigrees also were comparatively large, r_m tended to decrease as population size increased (relative to the number of generations in the pedigree). The imposition of spatial mating structure on the populations generally increased the mean cophenetic correlations. This probably reflects the smaller sizes of local

extant populations are further apart and more ancient. One difference between the current approach and those of the papers listed above is that we examine the total genealogical content within and among demes in a structured pedigree, rather than merely the topological agreement of gene trees with major nodes in the population phylogeny.

Also, from more traditional theory, differences in historical gene flow are well known to affect population genetic structures by altering gene correlations (e.g., as measured by F-statistics [Wright 1951] or related approaches) within and between populations (Slatkin 1985, 1989; Slatkin and Maddison 1989; Neigel 1997). Similarly, our current study examines genealogical patterns in structured populations and identifies a tendency for better agreement between allelic trees and composite organismal pedigrees in populations that are smaller and more subdivided.

The more novel and interesting aspects of the current study involve the sampling properties of genealogical information as increased numbers of allelic lineages are monitored. Maddison (1995, 1997) has stressed the cloudlike or fuzzy nature of organismal phylogeny due to the fact that gene trees within a species tree have a substantial variance. He and others (e.g., Ball et al. 1990; Hey 1994; Baum and Shaw 1995; Harding 1996) have urged further examination of the statistical sampling properties of that heterogeneity, some aspects of which are developed in this study.

neighborhoods in these populations and resulting increases in the genealogical structures of the pedigrees thereby made available for recovery in the summary statistics.

Similar effects are evident in the behavior of the cophenetic correlations in the sundered as opposed to nonsundered populations, that is, higher r_m-values generally appeared in the former sets of populations, particularly when the isolation events were ancient in the pedigree. Our models treated all subpopulations as part of a larger population of fixed total size. Thus, a stronger genealogical structure within the sundered pedigrees was available for capture in the transmission pathways. This structure stemmed from the isolation events themselves (particularly when ancient) and from decreases in the sizes of the subpopulations that were inherent in the nature of the models.

Influences of population size (relative to pedigree length) also were reflected in the higher r_m-values typically observed in the gender-defined as opposed to gender-stochastic transmission pathways. In theory, gender-defined pathways display a fourfold lower effective population size than do gender-stochastic pathways (where alleles in each generation can be passed through either sex). Smaller effective population sizes lead to higher rates of stochastic lineage sorting (see Tajima 1983; Avise et al. 1984), so gender-defined pathways are more sensitive to the detection of population structure.

Relevance to the Real World

As more transmission pathways were averaged into calculations of mean pairwise coancestry times in our models, estimates of true organismal relationships improved. This was particularly true for spatially structured or sundered pedigrees. The qualitative behaviors of the mean cophenetic correlations (Figs. 3–6) resemble patterns of agreement previously reported as increasing numbers of "gene trees" are employed to estimate organismal phylogenies in interspecific comparisons (e.g., see fig. 2 and table 3 in Takahata 1989). In effect, hereditary routes for alleles within and across loci are mere samples from the totality of possible genomewide transmission pathways that in principle have connected individuals and species in organismal pedigrees. The asymptotic shapes of the curves that describe the sampling dynamics of allelic transmission pathways bear a strong resemblance to analogous curves in other disciplines that characterize, for example, the level of sampling effort necessary for robust estimates of true biotic diversity within an ecological community (e.g., Pielou 1977, p. 302).

Our current models are mostly heuristic. They are not precisely translatable to current empirical efforts to analyze population structure through molecular assays of multiple nuclear (and cytoplasmic) gene trees. In the real world, extended gender-defined transmission histories of alleles (except for those in mtDNA, chloroplast DNA, and perhaps the Y chromosome) typically remain unknown. Even where reliable gene trees can be estimated empirically, resolution normally will not be at the levels of the modest numbers of organismal generations relatable directly to the current simulations.

Nonetheless, heuristic models can be useful in identifying concepts that otherwise might be neglected in interpretations of real data. Our current models examine the consequences of finite sampling from the astronomical numbers of allelic transmission pathways that exist within organismal pedigrees of even modest length. By quantifying the effects of such finite sampling, the models have begun to identify both the limits and strengths of empirical molecular assays of population structure.

LITERATURE CITED

AVISE, J. C. 1995. Mitochondrial DNA polymorphism and a connection between genetics and demography of relevance to conservation. Conserv. Biol. 9:686–690.

———. 1998. The history and purview of phylogeography: a personal reflection. Mol. Ecol. 7:371–380.

AVISE, J. C., AND K. WOLLENBERG. 1997. Phylogenetics and the origin of species. Proc. Natl. Acad. Sci. USA 94:7748–7755.

AVISE, J. C., J. E. NEIGEL, AND J. ARNOLD. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. J. Mol. Evol. 20:99–105.

BALL, R. M., JR., J. E. NEIGEL, AND J. C. AVISE. 1990. Gene genealogies within the organismal pedigrees of random-mating populations. Evolution 44:360–370.

BAUM, D. A., AND K. L. SHAW. 1995. Genealogical perspectives on the species problem. Pp. 289–303 in P. C. Hoch and A. G. Stephenson, eds. Experimental and molecular approaches to plant biosystematics. Monogr. Syst. Mo. Bot. Gard. 53.

BICKEBÖLLER, H., AND E. A. THOMPSON. 1996a. Distribution of genome shared IBD by half-sibs: approximation by the Poisson clumping heuristic. Theor. Popul. Biol. 50:66–90.

———. 1996b. The probability distribution of the amount of an individual's genome surviving to the following generation. Genetics 143:1043–1049.

DONNELLY, K. P. 1983. The probability that related individuals share some section of the genome identical by descent. Theor. Popul. Biol. 23:34–64.

GRIFFITHS, R. C. 1980. Lines of descent in the diffusion approximation of neutral Wright-Fisher models. Theor. Popul. Biol. 17:40–50.

GRIFFITHS, R. C., AND S. TAVARÉ. 1997. Computational methods for the coalescent. Pp. 165–182 in P. Donnelly and S. Tavaré, eds. Progress in population genetics and human evolution. Springer-Verlag, New York.

HARDING, R. M. 1996. New phylogenies: an introductory look at the coalescent. Pp. 15–22 in P. H. Harvey, A. J. Leigh Brown, J. Maynard Smith, and S. Nee, eds. New uses for new phylogenies. Oxford Univ. Press, Oxford, U.K.

HARTL, D. L., AND A. G. CLARK. 1989. Principles of population genetics. 2d ed. Sinauer, Sunderland, MA.

HEY, J. 1994. Bridging phylogenetics and population genetics with gene tree models. Pp. 435–449 in B. Schierwater, B. Streit, G. P. Wagner, and R. DeSalle, eds. Molecular ecology and evolution: approaches and applications. Birkhäuser Verlag, Basel, Switzerland.

HUDSON, R. R. 1983. Testing the constant-rate neutral allele model with protein sequence data. Evolution 37:203–217.

———. 1990. Gene genealogies and the coalescent process. Oxf. Surv. Evol. Biol. 7:1–44.

KINGMAN, J. F. C. 1982a. The coalescent. Stoch. Process. Appl. 13:235–248.

———. 1982b. On the genealogy of large populations. J. Appl. Prob. 19A:27–43.

MADDISON, W. P. 1995. Phylogenetic histories within and among species. Pp. 273–287 in P. C. Hoch and A. G. Stephenson, eds. Experimental and molecular approaches to plant biosystematics. Monogr. Syst. Mo. Bot. Gard. 53.

———. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

NEI, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.

NEIGEL, J. E. 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. Annu. Rev. Ecol. Syst. 28:105–128.

NEIGEL, J. E., AND J. C. AVISE. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pp. 515–534 in E. Nevo and S. Karlin, eds. Evolutionary processes and theory. Academic Press, New York.

PAMILO, P., AND M. NEI. 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5:568–583.

PIELOU, E. C. 1977. Mathematical ecology. Wiley, New York.

SAS INSTITUTE. 1992. SAS technical report P-229, SAS/STAT software: changes and enhancements. Rel. 6.07. SAS Institute, Inc., Cary, NC.

SLADE, R. W., C. MORITZ, AND A. HEIDEMAN. 1994. Multiple nuclear-gene phylogenies: application to pinnipeds and comparison with a mitochondrial DNA gene phylogeny. Mol. Biol. Evol. 11: 341–356.

SLATKIN, M. 1985. Gene flow in natural populations. Annu. Rev. Ecol. Syst. 16:393–430.

———. 1989. Detecting small amounts of gene flow from phylogenies of alleles. Genetics 121:609–612.

SLATKIN, M., AND W. P. MADDISON. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123: 603–613.

SOKAL, R. R., AND F. J. ROHLF. 1962. The comparison of dendrograms by objective methods. Taxon 11:33–40.

TAJIMA, F. 1983. Evolutionary relationships of DNA sequences in finite populations. Genetics 105:437–460.

TAKAHATA, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. Genetics 122:957–966.

TAVARÉ, S. 1984. Line-of-descent and genealogical processes, and their applications in population genetic models. Theor. Popul. Biol. 26:119–164.

WATTERSON, G. A. 1984. Lines of descent and the coalescent. Theor. Popul. Biol. 26:77–92.

WRIGHT, S. 1951. The genetical structure of populations. Ann. Eugen. 15:323–354.

WU, C.-I. 1991. Inferences of species phylogeny in relation to segregation of ancient polymorphisms. Genetics 127:429–435.

Corresponding Editor: S. Palumbi