Molecular Phylogenetics and Evolution
Vol. 20, No. 3, September, pp. 351–360, 2001
doi:10.1006/mpev.2001.0981, available online at http://www.idealibrary.com on IDEAL®

Can Microsatellites Be Used to Infer Phylogenies? Evidence from Population Affinities of the Western Canary Island Lizard (*Gallotia galloti*)

Murielle Richard[1] and Roger S. Thorpe[2]

School of Biological Sciences, University of Wales Bangor, Gwynedd LL57 2UW, United Kingdom

Received March 14, 2000; revised March 27, 2001

Population phylogeographic studies are generally based solely on mtDNA without corroboration, from an independent segregating unit (i.e., nuclear genes), that the mtDNA gene tree represents the organismal phylogeny. This paper attempts to evaluate the utility of microsatellites for this process by use of the Western Canary Island lacertid (*Gallotia galloti*) as a model. The geological times of island eruptions are known, and well-supported mtDNA phylogenies exist (corroborated as the organismal phylogeny rather than just a gene tree by nuclear random amplified polymorphic DNAs (RAPDs)). The allelic variation in 12 populations from four islands (representing five haplotype lineages) was investigated in five unlinked microsatellite loci. Analysis of molecular variance showed this data to be highly structured. A series of genetic distances among populations was computed based on both the variance in allele frequency (i.e., F_{st} related) and the variance in repeat numbers (i.e., R_{st} related). The genetic distances based on the former were more highly correlated with the mtDNA genetic distances than those based on the latter. All trees based on both models supported the primary division shown by mtDNA and RAPDs, which is dated at ca. 2.8 to 5.6 mybp (depending on calibration of the mtDNA clock) and which could, under the evolutionary species concept, be regarded separate species. This was achieved despite theoretical problems posed by the use of few loci, suspected bottlenecks, and large population sizes. The finer details were less consistently represented. Nevertheless, this study demonstrates that even a small number of microsatellites can be useful in corroborating the deeper divisions of a population phylogeny.

Key Words: **microsatellites; phylogeny; Canary Islands; lizards; mtDNA.**

INTRODUCTION

Microsatellites (1- to 5-bp tandem repeats) are codominant, widely distributed in the genome, and highly variable in length, which has resulted in them being widely used to assess population genetic structure and in molecular ecological studies (Jarne and Lagoda, 1996). It has also been suggested that microsatellites may be used to study the evolutionary relationships between groups that have evolved independently for up to several million years (Goldstein *et al.*, 1995a). However, they may possess attributes that limit their usefulness in this area. For example, the range of allele size within microsatellite loci may be limited (e.g., Garza *et al.*, 1995) and they have high mutation rates. As a consequence of the combination of these properties, the potential genetic divergence may be limited. Mutation may be a homogenizing factor due to size homoplasy (i.e., leading to the appearance of alleles already existing) that may overcome the diversifying effect of genetic drift, especially in large populations. Therefore, genetic divergence measurements that use microsatellite data may plateau quickly with time (Nauta and Weissing, 1996; Garza *et al.*, 1995; Paetkau *et al.*, 1997).

Even so, microsatellite data have been used to reconstruct population/species phylogenies in a range of insects and vertebrates (Bowcock *et al.*, 1994; Estoup *et al.*, 1995a; Angers and Bernatchez, 1998; Sültmann and Mayer, 1997; Forbes *et al.*, 1995; Berube *et al.*, 1998; Paetkau *et al.*, 1997; MacHugh *et al.*, 1997). Further evaluation of microsatellite data for reconstructing population phylogenies is important because currently population phylogenies are predominantly based on mtDNA data, and nuclear markers would have great utility in supplying an independent test of whether the mtDNA gene tree represents the organismal phylogeny (Moore, 1995). Use of additional mtDNA genes may increase the probability of resolving the gene tree, but all mtDNA constitutes a single linkage group and therefore additional mtDNA genes do not provide an independent test. The use of introns as

[1] Current address: Laboratoire Fonctionnement et Evolution des Ecosystemes, Université Paris VI, 7 quai Saint Bernard, case 237, 75252 Paris Cedex 05, France.

[2] To whom correspondence should be addressed. Fax: +44 1248 371644. E-mail: r.s.thorpe@bangor.ac.uk.

1055-7903/01 $35.00

nuclear genome markers tends to be more useful at higher taxonomic levels (Giannasi *et al.,* 2001a), and the use of amplified fragment polymorphisms (AFLPs) in phylogenetic reconstruction in vertebrates may be appropriate (Giannasi *et al.,* 2001b), but is not well established.

If microsatellites are to be used in this way, it is not yet clear which genetic distance measures should be used for studying microsatellite data. Several genetic metrics have been developed specifically for microsatellites (e.g., R_{st}; Slatkin, 1995), which use the variance in repeat numbers (VRN) and are compatible with the stepwise mutation model. Classic genetic distances (e.g., F_{st}; transformed for linearity with time) are based on the variance in allele frequencies (VAF) and are compatible with the infinite allele mutation model, which may, or may not, be appropriate to use to describe the microsatellite mutation processes. However, the mutation processes are unlikely to follow just one model, so over-generalization concerning the suitability of a given metric due to compatibility with a given mutation model is probably inappropriate (Gaggiotti *et al.,* 1999 and references therein). If one assumes that the range of allele sizes is unconstrained, VRN metrics should be well suited to phylogenetic reconstruction as they are expected to be linear with time (Kimmel *et al.,* 1996; Shriver *et al.,* 1995; Takezaki and Nei, 1996). However, Nauta and Weissing (1996) have shown that if this assumption does not hold then even these genetic distances will rapidly asymptote with time, particularly in large populations, and Gaggiotti *et al.*'s (1999) simulation study has shown that the large variance of VRN metrics may lead to a poorer performance than VAF metrics unless sample size and loci numbers are large. Considering migration and mutation further complicates the situation. The performance of VRN and VAF metrics in estimating Nm may be evaluated against varying migration/mutation rates (Gaggiotti *et al.,* 1999), but it is not apparent from this how these parameters influence genetic distance. Phylogenetic reconstruction assumes that genetic distances (whatever the type) are not perturbed by migration, i.e., that the effect of migration is trivial relative to mutation. This is likely to be the case when considering isolated populations, e.g., on islands, but less likely to be so for parapatric populations. Consequently, one would expect microsatellite genetic distances to be useful for phylogenetic reconstruction when mutation rates are much higher than migration rates, but not when they are high enough in relation to the time scale for homoplasy to become a problem. Population size is also pertinent. Hence, one would expect microsatellites to be most useful in reconstructing phylogenies between closely related, small, allopatric populations. These expectations must be tested, so it is important to evaluate the performances of these different statistics on data sets with well-supported evolutionary relationships (e.g., Forbes *et al.,* 1995; Paetkau *et al.,* 1997).

The western Canary Islands and the lizard *Gallotia galloti* constitute an extensively studied island/endemic species system. The geological history of the archipelago is well known. Whereas Tenerife probably arose from several precursor islands (Ancochea *et al.,* 1990), each western island is of independent volcanic origin and was not previously attached to another. These islands arose in an east to west series, Tenerife being the oldest at 15.7 million years and El Hierro the youngest at 1.2 million years (Carracedo, 1979; Guillou *et al.,* 1996).

G. galloti occupies these islands at very high densities with large, at least partly contiguous, populations within islands. There is a substantial amount of morphological and molecular (Thorpe, 1996; Thorpe *et al.,* 1994, 1996; Thorpe and Richard, 2001) variation within and between islands. The molecular phylogeny shows two primary lineages: one consisting of La Gomera and El Hierro and the other of Tenerife populations and La Palma, with La Palma and northeast Tenerife as sister groups (Thorpe *et al.,* 1994). The primary split shown by the mtDNA gene trees is independently corroborated by nuclear random amplified polymorphic DNAs (RAPDs). Consequently, this split probably represents the primary division in the organismal phylogeny. These molecular data reveal phylogeographic patterns of colonization (Thorpe *et al.,* 1994) and help to test evolutionary hypotheses (Thorpe, 1996; Thorpe *et al.,* 1996; Thorpe and Richard, 2001). Phylogeographic interpretation of these data (Thorpe *et al.,* 1994) suggests an origin on Tenerife with two westward colonizations to the younger islands as they arose from the seabed (from southwest Tenerife to La Gomera and then La Gomera to El Hierro and independently from northeast Tenerife to La Palma).

As this system has a wide range of molecular divergences (and associated times) it is a useful model to use to investigate the utility of microsatellites for phylogeny construction and to test the reliability of the different types of genetic distances in the case of large populations. Five microsatellite markers (Richard and Thorpe, 2000) have been used and a sample of populations chosen to represent different divergence times and demographic histories.

MATERIAL AND METHODS

Sampling

Noninvasive biopsies (tail tips naturally autotomized) were taken from 372 individuals representing a total of 12 populations of *G. galloti* (2 on each of La Palma and El Hierro, 1 on La Gomera, and 7 on Tenerife; Fig. 1). The Tenerife populations represent both the mtDNA lineages found on that island (Thorpe *et al.,*

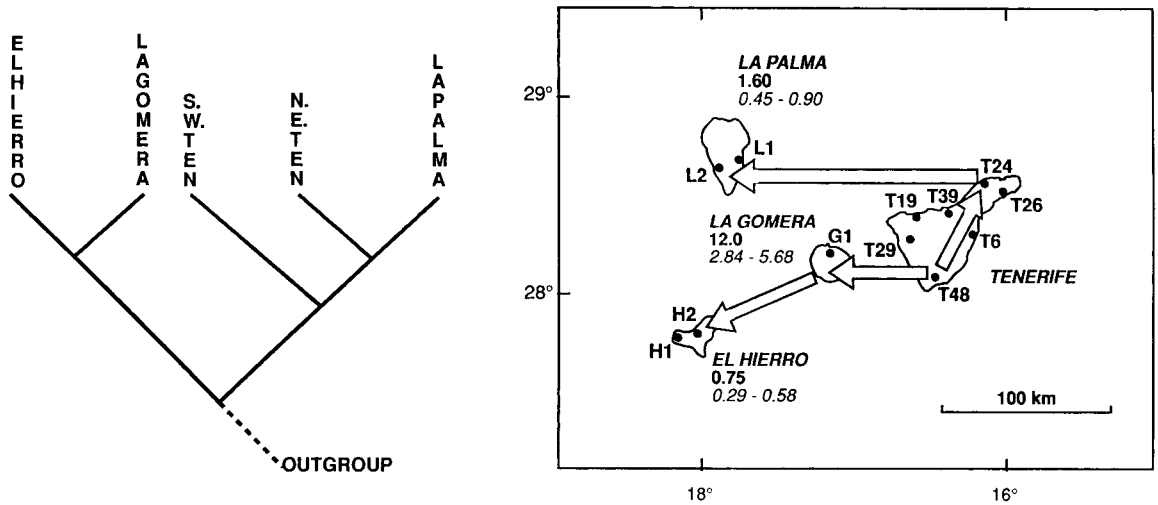


FIG. 1. Site location in the western Canary Islands and hypothesized colonization sequence of *Gallotia galloti* in the western Canary Islands from mtDNA phylogeny. (Left) Phylogeny of *G. galloti* populations based on mtDNA and RAPDs (Thorpe *et al.,* 1994). (Right) Colonization pathway of *G. galloti* among western Canary Islands. Times of island origin (above; boldface) and island colonization (below; italics) are given for each island (in my bp, from Thorpe *et al.,* 1994) based on 1–2% bp change per my (see text for the case for considering 1% change per my).

1996) with localities 19 and 6 being adjacent to the contact between them. The number of individuals analyzed per population range from 26 to 39. DNA extraction was performed with a standard phenol protocol (Sambrook *et al.,* 1989).

Microsatellite Loci

Microsatellite analyses were performed with five dinucleotide microsatellite markers isolated from a *G. galloti* genomic library (A348, $(AC)_{19}$; A49, $(CA)_{10}$; B81, $(TC)_{19}$; B821, $(AC)_{12}$; B967, $(GT)_3AT(GT)_{10}$) (Richard and Thorpe, 2000). Radioactive polymerase chain reaction (PCR) amplifications were carried out in 10 μl of a mixture containing 15 to 30 ng of DNA, 200 nM each primer (100 nM one of them labeled with γ^{32}P), 30 μM each dNTP, 1.5 mM $MgCl_2$, 1× reaction buffer (standard $MgC1_2$-free BRL buffer), and 0.4 units of *Taq* polymerase (BRL). An initial denaturing step of 3 min at 94°C was followed by 35 cycles (94°C for 30 s, 53 or 55°C for 30 s, and 72°C for 15 s) and 2 min at 72°C. PCR products were run through 6% denaturing sequencing polyacrylamide gels and visualized by autoradiography. Allele lengths were determined by comparison to the original clone and individuals with known allele size.

Statistical Analysis

Most of the data analysis was performed with ARLEQUIN 1.1 (Schneider *et al.,* 1996) and GENEPOP 1.2 (Raymond and Rousset, 1995). For each population–locus combination, allele frequency and observed (H_o) and expected (H_e) heterozygosities (gene diversity) were computed. The results were also pooled across loci by Fisher's combined probability method. We also tested our data set for departure from Hardy–Weinberg equilibrium (exact tests using a Markov chain, length 10,000). A permutation test (10,000 permutations) that used the expectation maximization (EM) algorithm (Slatkin and Excoffier, 1996) was used to test for linkage disequilibrium between each pair of loci.

Some genetic metrics, for example $(\delta\mu)^2$ (Goldstein *et al.,* 1995b), are designed specifically as genetic distances, whereas others, e.g., F_{st}, may be used to estimate gene flow (Slatkin, 1985), but can be converted to a dissimilarity coefficient (distance) by the appropriate transformation. Different pairwise genetic metrics have been chosen for evaluation on this data set. Two use the VAF: Nei's (1972) standard distance G_{st} and weighted F_{st} over loci, as in Weir and Cockerham (1984). Two use the VRN: $(\delta\mu)^2$ from Goldstein *et al.* (1995b) and R_{st} from Slatkin (1995). We also used the proportion of shared alleles, D_{ps}. Transformed $G_{st}(G_{st}/(1 - G_{st}))$ and $D_{ps}(D_{ps}/(1 - D_{ps}))$, together with $(\delta\mu)^2$, were computed with the package MICROSAT 1.5 (Minch *et al.,* 1996). The metrics F_{st} and R_{st} were computed with ARLEQUIN 1.1 and a transformation was applied to F_{st} and R_{st} to linearize the distance with population divergence time (Reynolds *et al.,* 1983; Slatkin, 1995). For convenience, transformed metrics are referred to in their untransformed state (i.e., F_{st}, G_{st}, R_{st}, and D_{ps}) and D_{ps} is grouped with VAF metrics.

The genetic structure of the populations was investigated by an analysis of molecular variance (AMOVA; Excoffier *et al.,* 1992; Michalakis and Excoffier, 1996), a hierarchical analysis of variance of allele frequencies. We have used two metrics, the number of different

TABLE 1

Range of Allele Sizes (Number of Repeats), Observed Heterozygosity (H_o), and Gene Diversity (H_e)

	Northeast Tenerife				Southwest Tenerife			La Gomera	El Hierro		La Palma	
	T19	T39	T24	T26	T6	T48	T29	G1	H1	H2	L1	L2
N	30	31	26	28	30	39	32	31	33	30	32	30
A348												
Range	10–19	10–20	10–18	9–19	10–20	10–19	3,10–20	7–13	9–14	9–14	11–19	11–19
H_o	0.467	0.806	0.577	0.607	0.733	0.590	0.781	0.677	0.152	0.167	0.677	0.767
H_e	0.525	0.785	0.554	0.768	0.757	0.679	0.743	0.572	0.146	0.159	0.696	0.859
A49												
Range	7–21	8–20	7–20	8–20	7–20	8–19	7–21	13–24	10–25	10–24	8–18	9–16
H_o	1.000	0.935	0.692	0.964	0.833	0.795	0.906	0.839	0.758	0.767	0.839	0.800
H_e	0.902	0.891	0.870	0.863	0.893	0.851	0.894	0.911	0.899	0.872	0.875	0.810
B81												
Range	6–22	6–21	6–21	6–17	6–23	7–21	7–23	7–10	8–12	9–12	8–12	9–19
H_o	0.900	0.871	0.808	0.857	0.867	0.872	0.813	0.548	0.515	0.467	0.290	0.400
H_e	0.814	0.856	0.903	0.884	0.895	0.845	0.852	0.606	0.459	0.496	0.338	0.508
B821												
Range	11–22	11–23	11–22	8–22	10–22	11–21	11–20	11–23	8–24	8–22	10–23	12–23
H_o	0.833	0.871	0.769	0.857	0.867	0.769	0.844	0.484	0.485	0.667	0.774	0.867
H_e	0.861	0.889	0.869	0.852	0.827	0.826	0.809	0.877	0.875	0.879	0.941	0.872
B967												
Range	8–22	9–23	11–23	11–23	8–21	8–22	9–22	10–28	11–28	11–26	9–23	9–21
H_o	0.867	1.000	0.885	0.893	0.900	0.795	0.969	0.903	0.848	0.700	0.903	0.900
H_e	0.916	0.909	0.851	0.925	0.901	0.882	0.921	0.896	0.896	0.798	0.919	0.876
Over all loci												
MNA	11.6	11.8	10.8	11.6	12.2	11.0	12.0	9.8	10.2	8.8	9.6	9.2
H_o	0.813	0.897	0.746	0.836	0.840	0.764	0.863	0.690	0.552	0.553	0.697	0.747
H_e	0.804	0.866	0.809	0.859	0.855	0.817	0.844	0.772	0.655	0.641	0.754	0.785

alleles (F_{st}) and the average squared size differences (R_{st}). This procedure was used to elucidate the among-island versus the within-island structure of population differentiation. The pairwise genetic distances generated were used to reconstruct trees of genetic similarity with the neighbor-joining algorithm in the package PHYLIP 3.57c (Felsenstein, 1995).

Comparison with Existing Data

Mitochondrial and whole-genome DNA data were available for each island (Thorpe *et al.*, 1993, 1994). They consisted of mtDNA restriction fragment length polymorphisms (RFLPs) (four and six base cutters), sequence of three mtDNA genes (cytochrome *b*, cytochrome oxidase, and 12s rRNA genes which together constitute over 1000 bp) and whole-genome RAPDs. The data were available for some of the localities sampled in the present study (Fig. 1). The pairwise genetic distances calculated for RFLPs from 19 restriction enzymes (D_{RFLP}) were taken from McGregor (1992). Kimura two-parameter genetic distance (Kimura, 1980) was calculated, with PHYLIP, from sequences presented in Thorpe *et al.* (1994). Genetic distances from mtDNA and microsatellite data were compared with the correlation coefficient (r) as a measure of relative correspondence between matrices. The trees were compared visually to elucidate congruence in the major phylogenetic divisions.

RESULTS

Within-Population Data

Allele frequency distributions are available from the corresponding author. The loci all have a rather similar range of allele sizes (17–21 repeats) and there are no consistent trends in which a set of populations has larger, or smaller, sizes across all loci (Table 1). The variation in Tenerife generally encompasses that on the other islands except for the larger alleles of loci A49 and B967 on the south western islands of La Gomera and El Hierro. The mean number of alleles (MNA), over a range of loci, is considered to be a reasonable indicator of genetic variation within a population, given the assumption of mutation-drift equilibrium and similar sample sizes among populations (Nei, 1987). On the more recently colonized islands (La Gomera, El Hierro, and La Palma) MNA is low compared to Tenerife. This is supported by observed heterozygosity (H_o) and gene diversity (H_e) estimates (Table 1) which are low in these more recently colonized islands, particular El Hierro. Only locus B821 has a strong heterozygote deficit and this only for the populations of La Gomera and El Hierro. This may be due to null alleles, selection, or population subdivision (Wahlund, 1928), although the latter is unlikely if the deficit is limited to one locus.

TABLE 2

AMOVA Design and Results

	Source of variation	*df*	Percentage of variation	Significance tests (10,000 permutations)
F_{st}	Among islands	4	11.29	$P < 0.0001$
	Among populations within islands	7	1.01	$P < 0.0001$
	Within populations	732	87.70	$P < 0.0001$
	Total	743		
R_{st}	Among islands	4	24.35	$P < 0.0001$
	Among populations within islands	7	1.95	$P < 0.0001$
	Within populations	732	73.70	$P < 0.0001$
	Total	743		

Note. The populations were grouped according to island (Fig. 1), with the populations from Tenerife forming two "island" groups according to their mtDNA haplotype representing their hypothesized derivation from ancient precursor islands (Thorpe *et al.*, 1996) (southwest = T6, T48 [TSW], T29; northeast = T19, T39, T24 [TNE], T26; with localities T6 and T19 on the border of the transition between the two Tenerife haplotypes).

The tests for departure from Hardy–Weinberg equilibrium showed that most populations were in equilibrium for most loci (except B821 and B967 for one El Hierro population and B821 for La Gomera). Tests for pairwise linkage disequilibrium between loci with the EM algorithm to estimate haplotype frequencies showed no significance and the loci are therefore considered to be at linkage equilibrium.

Among-Population Data

AMOVA tests indicate that there is significant structure in the data at all levels irrespective of whether F_{st} or R_{st} estimates are used (Table 2). The main source of genetic variation is at the within-population level with substantial variation between islands and little, but still highly significant, variation between populations within islands.

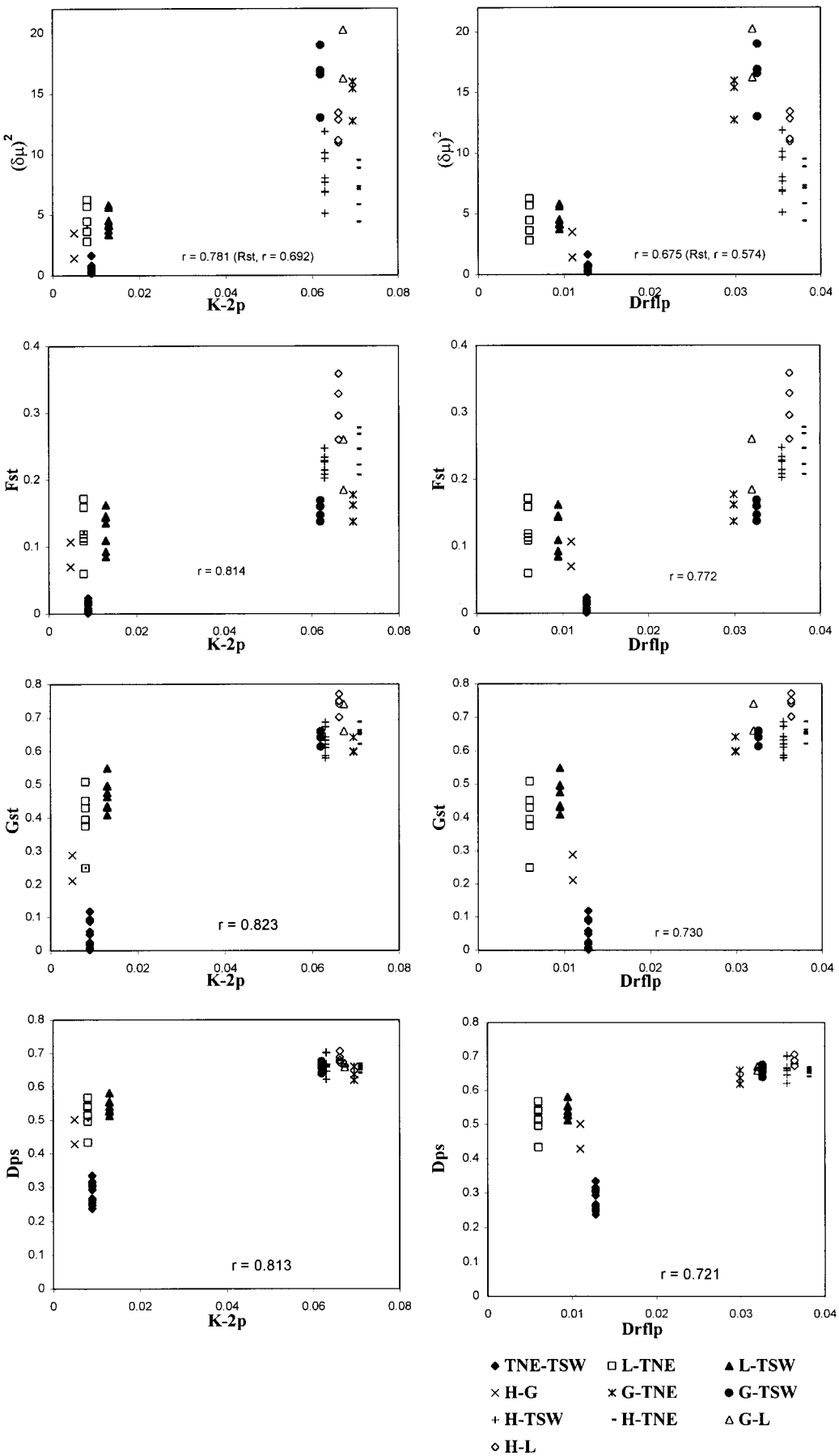
Pairwise genetic distances exhibit a wide range of values with the lowest values between populations within islands, e.g., southwestern Tenerife, and the highest values for comparisons between islands, e.g., La Palma and El Hierro. Scatter diagrams (Fig. 2) show the mtDNA distances to be bimodal with substantial overlap in both VRN and F_{st} microsatellite distances from high- and low-divergence mtDNA distances. However, the D_{ps} and G_{st} distances show no such overlap. Even so, the bimodal nature of the mtDNA distances allows no critical test of linearity with microsatellite distances. The correlations between VAF microsatellite distances on the one hand and RAPD, mtDNA, RFLP, and mtDNA sequence distances on the other hand are all consistently higher than comparable correlations with VRN distances (Fig. 2).

The neighbor-joining method was used to reconstruct trees based on the genetic distance matrices (12 *G. galloti* populations). Some of the five microsatellites would not amplify across the outgroup *G. stehlini* (Richard and Thorpe, 2000). Consequently, a midpoint is indicated on the trees, which in the absence of an outgroup, can be treated as a putative root, as in MacHugh *et al.* (1997). Trees (Fig. 3) were reconstructed from three VAF (D_{ps}, G_{st}, F_{st}) and two VRN ($(\delta\mu)^2$, R_{st}) distances (the tree based on R_{st} had a topology very similar to that based on $(\delta\mu)^2$ and is not illustrated). The nodes are well supported by bootstrap values except for some of the more terminal nodes within the Tenerife/La Palma lineage.

All five trees are fundamentally similar in that the El Hierro and La Gomera populations constitute one primary lineage and the Tenerife and La Palma populations constitute the other primary lineage. Details of the trees do differ, but in all trees based on VAF distances, the Tenerife populations constitute a lineage (which is the sister lineage to La Palma) and the El Hierro populations constitute a lineage (which is a sister lineage to La Gomera). The trees based on VRN distances do not show La Palma and Tenerife, or La Gomera and El Hierro, as sister lineages.

DISCUSSION

All microsatellite genetic distances assume no migration and a constant population size and are expected to perform optimally with small populations, large sample sizes, numerous loci, and lower levels of divergence (i.e., while they are still linear with time). Several facets of the organismal model are suboptimal with regard to these conditions. Migration between the islands is effectively zero compared to mutation, but migration between contiguous populations can be expected to occur within islands. However, this migration is not great enough to cause extensive admixture of mtDNA haplotypes (Thorpe *et al.*, 1996) or to prevent morphological (Thorpe *et al.*, 1996) or genetic (Table 2) (Thorpe and Richard, 2001) differentiation between localities within islands. Historically, if not recently, the population sizes would not be constant, because a bottleneck occurs during interisland colonization. The lower variation (MNA and H_e) on the younger, more recently colonized islands may be a reflection of this. Population sizes on entire islands may be extremely large (perhaps ca. 1×10^8 for Tenerife), certainly much larger than is considered "large" in simulation studies (Gaggiotti *et al.*, 1999), but may be relatively small at a specific locality. The sample sizes available in this study could not be considered small compared to values used in simulation studies (Gaggiotti *et al.*, 1999), but the number of loci is low (albeit comparable to many other such studies). The level of divergence represented in this study is nominally intraspecific. However, if one takes a rate of 2% base pair changes per


(δμ)²
r = 0.781 (Rst, r = 0.692)
K-2p
r = 0.675 (Rst, r = 0.574)
Drflp
Fst
r = 0.814
K-2p
r = 0.772
Drflp
Gst
r = 0.823
K-2p
r = 0.730
Drflp
Dps
r = 0.813
K-2p
r = 0.721
Drflp
◆ TNE-TSW
□ L-TNE
▲ L-TSW
× H-G
✱ G-TNE
● G-TSW
+ H-TSW
- H-TNE
△ G-L
◇ H-L

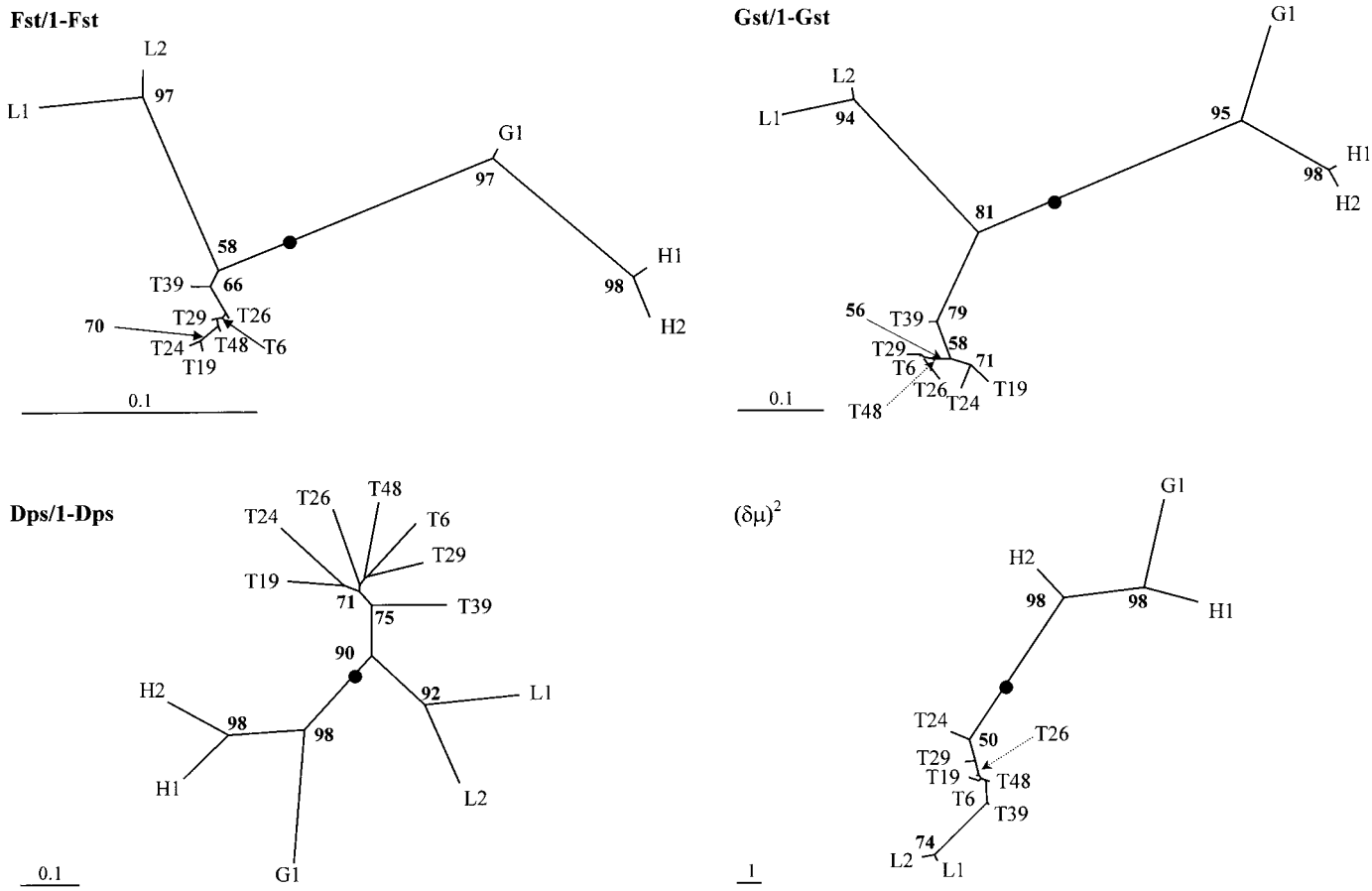


FIG. 3. Neighbor-joining trees reconstructed from different genetic distances: $(\delta\mu)^2$ and linearized D_{ps}, G_{st}, and F_{st}. The linearized R_{st} distance tree is not illustrated as it very similar to the tree based on $(\delta\mu)^2$. Bootstrap values over 50% are given and a solid circle symbolizes the midpoint root.

million years for mtDNA, then the divergence time between major lineages is thought to be 2.8 mybp (Thorpe *et al.*, 1994). However, the rate may be lower for ectothermal vertebrates (Gübitz *et al.*, 2000) and a rate of 1% base pair changes per million years would give a divergence time of 5.6 mybp for these two major lineages (ca. million generations). This range of times is compatible with the geological history of the islands (Fig. 1). Whereas, conventionally, *G. galloti* is considered a single species, under the evolutionary species concept (Frost *et al.*, 1992), some populations would probably be considered independent species. The La Gomera–El Hierro lineage (*G. caesaris*) could be considered a separate species relative to the Tenerife–La Palma lineage (*G. galloti*, *sensu stricto*) defined on the basis of fixed differences: i.e., differences in mtDNA base pairs (Thorpe *et al.*, 1994), whether the blue markings (perhaps sexual signals) are on the forelimbs (as in *G. caesaris*) or on the head/trunk (as in *G. galloti*, *s.s.*), a crenulated collar (present in *G. caesaris;* rare in *G. galloti*, *s.s.*), and mature male snout–vent length (generally less than 9.2 cm for *G. caesaris;* greater than 9.2 cm for *G. galloti*, *s.s.*) (data in Thorpe, 1996). In addition to being recognizable as separate species under the evolutionary species concept, the time scale involved encompasses species divergence in many other groups (Avise *et al.*, 1998). Consequently, the level of divergence between the two main clades in this model is comparable to species divergence in many vertebrates.

FIG. 2. Comparison of pairwise genetic distances from microsatellite (D_{ps}, G_{st}, F_{st}, and $(\delta\mu)^2$) and mtDNA (Kimura two-parameter and RFLP pairwise genetic distances) data. The R_{st} plots are excluded as they are very similar to the $(\delta\mu)^2$ plots, and RAPD plots are excluded as RAPDs did not differentiate between similar populations. The correlations between RAPD and microsatellite distances are $(\delta\mu)^2$, 0.798; R_{st}, 0.709; F_{st}, 0.830; D_{ps}, 0.852; and G_{st}, 0.862. The mtDNA genetic distances are used here as a "time scale" for the divergence times between populations, assuming a molecular clock (in relation with Fig. 1). MtDNA lineage name abbreviations: La Gomera, G; El Hierro, H; La Palma, L; northeast Tenerife, TNE; southwest Tenerife, TSW.

The expectation that all microsatellite genetic distances may lose linearity after several thousands of generations of divergence, essentially due to range constraints in allele sizes (Nauta and Weissing, 1996; Feldman *et al.*, 1997), is not borne out by the scatter diagrams. If the relationship between all microsatellite distances and mtDNA distances is asymptotic over the range presented then one would expect all of the VAF distances, including D_{ps} and G_{st}, and not just the VRN F_{SP} distances to show overlap in distances between mtDNA-similar and mtDNA-divergent populations. This is not the case; so, this study presents no evidence for all microsatellite distances reaching an asymptote over the time scale under consideration (3–6 mybp). However, the rather bimodal nature of mtDNA divergence in this example (Fig. 2) does not allow a critical test of linearity.

Microsatellite distances based on VAF are consistently more highly correlated to mtDNA and RAPD genetic distances than those based on VRN (Fig. 2). The scatter diagrams show that this is generally due to a wider range of VRN distances at high mtDNA distances, which, in turn, may be due to the greater variance of VRN metrics. This comparison of distance matrices argues in favor of the use of VAF distances to elucidate historical relationships in the situation represented by this data set (i.e., a limited number of loci and moderate sample sizes), a conclusion that is in accord with Gaggiotti *et al.*'s (1999) investigation of population structure and gene flow.

Phylogenetic analyses of the previous molecular information (both mtDNA and RAPDs) provide a strongly corroborated organismal phylogeny for *G. galloti, sensu lato,* which has two distinct lineages. One lineage comprises the adjacent islands of La Gomera and El Hierro and the other lineage consists of La Palma and northeast and southwest Tenerife. All the trees reconstructed from microsatellite data show these two lineages, irrespective of which mutation model was the basis for the genetic distance metric and despite the organismal model failing to conform entirely to the assumptions and optimal conditions for the use of microsatellites. Whereas, in a few species, numerous microsatellite loci are available, the number of loci used in this study is more typical of the number that has been available for most species. Consequently, the fact that the essentials of a well corroborated tree can be reconstructed from such a relatively small number of microsatellites argues for their utility in this area. This is in line with Estoup *et al.*'s (1995a,b) phylogenetic study of bees based on seven microsatellite loci, Berube *et al.*'s (1998) study of fin whales based on six microsatellite loci, and Forbes *et al.*'s (1995) study of sheep based on six microsatellite loci. These phylogenetic studies, which used relatively few loci, produced results broadly compatible with mtDNA and other data. However, several other phylogenetic studies have not shown this. For example, Paetkau *et al.*'s (1997) study of bears, based on eight microsatellite loci, gave less phylogenetic resolution at higher levels, a study of Pacific trout (Nielsen *et al.*, 1997), based on three microsatellite loci, showed incongruence to mtDNA, and a study of brook charr (Angers and Bernatchez, 1998), based on five microsatellite loci showed incongruence between mtDNA and microsatellites, as the latter gave information at a finer scale.

Although cross-species amplification occurs, we are probably near the upper limit of the phylogenetic usefulness of microsatellites in this group as not all the primers are conserved in *G. stehlini,* the available outgroup. Consequently we had to use midpoint rooting, with its limitation of the assumption of equal rate of divergence. An alternative would be to use a more closely related outgroup, such as *G. atlantica* or one of the rare and endangered species, e.g., *G. simonyi.*

Failure to meet the assumptions and optimal conditions for the use of microsatellites for phylogenetic reconstruction is evident when one considers the finer details of the trees. During the colonization of La Gomera from Tenerife and that of El Hierro from La Gomera, strong bottlenecks and founder effects are expected to have occurred (although migration between these islands is negligible). Microsatellite genetic distances are expected to be strongly affected by bottlenecks and fluctuations of population size (Nauta and Weissing, 1996; Takezaki and Nei, 1996). Whereas all the trees based on VAF distances revealed the El Hierro populations as a coherent lineage (consistent with geography and mtDNA), the VRN trees failed in this respect. The VRN (but not the VAF) El Hierro–Tenerife distances are lower than the La Gomera–Tenerife distances even though El Hierro was probably colonized from La Gomera. This suggests either that these VRN distances have been more disrupted by these relatively recent bottlenecks than VAF distances and are more sensitive to the recent demographic history of the populations or that the greater variance of the VRN distances renders the finer details unreliable with moderate sample sizes and few loci.

The finer relationships within the La Palma/Tenerife lineage represent a different demography as migration among Tenerife populations may be high compared to mutation rates. In this case neither the VAF nor the VRN trees represent the relationships seen in the mtDNA trees (Thorpe *et al.*, 1994, 1996), i.e., La Palma as a sister group to the northeast Tenerife lineage (from whence it was colonized), with a southwest Tenerife lineage as sister group to this northeast Tenerife/La Palma lineage. Hence, it appears that migration in Tenerife, as expected, has prevented a phylogenetic reconstruction congruent with the mtDNA data for this part of the tree.

In summary, the deeper phylogenetic division is consistently represented by all microsatellite trees while

the finer details are not, although trees based on VAF metrics appear to do marginally better at this than those based on VRN metrics. Against expectations, bottlenecks do not necessarily perturb microsatellite distances (at least not VAF distances), although, as expected, migration does appear to prevent phylogenetic reconstruction with microsatellites. Nevertheless, even with few loci and severe bottlenecks, both VAF and VRN distances may recover phylogenetic relationships among island populations that are millions of years old and may be useful in studying the phylogeography of populations and closely related species.

ACKNOWLEDGMENTS

We thank Marcos Báez, University of La Laguna, Tenerife (corecipient of the EC Grant ERBCHRXCT. 940585 which supported this study) for obtaining permissions. We also thank Anita Malhotra, Per Palsbøll, Oscar Gaggiotti, Andy Stenson, and Wolfgang Wüster for their valuable comments on earlier drafts of the manuscript.

REFERENCES

Ancochea, E., Fuster, J. M., Ibarrola, E., Cendrero, A., Coello, J., Hernan, F., Cantagrel, J. M., and Jamond, C. (1990). Volcanic evolution of the island of Tenerife (Canary Islands) in the light of new K-Ar data. *J. Volcanol. Geotherm. Res.* **44:** 231–249.

Angers, B., and Bernatchez. L. (1998). Combined use of SMM and non-SMM methods to infer fine structure and evolutionary history of closely related brook charr (*Salvelinus fontinalis*, Salmonidae) populations from microsatellites. *Mol. Biol. Evol.* **15:** 143–159.

Avise, J. C., Walker, D., and Johns, G. C. (1998). Speciation durations and Pliestocene effects on vertebrate phylogeography. *Proc. R. Soc. Lond. B* **265:** 1707–1712.

Berube, M., Aguilar, A., Dendanto, D., Larsen, F., Notarbartolo di Sciara, G., Sears, R., Sigurjonsson, J., Urban-R, J., and Palsboll, P. J. (1998). Population genetic structure of North Atlantic, Mediterranean and Sea of Cortez fin whales, *Balaenoptera physalis* (Linnaeus 1758): Analysis of mitochondrial and nuclear loci. *Mol. Ecol.* **7:** 585–599.

Bowcock, A. M., Ruizlinares, A., Tomfohrde, J., Minch, E., Kidd, J. R., and Cavalli-Sforza, L. L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368:** 455–457.

Carracedo, J. C. (1979). "Paleomagnetismo e Historia Volcanica de Tenerife," Aula de Cultura Tenerife, Tenerife, Islas Canarias.

Estoup, A., Garnery, L., Solignac, M., and Cornuet, J-M. (1995a). Microsatellite variation in honey bee (*Apis mellifera* L.) populations: Hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics* **140:** 679–695.

Estoup, A., Tailliez, L., Cornuet, J-M., and Solignac, M. (1995b). Size homoplasy and mutational process of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.* **12:** 1074–1084.

Excoffier, L., Smouse, P. E., and Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131:** 479–491.

Feldman, M. W., Bergman, A., Pollock, D. D., and Goldstein, D. B. (1997). Microsatellite genetic distances with range constraints: Analytical description and problems of estimation. *Genetics* **29:** 207–216.

Felsenstein, J. (1995). PHYLIP (Phylogeny Inference Package) version 3.57c. University of Washington, Seattle, and University Herbarium, University of California, Berkeley.

Forbes, S. H., Hogg, J. T., Buchanan, A. M., Crawford, A. M., and Allendorf, F. W. (1995). Microsatellite evolution in congeneric mammals: Domestic and bighorn sheep. *Mol. Biol. Evol.* **12:** 1106–1113.

Frost, D. R., Kluge, A. G., and Hillis, D. M. (1992). Species in contemporary herpetology: Comments on phylogenetic inference and taxonomy. *Herpetol. Rev.* **23:** 46–54.

Gaggiotti, O. E., Lange, O., Rassmann, K., and Gliddon, C. (1999). A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Mol. Ecol.* **8:** 1513–1520.

Garza, J. C., Slatkin, M., Freimer, N. B. (1995). Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.* **12:** 594–603.

Giannasi, N., Malhotra, A., and Thorpe, R. S. (2001a) Nuclear and mitochondrial DNA phylogenies of the *Trimeresurus* complex: Implications for the gene versus species tree debate. *Mol. Phylogenet. Evol.* **19:** 463–471.

Giannasi, N., Thorpe, R. S., and Malhotra, A. (2001b). The use of amplified fragment length polymorphism (AFLP) in determining species trees at fine taxonomic levels: Analysis of a medically important snake, *Trimeresurus albolabris*. *Mol. Ecol.* **10:** 419–426.

Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L., and Feldman, M. W. (1995a). An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139:** 463–471.

Goldstein, D. B., Linares, A. R., Cavalli-Sforza, L. L., and Feldman, M. W. (1995b). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* **92:** 6723–6727.

Gübitz, T., Thorpe, R. S., and Malhotra, A. (2000) Phylogeography and natural selection in the Tenerife gecko *Tarentola delalandii:* Testing historical and adaptive hypotheses. *Mol. Ecol.* **9:** 1213–1221.

Guillou, H., Carracedo, J. C., Torradom, F. P., and Badiola, E. R. (1996). K-Ar ages and magnetic stratigraphy of a hotspot-induced, fast grown oceanic island—El Hierro, Canary Islands. *J. Volcanol. Geotherm. Res.* **73:** 141–155.

Jarne, P., and Lagoda, P. J. L. (1996). Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **11:** 424–429.

Kimmel, M., Chakraborty, R., Stivers, D. N., and Deka, R. (1996). Dynamics of repeat polymorphisms under a forward–backward mutation model: Within- and between-population variability at microsatellite loci. *Genetics* **143:** 549–555.

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16:** 111–120.

MacHugh, D. E., Shriver, M. D., Loftus, R. T., Cunningham, P., and Bradley, D. G. (1997). Microsatellite DNA variation and the evolution, domestication and phylogeography of taurine and Zebu cattle (*Bos taurus* and *Bos indicus*). *Genetics* **146:** 1071–1086.

McGregor, D. P. (1992). "Mitochondrial DNA Evolution in the Canary Island Lizards (Genus: *Gallotia*)," Ph.D. dissertation, University of Aberdeen, Scotland.

Michalakis, Y., and Excoffier, L. (1996). A generic comparison of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142:** 1061–1064.

Minch, E., Ruiz-Linares, A., Goldstein, D., Feldman, M., and Cavalli-Sforza, L. L. (1996). Microsat (version 1.5): A computer program for calculating various statistics on microsatellite allele data. Stanford University Medical Center, Stanford, CA.

Moore, W. S. (1995). Inferring phylogenies from mtDNA variation:

mtDNA gene trees versus nuclear gene trees. *Evolution* **49:** 718–726.

Nauta, M. J., and Weissing, F. J. (1996). Constraints on allele size at microsatellite loci—Implications for genetic differentiation. *Genetics* **143:** 1021–1032.

Nei, M. (1972). Genetic distance between populations. *Am. Nat.* **106:** 283–292.

Nei, M. (1987). "Molecular Evolutionary Genetics," Columbia Univ. Press, New York.

Nielsen, J. L., Fountain, M. C., and Wright, J. M. (1997). Biogeographic analysis of Pacific trout (*Oncorhynchus mykiss*) in California and Mexico based on mitochondrial DNA and nuclear microsatellites. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 53–73. Academic Press, London.

Paetkau, D., Waits, L. P., Clarkson, P. L., Craighead, L., and Strobeck, C. (1997). An empirical evaluation of genetic distance statistics using microsatellite data from bear (*Ursidae*) populations. *Genetics* **147:** 1943–1957.

Raymond, M., and Rousset, F. (1995). GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *J. Hered.* **83:** 239.

Reynolds, J., Weir, B. S., and Cockerham, C. C. (1983). Estimation for the coancestry coefficient: Basis for a short-term genetic distance. *Genetics* **105:** 767–779.

Richard, M., and Thorpe, R. S. (2000). Highly polymorphic microsatellites in the lacertid *Gallotia galloti* from the western Canary Islands. *Mol. Ecol.* **9:** 1919–1920.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schneider, S., Kueffer, J-M., Roessli, D., and Excoffier, L. (1996). ARLEQUIN (Ver. 1.1): a software environment for the analysis of population genetics data. Genetics and Biometry Lab., University of Geneva, Switzerland.

Shriver, M. D., Jin, L., Boerwinkle, E., Deka, R., Ferrell, R. E., and Chakraborty, R. (1995). A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Mol. Biol. Evol.* **12:** 914–920.

Slatkin, M. (1985). Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* **16:** 393–430.

Slatkin, M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139:** 457–462.

Slatkin, M., and Excoffier, L. (1996). Testing for linkage disequilibrium in genotypic data using the Expectation Maximization Algorithm. *Heredity* **76:** 377–383.

Sültmann, H., and Mayer, W. E. (1997). Reconstruction of cichlid fish phylogeny using nuclear DNA markers. *In* "Molecular Systematics of Fishes" (T. D. Kocher and C. A. Stepien, Eds.), pp. 39–51. Academic Press, London.

Takezaki, N., and Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144:** 389–399.

Thorpe, R. S. (1996). The use of DNA divergence to help determine the correlates of evolution of morphological characters. *Evolution* **50:** 524–531.

Thorpe, R. S., Black, H., and Malhotra, A. (1996). Matrix correspondence tests on the DNA phylogeny of the Tenerife lacertid elucidate both historical causes and morphological adaptation. *Syst. Biol.* **45:** 335–343.

Thorpe, R. S., McGregor, D. P., and Cumming, A. M. (1993). Molecular phylogeny of the Canary Island lacertids (*Gallotia*): Mitochondrial DNA restriction site divergence in relation to sequence divergence and geological time. *J. Evol. Biol.* **6:** 725–735.

Thorpe, R. S., McGregor, D. P., Cumming, A. M., and Jordan, W. C. (1994). DNA evolution and colonization sequence of island lizards in relation to geological history: mtDNA RFLP, cytochrome *b*, cytochrome oxidase, 12s rRNA sequence, and nuclear RAPD analysis. *Evolution* **48:** 230–240.

Thorpe, R. S., and Richard, M. (2001). Evidence that ultraviolet markings are associated with patterns of molecular gene flow. *Proc. Natl. Acad. Sci. USA* **98:** 3929–3934.

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

Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.