

# Deriving Evolutionary Relationships Among Populations Using Microsatellites and $(\delta\mu)^2$ : All Loci Are Equal, but Some Are More Equal Than Others . . .

Pierre-Alexandre Landry,<sup>\*,1</sup> Mikko T. Koskinen<sup>†</sup> and Craig R. Primmer<sup>†</sup>

<sup>\*</sup>Metapopulation Research Group, Division of Population Biology, Department of Ecology and Systematics, FIN-00014, University of Helsinki, Helsinki, Finland and <sup>†</sup>Division of Population Biology, Department of Ecology and Systematics, FIN-00014, University of Helsinki, Helsinki, Finland

Manuscript received September 6, 2001

Accepted for publication May 6, 2002

## ABSTRACT

Numerous studies have relied on microsatellite DNA data to assess the relationships among populations in a phylogenetic framework, converting microsatellite allelic composition of populations into evolutionary distances. Among other coefficients,  $(\delta\mu)^2$  and  $R_t$  are often employed because they make use of the differences in allele sizes on the basis of the stepwise mutation model. While it has been recognized that some microsatellites can yield disproportionate interpopulation distance estimates, no formal investigation has been conducted to evaluate to what extent such loci could affect the topology of the corresponding dendrograms. Here we show that single loci, displaying extremely large among-population variance, can greatly bias the topology of the phylogenetic tree, using data from European grayling (*Thymallus thymallus*, Salmonidae) populations. Importantly, we also demonstrate that the inclusion of a single disproportionate locus will lead to an overestimation of the stability of trees assessed using bootstrapping. To avoid this bias, we introduce a simple statistical test for detecting loci with significantly disproportionate variance prior to phylogenetic analyses and further show that exclusion of offending loci eliminates the false increase in phylogram stability.

**N**UCLEAR microsatellite DNA loci are increasingly employed to assess evolutionary relationships among populations (GOLDSTEIN and SCHLÖTTERER 1999). Because of their high variability, microsatellites can allow for the discrimination of populations where other methods (*e.g.*, DNA sequencing) have failed to detect any polymorphism (BOWCOCK *et al.* 1994; ANGERS and BERNATCHEZ 1997; BRÜNNER *et al.* 1998). Another attractive feature of microsatellites is that the phylogeny of populations can be retraced from a large number of independent loci, whereas, *e.g.*, with mitochondrial DNA, conclusions rely essentially on only one locus. Microsatellites evolve predominantly according to the stepwise mutation model (SMM; OHTA and KIMURA 1973; SCHLÖTTERER 2000; but see also BALLOUX *et al.* 2000), according to which every allele has an equal probability to mutate up or down by a single repeat unit (SHRIVER *et al.* 1993; VALDES *et al.* 1993). Consequently, numerous genetic distance indices have been developed to make use of the evolutionary information contained within differences in repeat numbers among alleles, treating them as quantitative traits (*e.g.*, GOLDSTEIN *et al.* 1995a,b; SHRIVER *et al.* 1995; SLATKIN 1995). GOLDSTEIN *et al.* (1995b) have suggested the use of a coefficient

known as  $(\delta\mu)^2$ , which relies solely upon the differences in mean allele sizes between a pair of populations,

$$(\delta\mu)^2 = \sum_{j=1}^r (m_x - m_y)^2 / r, \quad (1)$$

where  $m_x$  and  $m_y$  are the mean allele sizes (in repeat units) for a given locus  $j$  in populations  $x$  and  $y$ , respectively, and  $r$  represents the number of loci. This SMM-based genetic distance measure was developed specifically to accommodate circumstances in which populations have been isolated for long time periods, *i.e.*, when mutations might account for a marked proportion of the interpopulation microsatellite variation (GOLDSTEIN *et al.* 1995b). Indeed, computer simulation studies have demonstrated  $(\delta\mu)^2$  to remain linear with time (GOLDSTEIN *et al.* 1995b), even when the strict SMM assumption for the occurrence of solely single-step mutations is violated (KIMMEL *et al.* 1996). Moreover,  $(\delta\mu)^2$  has been shown to be relatively robust to fluctuations in population sizes (TAKEZAKI and NEI 1996). In theory, this distance index is expected to provide unbiased estimates of divergence among populations that have been maintained in mutation-drift equilibrium during their evolution, have had constant sizes, and were subjected to no gene flow (ZHIVOTOVSKY 2001). Due to these desired properties,  $(\delta\mu)^2$  has been widely used in evolutionary studies of humans (*e.g.*, GOLDSTEIN *et al.* 1995b; PÉREZ-LEZAUN *et al.* 1997; CALAFELL *et al.* 1998; JIN *et al.* 2000), domestic animals (*e.g.*, RITZ *et al.* 2000), and also numerous wild species (*e.g.*, ANGERS and BERNAT-

<sup>1</sup>Corresponding author: Metapopulation Research Group, Division of Population Biology, Department of Ecology and Systematics, PO Box 65 (Viikinkaari 1), FIN-00014, University of Helsinki, Helsinki, Finland. E-mail: alexandre.landry@helsinki.fi

CHEZ 1997; VALSECCHI *et al.* 1997; GOODMAN 1998; GOLDSTEIN *et al.* 1999; TESSIER and BERNATCHEZ 2000).

However, computer simulation studies have also pointed out that  $(\delta\mu)^2$  distances have an inherently high variance, suggesting that hundreds of loci may be required to attain stable estimates (ZHIVOTOVSKY and FELDMAN 1995). Furthermore, a recent study of four human populations indicated that  $(\delta\mu)^2$  distances can be extremely sensitive to the influence of a very small number of loci, even when >200 loci are utilized; indeed, it was shown that almost one-half of the average interpopulation distance could be attributed to only 2 out of 213 loci (COOPER *et al.* 1999). Therefore, the contribution of each microsatellite to the overall  $(\delta\mu)^2$  distance estimates (hereafter referred to as the “contribution of a locus”) can vary tremendously, to a point where a small number of the assessed loci can dictate the value of the  $(\delta\mu)^2$  distances. Despite the fact that large variance of the  $(\delta\mu)^2$  distance has been recognized, there has been no formal attempt to ascertain the effect this variance has on the topology of phylogenetic trees derived from the distance matrix. In fact, there is good reason to expect that inequalities among loci could create a bias in the topological validation of trees, as assessed with resampling procedures (such as bootstrapping or jackknifing; see LAPOINTE 1998 for a review). Indeed, the main underlying assumptions of bootstrapping are that characters (here loci) are independently and identically distributed (IID; WEST and FAITH 1990; CARPENTER 1992). Identical distribution requires that each locus “must obey one common stochastic model of evolution” (SANDERSON 1995), an assumption that is likely to be violated with microsatellite loci that are characterized typically by heterogeneous mutation rates (WEBER and WONG 1993; PRIMMER *et al.* 1996; DI RIENZO *et al.* 1998; HARR *et al.* 1998) and possible differences in range constraints (GARZA *et al.* 1995; LEHMANN *et al.* 1996). In a case where a small number of loci could disproportionately affect the overall  $(\delta\mu)^2$  distance matrix, resampling (*e.g.*, using the bootstrap) of microsatellites having a small contribution to distances may have little or no effect on the determination of the topology of a tree. For that reason, it could be expected that unequal contribution of loci to the interpopulation distances should upwardly bias the stability estimates (bootstrap support) of  $(\delta\mu)^2$  trees.

To evaluate this potential bias, we have used an innovative randomization method, based on the permutation of alleles within a real microsatellite data set. This approach was inspired by the family of permutation tail probability tests (PTP; FAITH and CRANSTON 1991). In this general framework, statistics pertaining to the topological stability of a phylogenetic tree were computed from real data and then compared to those attained with randomized data, *i.e.*, following the permutation of alleles among individuals or loci. This procedure intends to establish whether the topology of the tree

derived from the real data is more stable than that of a tree built from random data. Any pattern of topological stability emerging from the permuted (randomized) data set would reveal the existence of a bias in the analyses and further allows qualitative evaluation of the extent of this bias.

In this study, we explored the effect of unequal contribution of loci in resolving population evolutionary relationships, using the  $(\delta\mu)^2$  genetic distance. This has been examined using data from 17 microsatellite loci obtained from widely distributed natural populations of European grayling, *Thymallus thymallus* (Salmonidae). We show that a single locus that exhibits strikingly large interpopulation variance can completely dominate the calculation of the genetic distances among populations, introducing a substantial bias in the topology of the corresponding phylogram. Using a permutational approach, we show that the inclusion of disproportionately variable loci falsely increases the similarity of trees between resamples and, consequently, overestimates the bootstrap support values of the  $(\delta\mu)^2$  phylogenetic tree. The extension of these analyses to a second index of interpopulation genetic distances, the so-called  $R_{st}$  (SLATKIN 1995), revealed that such problems are not limited to the  $(\delta\mu)^2$  index, but are to be expected with most SMM-based distance coefficients. These findings imply that caution is warranted when applying  $(\delta\mu)^2$  or  $R_{st}$  distances on microsatellite loci that display heterogeneous levels of diversity.

## METHODOLOGY

**Data:** Seventeen microsatellites were employed to genotype 594 *T. thymallus* individuals sampled from 17 populations across Europe (genotyping details in KOSKINEN and PRIMMER 2001). All specimens were caught from the wild between 1994 and 1999 from areas considered to be mostly unaffected by stocking of European grayling. Linkage equilibrium tests of the loci and Hardy-Weinberg equilibrium tests of the loci and populations did not reveal any significant deviations potentially violating the assumptions of genetic distance estimation and  $(\delta\mu)^2$  phylogram construction (KOSKINEN *et al.* 2002). The microsatellite-based evolutionary relationships of the European grayling populations are relatively clear: Phylogenetic trees based on CAVALLI-SFORZA and EDWARDS's (1967) chord distance ( $D_{CE}$ ) and NEI *et al.*'s (1983)  $D_A$  distance are congruent with mitochondrial DNA-based results (KOSKINEN *et al.* 2002). Yet, the  $(\delta\mu)^2$  distances yielded by each locus were extremely heterogeneous and motivated the investigation of unequal contributions of loci on phylogenetic tree topology.

**Testing the unequal contributions of different loci:** Following Equation 1, a locus displaying large differences in mean allele sizes among populations will contribute more to the overall  $(\delta\mu)^2$  distance than a locus exhibiting small size differences. Therefore, we assessed

the expected overall contribution of each locus  $i$  by calculating its corresponding variance of mean allele size among populations, hereafter referred to as the variance of mean sizes ( $\mu \text{Var}_i$ ),

$$\mu \text{Var}_i = \frac{\sum_{j=1}^k n_j (m_j - \mu)^2}{(k-1)}, \quad (2)$$

where  $k$  is the number of populations,  $n_j$  and  $m_j$  are the sample size and the mean allele size of the  $j$ th population, respectively, and  $\mu$  is the mean allele size of locus  $i$  across all populations. This equation is analogous to the interpopulation variance in an analysis of variance framework (see SOKAL and ROHLF 1995). Let us draw attention to the term  $n_j (m_j - \mu)^2$ , which in fact represents the share of each population  $j$  to the contribution of locus  $i$ ; this is relevant for testing the inequalities among loci (see below).

To evaluate whether a locus had a significantly larger contribution than the other loci to the  $(\delta\mu)^2$  distance matrix, we introduce a new statistic comparing the contribution of a locus  $i$  to the average contribution of the remaining  $l$  loci:

$$F_{\text{cr}_i} = \frac{\mu \text{Var}_i}{\sum_{l=1}^r \mu \text{Var}_l / (r-1)}, \quad \text{where } l \neq i. \quad (3)$$

Under the null hypothesis, the contribution of a single locus should not differ from the average contribution of others, and therefore  $F_{\text{cr}_i}$  will fluctuate around unity. The statistical significance of this ratio can be assessed by computing the null distribution of this statistic using a permutation procedure. Here, loci are assumed to be independent from one another, but the same presumption cannot hold true for populations that are linked by phylogenetic relationships. Under these conditions, a relevant permutation scheme should permute data among loci but not among populations. The probability associated with the null hypothesis was then obtained by permuting the shares of populations pertaining to single-locus contributions (see above). Share values were randomly reassigned to loci, the value of  $F_{\text{cr}_i}$  was reestimated, and the probability of the null hypothesis was obtained by computing the proportion of permuted cases for which the statistic is equal to or larger than the original value. This test is unilateral by design; *i.e.*, only loci displaying contributions larger than others will be declared disproportionate. Significance levels must then be adjusted for multiple testing (*e.g.*, Bonferroni correction). The program to test for unequal contribution (*AnimalFarm*, ver. 1.0) is available from <http://www.helsinki.fi/~primmer> (under publications and data).

While the  $(\delta\mu)^2$  index is the focal point of this study, disparities among loci contributions are also expected to influence other SMM-based distance measures. Handling repeat unit numbers as a quantitative trait can indeed induce variations in the relative importance of differing loci, which would not be the case for indices

based on allele frequencies that always sum up to one (*e.g.*, CAVALLI-SFORZA and EDWARDS 1967; NEI *et al.* 1983). For comparative purposes, we extended our analyses to a second distance coefficient derived for microsatellite markers, namely  $R_{\text{st}}$  (SLATKIN 1995). The two measures are related to a certain point, as allele size variance estimators (within and among populations) unavoidably rely on the sum of squared differences. However, while  $(\delta\mu)^2$  considers only the average allele size differences among populations,  $R_{\text{st}}$  also takes into account the within-population variance, which might compensate for the influence of disproportionately variable loci.

#### Phylogenetic analyses and topological comparisons:

All trees analyzed in this study were recovered using the following procedure: First, interpopulation genetic distance matrices were obtained by summing the  $(\delta\mu)^2$  values across loci or by calculating the  $R_{\text{st}}$  variance ratio (computations carried out with *MsatBootstrap* 1.1, available from: <http://www.helsinki.fi/~primmer>, under publications and data). Then, the corresponding trees were recovered with the Fitch-Margoliash algorithm [computations performed in *FITCH* from the *PHYLIP* package, version 3.752 ( $P = 2$ ,  $G$ , and no negative branches allowed); Felsenstein 1995].

The topological similarity between trees was evaluated with the partition metric ( $P_m$ ; ROBINSON and FOULDS 1981). For a given pair of trees, this index counts the number of clusters occurring in the first or the second tree, but not in both (*i.e.*, the number of topological differences between a pair of trees; PENNY and HENDY 1985). The  $P_m$  values were standardized with the maximum number of topological differences between two trees (*i.e.*,  $2n - 6$ ; STEEL and PENNY 1993), and the 1-complement of this measure was recorded [hereafter referred to as the topological similarity index (TSI); see also LANDRY *et al.* 1996]. Consequently, topologically identical trees have a TSI of 1, whereas completely different trees have a TSI of 0.

#### Contribution of individual loci in the determination of the topology of a phylogram:

Two strategies were employed to evaluate the contribution of single loci to the structure of a phylogenetic tree. First, we compared the tree that is based on the complete data set (17 loci) to trees obtained when each single locus was excluded in turn, *i.e.*, 17 trees based on 16 loci each (using TSI; see above). The *raison d'être* of this procedure was to assess whether tree topology would be altered more by the removal of a locus with a larger  $\mu \text{Var}$  than it would when a locus with a smaller  $\mu \text{Var}$  was excluded from the data set. Second, the topology of the tree derived from each single locus was compared to the tree based on all loci, *i.e.*, 17 trees based on 1 locus each (using TSI; see above). The rationale of this second procedure was that a tree based on a locus contributing more to the overall interpopulation distances should be more similar to the tree based on all 17 loci than a tree ob-

tained from a locus with a smaller contribution to the distance matrix.

**Topology and bootstrap support of phylograms as a function of the number of assessed loci:** Computer simulation studies have revealed that increasing the number of microsatellites ought to decrease the variance of  $(\delta\mu)^2$  (GOLDSTEIN *et al.* 1995b). Accordingly, increasing the number of microsatellites should stabilize the structure of the corresponding phylogram (GOLDSTEIN *et al.* 1995b; TAKEZAKI and NEI 1996; JIN *et al.* 2000). To investigate this, we applied two distinct analyses. First, for every given fixed number of loci (from 2 to 17), 100 bootstrap replicates were generated from the 17-locus data set, and their matching  $(\delta\mu)^2$  or  $R_{st}$  distance matrices were calculated (computations carried out in MsatBootstrap 1.1). Then, the FITCH tree corresponding to each matrix was recovered as described above. Subsequently, the topology of each replicated tree was compared to the topology of the tree based on all 17 loci (using TSI), and the mean TSI value across each set of 100 replicates was recorded as an index of general topological similarity between trees based on a subset of loci and the tree based on all 17 microsatellites.

Second, the topological stability (*i.e.*, bootstrap support) of trees built from either  $(\delta\mu)^2$  or  $R_{st}$  matrices was evaluated as a function of the number of loci, using the analytical design described above. For every set of 100 phylograms previously obtained, a majority-rule consensus tree was calculated (computations in CONSENSE in PHYLIP; FELSENSTEIN 1995). Then, the mean node-support value of the consensus tree was computed and further used as an indicator of topological stability. Large values of average node support characterize trees with clusters that are well supported by the data (*i.e.*, stable trees) whereas small values indicate that some clusters are poorly supported (*i.e.*, unstable trees).

**Data randomization procedures for evaluating contribution of loci to interpopulation distances and subsequent tree topology:** Two related permutation models were employed to randomize the data:

1. *Single permutation:* To eliminate any evolutionary signal that could be related to the differences in the number of repeats between populations, alleles scored at a given locus were randomly assigned to individuals. Under this permutational hypothesis, mean allele sizes are expected to be equal in all populations, leading to expected interpopulation distances approaching zero; any differences in variance between loci are, however, retained.
2. *Double permutation:* In this scheme, the single permutation was followed by a second shuffling that permuted alleles among loci within each individual. Thereby any given allele was assigned to a randomly chosen individual and to a randomly chosen locus. In addition to bringing expected interpopulation distances near zero, this procedure aimed at equalizing

the variance of allele sizes among loci. All permutation procedures were repeated 10 times, and the same analyses of convergence and stability were repeated for each permuted data set.

## RESULTS

The  $(\delta\mu)^2$  genetic interpopulation distances yielded by each locus were found to be extremely variable, maximum values ranging from 4.0 to 169.0 squared repeat units ( $ru^2$ ), with a mean estimate of 27.6  $ru^2$  (Table 1). Marked variations in average  $(\delta\mu)^2$  distances were also observed among loci, ranging from 1.0  $ru^2$  (*BFRO016*) to 194.8  $ru^2$  (*One2*). Thus, it was clear that the influence of individual microsatellites on the overall  $(\delta\mu)^2$  distance matrix greatly differed among loci. Indeed, only 2 of the 17 loci (namely *One2* and *BFRO012*) accounted for 63.6% of the total average interpopulation distance (Table 1). Accordingly, these two microsatellites exhibited much larger allele size variances ( $\mu$  Var) than did the remaining 15 loci (Table 1). Rigorous statistical testing revealed that the contribution of the *One2* locus was significantly larger than the average contribution of other loci, an assertion that did not hold true for any other locus after Bonferroni correction (see Table 1).

Consequently, the importance of each locus in the determination of the structure of the evolutionary relationships among populations was found to differ strikingly among loci. Single removal of strategic loci (*e.g.*, *One2* or *BFRO013*) resulted in dramatic alterations to the topology of the tree (Figure 1). On the other hand, the majority of the other loci could be individually excluded without observing any major change in the topology. In fact, nine loci (53%) could be individually removed without causing a single topological modification (Figure 1). From a reverse angle, the single-locus trees obtained from either *One2* or *BFRO013* showed the highest similarity with the tree based on complete data set (Figure 1), providing additional evidence that these loci were predominant in the determination of the general structure of the phylogenetic tree.

Analyses of the topological structure of  $(\delta\mu)^2$  phylograms as a function of an increasing number of loci also indicated the drastic effects that diverging contributions of the microsatellites can have on the recovery of evolutionary relationships (Figure 2). As expected, increasing the number of resampled loci led consistently to a tree that was increasingly more similar to the one based on all 17 loci. Interestingly, however, the rate of increase in the TSI estimates based on real data did not differ from the one obtained with the single-permuted data (Figure 2A). On the other hand, the pattern observed with the double-permutation procedure showed that the complete randomization of alleles (*i.e.*, among individuals *and* loci) removed any false evolutionary pattern more successfully than the single permutation did (Figure 2A). Interestingly, the same conclusions can be drawn from the trees derived from the variance-based

TABLE 1  
Diversity characteristics of the 17 loci analyzed, shown in decreasing order of  $\mu$  Var

Locus	Maximum $(\delta\mu)^2$ distance ( $\text{ru}^2$ )	Variance of mean sizes ( $\mu$ Var)	Differential contribution test ( $F_{\text{ctr}}$ )	Average interpopulation distance ( $\text{ru}^2$ )	Standard deviation of distances	% of the overall average $(\delta\mu)^2$ distance
<i>One2</i>	1691.0	3280.1	11.42, $p < 0.0002$	194.8	274.7	41.5
<i>BFRO012</i>	627.3	1733.4	4.48, $p < 0.0128$	103.8	174.9	22.1
<i>BFRO013</i>	256.6	678.3	1.48, $p < 0.2485$	39.1	47.7	8.3
<i>Ogo2</i>	273.8	560.6	1.21, $p < 0.2885$	32.0	61.6	6.8
<i>BFRO010</i>	160.2	540.3	1.16, $p < 0.2987$	31.6	47.7	6.7
<i>BFRO018</i>	120.1	240.9	0.50, $p < 0.5757$	12.9	22.0	2.8
<i>BFRO007</i>	49.0	142.7	0.29, $p < 0.7977$	9.4	12.6	2.0
<i>BFRO005</i>	95.1	130.2	0.27, $p < 0.8278$	9.5	14.6	2.0
<i>Str85INRA</i>	39.6	120.7	0.25, $p < 0.8467$	5.8	9.6	1.2
<i>Cocl23</i>	65.8	114.9	0.23, $p < 0.8628$	6.3	10.4	1.3
<i>BFRO015</i>	62.4	85.8	0.17, $p < 0.9318$	9.5	5.0	2.0
<i>Str73INRA</i>	24.7	78.5	0.16, $p < 0.9445$	5.1	5.8	1.1
<i>BFRO017</i>	26.0	48.8	0.10, $p < 0.9831$	2.9	4.1	0.6
<i>BFRO011</i>	12.1	39.1	0.08, $p < 0.9927$	2.1	2.4	0.4
<i>BFRO004</i>	14.2	38.6	0.08, $p < 0.9948$	2.2	3.3	0.5
<i>BFRO009</i>	6.4	17.3	0.04, $p < 0.9998$	1.1	1.4	0.2
<i>BFRO016</i>	4.0	16.2	0.03, $p < 1.0000$	1.0	1.2	0.2

index,  $R_{\text{st}}$ , despite the fact that this distance should account for within-population variance (Figure 2C).

Another striking conclusion of this study was that the average bootstrap support values of  $(\delta\mu)^2$  phylograms based on randomized data (single permutation) were comparable to the bootstrap values observed with real data (Figure 2B). Even in the absence of any evolutionary signal (singly permuted data), bootstrap support values averaging up to 50% were recorded, with the relationships of some nodes being supported by bootstrap values as high as 79%. However, the bootstrap support values of the trees derived from randomized data following the double-permutation procedure were considerably decreased (Figure 2B). Results for the  $R_{\text{st}}$  exhibited a similar pattern (Figure 2D). As for TSI analyses, the  $R_{\text{st}}$  trees based on >14 loci appeared to some extent more stable than those from the singly permuted data.

To further demonstrate the effects of a dominant locus, the analyses were rerun after removal of the locus displaying a contribution significantly larger than all others (*i.e.*, *One2*); this partial data set was also submitted to the single permutation procedure [ $(\delta\mu)^2$  only]. Following this, the topological convergence (*i.e.*, increase in the TSI; Figure 2A) and stability levels (bootstrap support; Figure 2B) of trees observed were found to be comparable to those obtained with the double permutation of the complete data set.

## DISCUSSION

Collectively, these results, based on data from 17 populations spanning the natural range of European grayling, reinforce earlier findings from four human popula-

tions (COOPER *et al.* 1999) that some loci can have a markedly larger influence than others in the determination of  $(\delta\mu)^2$  genetic distances when several loci are combined (see also GOODMAN 1997). More importantly, the results presented here demonstrate for the first time that loci contributing disproportionately to the distance matrix can also strongly bias subsequent phylogenetic tree construction and internal validation procedures. To diagnose this problem within a data set before undertaking any phylogenetic analysis, we have formulated a simple test based on permutations aimed at detecting

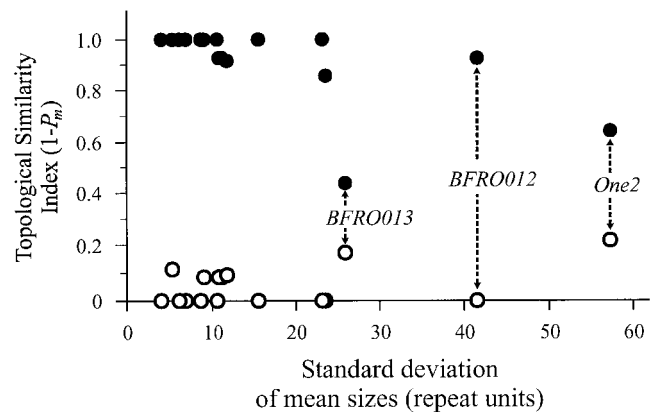


FIGURE 1.—Association between standard deviation of mean allele size and the influence of each locus on the topology of the  $(\delta\mu)^2$  phylogenetic tree. Topological comparisons (as measured by TSI) of trees based on sets of 16 loci (obtained by excluding each particular locus) with the tree based on the complete data of 17 loci are indicated by open circles, whereas topological comparisons of each single-locus tree with the 17-locus tree are represented by solid circles.

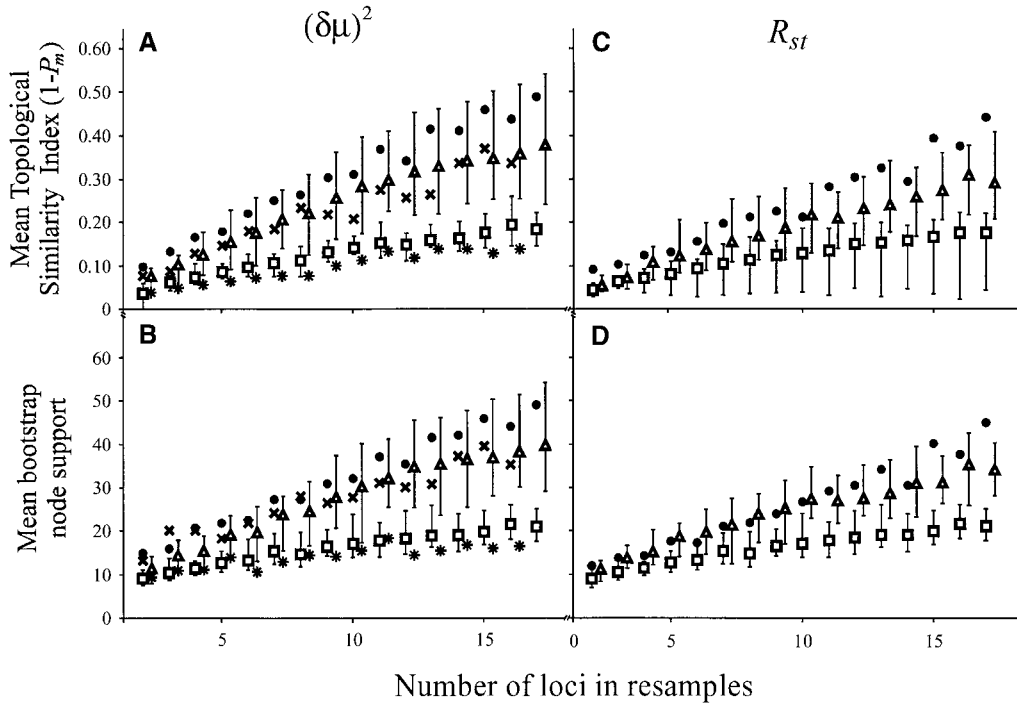


FIGURE 2.—Topological similarity index (TSI, A and C) and bootstrap support (B and D) of resampled phylograms as a function of the number of loci for two distance indices,  $(\delta\mu)^2$  (A and B) and  $R_{st}$  (C and D). Original data (solid circles) and single permutation data (open triangles) show similar trends, while double permutation data (open squares) show a much slower increase in both analyses (see text for permutation methodology). Intervals illustrate the range of values obtained with the replicated permuted data. Results from excluding the most divergent locus are indicated by crosses (original data) and asterisks (single permutation). Some symbols are slightly displaced to the right for clarity.

unequal locus contributions, providing a robust framework to address this problem and identify outliers.

The contribution of a locus (differences in size among alleles) is not necessarily related to its phylogenetic informativeness (ANGERS and BERNATCHEZ 1997; ESTOUP and ANGERS 1998), due to factors including size homoplasy (ESTOUP *et al.* 1995; ANGERS and BERNATCHEZ 1997), variation in mutation rates within (PRIMMER *et al.* 1998; SCHLÖTTERER *et al.* 1998; CROZIER *et al.* 1999) and between (BRINKMANN *et al.* 1998; KAYSER *et al.* 2000) loci, and allele size constraints (GARZA *et al.* 1995; LEHMANN *et al.* 1996). Another fundamental problem is that incompatibility of phylogenetic information contained in different microsatellite loci is common. Nevertheless, validation procedures (such as bootstrapping) are aimed at revealing the degree of confidence of evolutionary trees derived from multiple loci. However, our results illustrate that inequalities among loci can falsely increase the topological congruence (TSI) and stability (bootstrap support) of corresponding phylogenetic trees. Therefore, standard resampling procedures cannot accurately assess the reliability of a tree in the presence of important inequalities among contribution of loci, neutralizing an important instrument to identify the possible phylogenetic incongruence within the data. This also implies that it could be very difficult to recover the true evolutionary relationships among populations in a case where one dominating locus (*i.e.*, with disproportionate contribution) would lead to a phylogenetic tree that is erroneous.

This is evidenced in the European grayling data set, where single-permuted (randomized) data that pro-

duced trees of similar “quality” to the trees based on the original data indicated that the increase in TSI estimates of the nonpermuted microsatellite data with increasing locus number was not a reflection of a meaningful evolutionary pattern. Given that alleles were permuted within each locus separately so that the locus-specific allele size variances were retained, the most likely explanation is that the observed TSI estimates were governed by differences in the magnitude of  $(\delta\mu)^2$  or  $R_{st}$  distances among loci. Indeed, equalizing the variance among loci with the double permutation procedure was enough to remove most of the spurious increase of the TSI in the randomized data, confirming that the increase of TSI within the singly permuted data can be explained solely by differences of mean allele size variance among loci.

Analyses of the bootstrap support values of phylogenetic trees indicated a comparable trend. The similar increase of the average bootstrap values observed in the single-permuted, compared to the original, data substantiates the idea that the reliability of the topology of a phylogenetic tree obtained with  $(\delta\mu)^2$  can be governed by factors that are not related to any evolutionary pattern. The most likely explanation for this finding is that the topological stability was artificially increased due to the effects of larger distances of loci displaying higher variance of mean allele sizes. In support of this, equalizing the variance using the double permutation procedure confirmed that the patterns observed with the single permutation of the complete data (17 loci) are attributable exclusively to differences in variance among loci. It is worth noting that the same analyses, when applied to non-SMM-based distance coefficients [ $D_A$  (NEI *et al.* 1983) and  $D_{CE}$  (CAVALLI-SFORZA and

EDWARDS 1967)], did not reveal any noteworthy increase in the topological stability of phylogenetic trees derived from the single-permuted data (either TSI or bootstrap support; M. T. KOSKINEN, H. HIRVONEN, P.-A. LANDRY and C. R. PRIMMER, unpublished results). Therefore, the surprising stability of trees derived from randomized data does not result from a bias caused by the permutation procedure.

Several studies reported that trees based on  $(\delta\mu)^2$  exhibited lower bootstrap values than did trees based on other distance measures (such as CAVALLI-SFORZA and EDWARDS 1967 or NEI *et al.* 1983; GOLDSTEIN *et al.* 1995b; TAKEZAKI and NEI 1996; JIN *et al.* 2000; M. T. KOSKINEN, H. HIRVONEN, P.-A. LANDRY and C. R. PRIMMER, unpublished results). It has been advocated that the poor stability of  $(\delta\mu)^2$  trees resulted from the high sampling variance associated with  $(\delta\mu)^2$  distances and that increasing the number of loci should increase the stability of such trees (GOLDSTEIN *et al.* 1995b). While this appeared to be the case with our data, deeper analysis revealed that the increase of stability could be explained solely by the large inequalities among loci to the  $(\delta\mu)^2$  distances. Findings of this study reinforce that one should be cautious when utilizing the  $(\delta\mu)^2$  coefficient to reconstruct the evolutionary history of populations, especially when the among-population variance of loci is heteroscedastic. The effects of inequalities among loci are likely to be more important in relatively small data sets, because the variance of  $(\delta\mu)^2$  is expected to decrease when increasing the number of loci (GOLDSTEIN *et al.* 1995b). This consequence might be of primary importance for studies of wild species in which, on average, only six microsatellites are currently utilized in the estimation of evolutionary relationships (M. T. KOSKINEN, H. HIRVONEN, P.-A. LANDRY and C. R. PRIMMER, unpublished results). These findings also suggest that caution is warranted when using so-called "hybrid trees," in which the tree topology is determined using non-SMM distance measure, and interpopulation  $(\delta\mu)^2$  distances are applied to the tree (ANGERS and BERNATCHEZ 1997; ESTOUP and ANGERS 1998). Interpopulation distances based on  $(\delta\mu)^2$  can be erroneously influenced by a small number of loci, implying that even *a posteriori* distance adjustments could in practice be very sensitive to highly disproportionate loci.

Data sets comprising large numbers of loci can indeed display reduced  $(\delta\mu)^2$  variance; nevertheless, it was shown that a very small number of loci (*i.e.*, 2 out of 213) can contribute to almost one-half of the overall distances (COOPER *et al.* 1999). Thus, increasing the number of loci does not necessarily provide a solution to the influence of disproportionate loci. COOPER *et al.* (1999) have also proposed several corrections to the data prior to calculating the  $(\delta\mu)^2$  distances to normalize the outlying loci, but with limited improvements. The main difficulty is to make an adjustment that will not modify the property of linearity with time of the  $(\delta\mu)^2$  index. For exam-

ple, allele sizes could be standardized before analyses, ascertaining that each locus would then have an equal weight in the distance calculations (GOODMAN 1997). While this procedure is warranted for  $R_{st}$  calculations, it is not advisable for  $(\delta\mu)^2$  because large distances will be more constrained than small ones by the standardization, giving up the linearity with divergence time.

Given that differences in locus contributions arise in part because of different range constraints and heterogeneous mutations rates among loci (ESTOUP and ANGERS 1998), a distance measure accounting for these two parameters should produce more accurate estimates of the distances. Some corrections involving these parameters were indeed intended to account for different mutations rates and size constraints among loci (*e.g.*,  $D_{GLS}$ ; POLLOCK *et al.* 1998), but the use of these coefficients requires reasonably accurate estimates of evolutionary parameters of each microsatellite, which at present cannot be reliably achieved (ESTOUP and ANGERS 1998; COOPER *et al.* 1999; ELLEGREN 2000). Therefore, in cases where outlying loci can be identified using the test reported here, the simplest and most straightforward option for eliminating any false phylogram stability (bootstrap support) and convergence toward a final solution (TSI) is to remove the offending loci from subsequent analyses (*e.g.*, ZHIVOTOVSKY *et al.* 2000).

## CONCLUSION

The problem of unequal contribution of microsatellites combined with the use of an SMM-based distance coefficient [ $(\delta\mu)^2$  or  $R_{st}$ ] should be considered when assessing the evolutionary relationships among populations and especially when utilizing validation procedures based on resampling. Results presented here suggest that the use of  $(\delta\mu)^2$  or  $R_{st}$  should be restricted to arrays of loci displaying comparable amounts of variance, to minimize the influence of exceptionally highly variable loci. Thus, in establishing the phylogenetic relationships among populations with  $(\delta\mu)^2$ , all microsatellites are considered equal, but it clearly appears that some are more equal than others . . . (to paraphrase ORWELL 1945).

The authors are grateful to F.-J. Lapointe for stimulating discussions in the early stages of this study and to J. N. Painter, D. L. Johnson, and three anonymous reviewers for helpful suggestions to improve this manuscript. Researchers providing the *T. thymallus* samples from across Europe are also much appreciated. This work was supported by a National Sciences and Engineering Research Council of Canada postdoctoral fellowship awarded to P.-A. Landry and by the Biological Interactions Graduate School, the University of Helsinki, and the Academy of Finland (project no. 172964 and Centre of Excellence Program 2000-2005, grant no. 44887).

## LITERATURE CITED

- ANGERS, B., and L. BERNATCHEZ, 1997 Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Mol. Biol. Evol.* **14**: 230-238.

- BALLOUX, F., H. BRÜNNER, N. LUGON-MOULIN, J. HAUSSER and J. GOUDET, 2000 Microsatellites can be misleading: an empirical and simulation study. *Evolution* **54**: 1414–1422.
- BOWCOCK, A. M., A. RUIZ-LINARES, J. TOMFOHRDE, E. MINCH, J. R. KIDD *et al.*, 1994 High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**: 455–457.
- BRINKMANN, B., M. KLINTSCHAR, F. NEUHUBER, J. HUHNE and B. ROLF, 1998 Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.* **62**: 1408–1415.
- BRÜNNER, P. C., M. R. DOUGLAS and L. BERNATCHEZ, 1998 Microsatellite and mitochondrial DNA assessment of population structure and stocking effects in Arctic charr *Salvelinus alpinus* (Teleostei: Salmonidae) from central Alpine lakes. *Mol. Ecol.* **7**: 209–223.
- CALAFELL, F., A. SHUSTER, W. C. SPEED, J. R. KIDD and K. KIDD-KENNETH, 1998 Short tandem repeat polymorphism evolution in humans. *Eur. J. Hum. Genet.* **6**: 38–49.
- CARPENTER, J. M., 1992 Random cladistics. *Cladistics* **8**: 147–153.
- CAVALLI-SFORZA, L. L., and A. W. F. EDWARDS, 1967 Phylogenetic analysis: models and estimation procedures. *Am. J. Hum. Genet.* **19**: 233–257.
- COOPER, G., W. AMOS, R. BELLAMY, M. R. SIDDIQUI, A. FRODSHAM *et al.*, 1999 An empirical exploration of the  $(\delta\mu)^2$  genetic distance for 213 human microsatellite markers. *Am. J. Hum. Genet.* **65**: 1125–1133.
- CROZIER, R. H., B. KAUFMANN, M. E. CAREW and Y. C. CROZIER, 1999 Mutability of microsatellites developed for the ant *Camponotus sobrinus*. *Mol. Ecol.* **8**: 271–276.
- DI RIENZO, A., P. DONNELLY, C. TOOMAJIAN, B. SISK, A. HILL *et al.*, 1998 Heterogeneity of microsatellite mutations within and between loci and implications for human demographic histories. *Genetics* **148**: 1269–1284.
- ELLEGREN, H., 2000 Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet.* **16**: 551–558.
- ESTOUP, A., and B. ANGERS, 1998 Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations, pp. 55–86 in *Advances in Molecular Ecology*, edited by G. R. CAVALLI-SFORZA. IOS Press, Amsterdam.
- ESTOUP, A., C. TAILLIEZ, J. M. CORNUET and M. SOLIGNAC, 1995 Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.* **12**: 1074–1084.
- FAITH, D. P., and P. S. CRANSTON, 1991 Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. *Cladistics* **7**: 1–28.
- FELSENSTEIN, J., 1995 *Phylib (Phylogeny Inference Package) Version 3.572*. Department of Genetics, University of Washington, Seattle.
- GARZA, J. C., M. SLATKIN and N. B. FREIMER, 1995 Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.* **12**: 594–603.
- GOLDSTEIN, D. B., and C. SCHLÖTTERER (Editors), 1999 *Microsatellites: Evolution and Applications*. Oxford University Press, Oxford.
- GOLDSTEIN, D. B., A. RUIZ LINARES, L. L. CAVALLI-SFORZA and M. L. FELDMAN, 1995a Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* **92**: 6723–6727.
- GOLDSTEIN, D. B., A. RUIZ-LINARES, L. L. CAVALLI-SFORZA and M. W. FELDMAN, 1995b An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**: 463–471.
- GOLDSTEIN, D. B., G. W. ROEMER, D. A. SMITH, D. E. REICH, A. BERGMAN *et al.*, 1999 The use of microsatellite variation to infer population structure and demographic history in a natural model system. *Genetics* **151**: 797–801.
- GOODMAN, S. J., 1997 Rst Calc: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Mol. Ecol.* **6**: 881–885.
- GOODMAN, S. J., 1998 Patterns of extensive genetic differentiation and variation among European harbor seals (*Phoca vitulina vitulina*) revealed using microsatellite DNA polymorphisms. *Mol. Biol. Evol.* **15**: 104–118.
- HARR, B., B. ZANGERL, G. BREM and C. SCHLÖTTERER, 1998 Conservation of locus specific microsatellite variability across species: a comparison of two *Drosophila* sibling species *D. melanogaster* and *D. simulans*. *Mol. Biol. Evol.* **15**: 176–184.
- JIN, L., M. L. BASKETT, L. L. CAVALLI-SFORZA, L. A. ZHIVOTOVSKY, M. W. FELDMAN *et al.*, 2000 Microsatellite evolution in modern humans: a comparison of two data sets from the same populations. *Ann. Hum. Genet.* **64**: 117–134.
- KAYSER, M., L. ROEWER, M. HEDMAN, L. HENKE, J. HENKE *et al.*, 2000 Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am. J. Hum. Genet.* **66**: 1580–1588.
- KIMMEL, M., R. CHAKRABORTY, D. N. STIVERS and R. DEKA, 1996 Dynamics of repeat polymorphisms under a forward-backward mutation model: within- and between-population variability at microsatellite loci. *Genetics* **143**: 549–555.
- KOSKINEN, M. T., and C. R. PRIMMER, 2001 High throughput analysis of 17 microsatellite loci in grayling (*Thymallus* spp. Salmonidae). *Conserv. Genet.* **2**: 173–177.
- KOSKINEN, M. T., E. RANTA, J. PIIRONEN, A. VESELOV, S. TITOV *et al.*, 2000 Genetic lineages and postglacial colonization of grayling (*Thymallus thymallus*, Salmonidae) in Europe, as revealed by mitochondrial DNA analyses. *Mol. Ecol.* **9**: 1609–1624.
- KOSKINEN, M. T., J. NILSSON, A. JE. VESELOV, A. G. POTUTKIN, E. RANTA *et al.*, 2002 Microsatellite data resolve phylogeographic patterns in European grayling, *Thymallus thymallus*, Salmonidae. *Heredity* **88**: 391–401.
- LANDRY, P.-A., F.-J. LAPOINTE and J. A. W. KIRSCH, 1996 Estimating phylogenies from lacunose distance matrices: additive is superior to ultrametric estimation. *Mol. Biol. Evol.* **13**: 818–823.
- LAPOINTE, F.-J., 1998 How to validate phylogenetic trees: a stepwise procedure, pp. 71–88 in *Data Science, Classification and Related Methods*, edited by C. HAYASHI, N. OHSUMI, K. YAJIMA, Y. TANAKA, H.-H. BOCK *et al.* Springer, Tokyo.
- LEHMANN, T., W. A. HAWLEY and F. H. COLLINS, 1996 An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics* **144**: 1155–1163.
- NEI, M., F. TAJIMA and Y. TATENO, 1983 Accuracy of estimated phylogenetic trees from molecular data. *J. Mol. Evol.* **19**: 153–170.
- OHTA, T., and K. KIMURA, 1973 The model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genet. Res.* **22**: 201–204.
- ORWELL, G., 1945 *Animal Farm*. Secker & Warburg, London.
- PENNY, D., and M. D. HENDY, 1985 The use of tree comparison metrics. *Syst. Zool.* **34**: 75–82.
- PÉREZ-LEZAUN, A., F. CALAFELL, E. MATEU, D. COMAS, R. RUIZ-PACHECO *et al.*, 1997 Microsatellite variation and the differentiation of modern humans. *Hum. Genet.* **1**: 1–7.
- POLLOCK, D. D., A. BERGMAN, M. W. FELDMAN and D. B. GOLDSTEIN, 1998 Microsatellite behavior with range constraints: parameter estimation and improved distances for use in phylogenetic reconstruction. *Theor. Popul. Biol.* **53**: 256–271.
- PRIMMER, C. R., H. ELLEGREN, N. SAINO and A. P. MØLLER, 1996 Directional evolution in germline microsatellite mutations. *Nat. Genet.* **13**: 391–393.
- PRIMMER, C. R., N. SAINO, A. P. MØLLER and H. ELLEGREN, 1998 Unravelling the processes of microsatellite evolution through analysis of germline mutations in barn swallows *Hirundo rustica*. *Mol. Biol. Evol.* **15**: 1047–1054.
- RITZ, L. R., M.-L. GLOWATZKI-MULLIS, D. E. MACHIGH and C. GAILLARD, 2000 Phylogenetic analysis of the tribe Bovini using microsatellites. *Anim. Genet.* **31**: 178–185.
- ROBINSON, D. F., and L. R. FOULDS, 1981 Comparison of phylogenetic trees. *Math. Biosci.* **53**: 131–147.
- SANDERSON, M. J., 1995 Objections to bootstrapping phylogenies: a critique. *Syst. Biol.* **44**: 299–320.
- SCHLÖTTERER, C., 2000 Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**: 365–371.
- SCHLÖTTERER, C., R. RITTER, B. HARR and G. BREM, 1998 High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. *Mol. Biol. Evol.* **15**: 1269–1274.
- SHRIVER, M. D., L. JIN, R. CHAKRABORTY and E. BOERWINKLE, 1993 VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. *Genetics* **134**: 983–993.
- SHRIVER, M. D., L. JIN, E. BOERWINKLE, R. DEKA, R. E. FERRELL *et al.*, 1995 A novel measure of genetic distance for highly polymorphic tandem repeat loci. *Mol. Biol. Evol.* **12**: 914–920.
- SLATKIN, M., 1995 A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457–462.



- SOKAL, R. R., and F. J. ROHLF, 1995 *Biometry—The Principles and Practices of Statistics in Biological Research*, Ed. 3. W. H. Freeman, New York.
- STEEL, M. A., and D. PENNY, 1993 Distributions of tree comparison metrics—some new results. *Syst. Biol.* **42**: 126–141.
- TAKEZAKI, N., and M. NEI, 1996 Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**: 389–399.
- TESSIER, N., and L. BERNATCHEZ, 2000 A genetic assessment of single versus double origin of landlocked Atlantic salmon (*Salmo salar*) from Lake Saint-Jean, Québec, Canada. *Can. J. Fish. Aquat. Sci.* **57**: 797–804.
- VALDES, A. M., M. SLATKIN and N. B. FREIMER, 1993 Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* **133**: 739–749.
- VALSECCHI, E., P. PALSBOELL, P. HALE, D. GLOCKNER-FERRARI, M. FERRARI *et al.*, 1997 Microsatellite genetic distances between oceanic populations of the humpback whale (*Megaptera novaeangliae*). *Mol. Biol. Evol.* **14**: 335–362.
- WEBER, J. L., and C. WONG, 1993 Mutation of human short tandem repeats. *Hum. Mol. Genet.* **2**: 1123–1128.
- WEST, J. G., and D. P. FAITH, 1990 Data, methods and assumptions in phylogenetic inferences. *Aust. Syst. Bot.* **3**: 9–20.
- ZHIVOTOVSKY, L. A., 2001 Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. *Mol. Biol. Evol.* **18**: 700–709.
- ZHIVOTOVSKY, L. A., and M. W. FELDMAN, 1995 Microsatellite variability and genetic distances. *Proc. Natl. Acad. Sci. USA* **92**: 11549–11552.
- ZHIVOTOVSKY, L. A., L. BENNETT, A. M. BOWCOCK and M. W. FELDMAN, 2000 Human population expansion and microsatellite variation. *Mol. Biol. Evol.* **17**: 757–767.

Communicating editor: M. W. FELDMAN