MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 1083–1091 doi: 10.1111/j.1365-294X.2010.04996.x

OPINION

G'_{ST} and D do not replace F_{ST}

MICHAEL C. WHITLOCK

Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4

Abstract

The genetic differentiation among populations is affected by mutation as well as by migration, drift and selection. For loci with high mutation rates, such as microsatellites, the amount of mutation can influence the values of indices of differentiation such as G_{ST} and F_{ST}. For many purposes, this effect is undesirable, and as a result, new indices such as G'_{ST} and D have been proposed to measure population differentiation. This paper shows that these new indices are not effective measures of the causes or consequences of population structure. Both G'_{ST} and D depend heavily on mutation rate, but both are insensitive to any population genetic process when the mutation rate is high relative to the migration rate. Furthermore, D is specific to the locus being measured, and so little can be inferred about the population demography from D. However, at equilibrium, D may provide an index of whether a particular marker is more strongly affected by mutation than by migration. I argue that F_{ST} is a more important summary of the effects of population structure than D and that R_{ST} or other measures that explicitly account for the mutation process are much better than G_{ST}, G'_{ST}, or D for highly mutable markers. Markers with lower mutation rates will often be easier to interpret.

Keywords: diversity measures, F_{ST}, G_{ST}, G'_{ST}, Jost's D, population differentiation, population structure

Received 20 September 2010; revision received 15 November 2010; accepted 1 December 2010

F_{ST} and its related estimators have been used in at least three ways: to summarize genetic variation among populations for evolutionary models, to estimate demographic properties of populations and, using pairwise differences between populations, to predict which of the populations are most genetically unique, perhaps for conservation prioritization. For similar reasons, an alphabet soup of related indices have been developed that try to estimate genetic differentiation, including θ, G_{ST}, G'_{ST}, R_{ST}, ϕ_{ST}, D and many others. The intent of this article is to discuss the relative merits of some of these measures. In particular, I argue that both D (Jost 2008) and G'_{ST} (Hedrick 2005) tell us little about the evolutionary or demographic processes that lead to genetic variation among populations for most of the genome. When we are interested in measuring genetic variance among populations or identifying the degree of evolutionary independence of populations, these measures are poor substitutes for F_{ST} and other estimates like R_{ST} (Slatkin 1995) that explicitly account for the mutation process.

Correspondence: Michael C. Whitlock, Fax: 604 822 2416;
E-mail: whitlock@zoology.ubc.ca

The uses of F_{ST}

F_{ST}, as originally derived to describe the genetic patterns observed in a locus with two alleles, has two very useful interpretations. First, it is a standardized measure of the variance among demes of local allele frequency ($Var[p]$),

$$F_{ST} = \frac{Var[p]}{\bar{p}(1-\bar{p})},$$

where the standardization is via division by a function of the mean allele frequency, $\bar{p}$. Wright (1943) showed that this quantity would have the same expectation for all neutral alleles with low mutation rates. As such, F_{ST} can often be used together with the allele frequency of a locus to predict the distribution of allele frequencies across populations and therefore to understand evolution in structured populations. Second, Wright also

showed that F_{ST} can be interpreted as the proportional loss in heterozygosity at a locus caused by spatial population structure. The heterozygosity at a locus is reduced by a factor F_{ST} compared to what would be expected for a panmictic population with the same allele frequencies. In this way, F_{ST} can be used to predict the effects of population structure on the expression of alleles, which in turn can be used to predict the many aspects of evolution affected by changes in genotypes frequencies.

F_{ST} can be viewed as a description of the relative time to the most recent common ancestor for alleles chosen within and between populations (Slatkin 1995; Rousset 2004). F_{ST} defined by this coalescent approach captures the demographic history of the populations, but it will be unaffected by mutation. A coalescent F_{ST} is a common description of the average evolutionary history of all neutral loci. It describes the relative reduction in the time of divergence of a pair of alleles taken from the same population compared to two alleles from the population as a whole. The coalescent F_{ST}, which we will denote by $F_{ST,coal}$ here, is equal to

$$F_{ST,coal} = \frac{\bar{t} - \bar{t}_0}{\bar{t}},$$

where $\bar{t}$ is the mean time to the most recent common ancestor of two alleles chosen from the metapopulation as a whole and $\bar{t}_0$ is the same quantity for two alleles from the same population. $F_{ST,coal}$ measure the proportion of recent evolutionary history that is shared by alleles from the same population and as such provides a measure of the evolutionary isolation of populations. (These properties of $F_{ST,coal}$ do not depend on the island model or other assumptions about demography.) This is a very useful quantity because we will expect a roughly similar $F_{ST,coal}$ for all loci.

If the mutation process of marker alleles leaves a traceable history of the coancestry of alleles, then the coalescent F_{ST} can be inferred from data. For example, with sequence data, most mutations do not erase all evidence of old mutations and two alleles that differ by more mutations are likely to have been evolving independently for longer. For microsatellites, R_{ST} (Slatkin 1995) or ϕ_{ST} (Michalakis & Excoffier 1996) can implicitly use the similarity between alleles to infer the time to a common ancestor, so as to create an estimate of F_{ST} in the limit of no mutation. However, these measures are sensitive to deviations from the assumed mutation model.

A useful element of $F_{ST,coal}$ when it can be measured is that it increases monotonically with increasing isolation of populations. Alleles within populations that are not shared with other populations typically arose within that population, and they will coalesce within that population since the last migration event. As such, the coalescent F_{ST} also gives a good measure of the evolutionary uniqueness of separate populations.

Moreover, if genetic variation increases proportionally with time of divergence of alleles, then the coalescent F_{ST} allows a relative partitioning of the proportion of genetic variance that is between populations. The coalescent F_{ST} is particularly useful, because it provides an index of the pattern of genetic partitioning likely to describe most neutral loci with low mutation rates. Given that the mutation rates at most base pairs are quite small (in the order of 10^{-9} to 10^{-8}), for most species the coalescent F_{ST} will describe well the pattern of differentiation at most loci not under strong selection.

F_{ST} can be affected by all evolutionary forces. F_{ST} is increased by genetic drift within populations (either through persistent small population size, bottlenecks or founder effects) but reduced by migration of alleles among populations. Mutation can lower F_{ST}, especially because mutation can lead to high levels of heterozygosity both within and between populations. Mutation can also lower F_{ST} by making diverged populations more similar, if an allele present in one population is introduced by mutation into a diverged sister population (homoplasy), and homoplasy can also increase F_{ST} (Rousset 1996). Spatially heterogeneous selection on a locus increases F_{ST}, while spatially homogeneous selection can reduce F_{ST}.

Of these processes, mutation and selection tend to be idiosyncratic to the locus. Both selection and mutation vary widely from locus to locus, even within the same species. However, most of the genome experiences migration and drift roughly equally at all autosomal loci. Thus, the spatial genetic patterns at loci that are only strongly affected by migration and drift should have roughly similar F_{ST} for all loci, whereas loci with high mutation rates or strong selection may have F_{ST} values that differ greatly from other loci in the same genome. [Even for neutral alleles with low mutation, F_{ST} can vary stochastically among loci (Beaumont & Nichols 1996; Whitlock 2008), but the same expected value of F_{ST} applies to each such locus.]

The repeatability across the genome for neutral loci with low mutation rates is a powerful boon to evolutionary research. Given that F_{ST} (or similar indices) can now be measured easily from multiple loci in nearly all organisms, the repeatability across loci for F_{ST} makes is possible to establish a neutral baseline from which to infer selection at some loci (Beaumont & Balding 2004) or to measure demographic parameters. Migration rates and population sizes can be estimated from genetic data because multiple loci are expected to give similar information, allowing replication for such estimates. For example, F_{ST} has also been very useful in providing an

index of the relative isolation of populations. With certain strained assumptions, F_{ST} can give some idea of the amount of migration between populations (but see Whitlock & McCauley 1999). However, even when the demographic assumptions of the inference model are true, if genetic structure is strongly affected by mutation or selection, then such inference is impossible.

In all cases, the usefulness of F_{ST} hinges on two facts: that it can be empirically measured from genetic marker data and that a set of loci can be identified that give nearly independent replicate measures of the demographic history of the populations. It is very rare that we care about the differentiation of a particular locus, and even then we also need a baseline measure from other loci. For example, we need F_{ST} from many neutral loci to understand the evolution of the F_{ST} of a selected locus.

G_{ST} and θ

G_{ST} is an index similar to F_{ST}, derived explicitly to deal with multiple alleles at a locus (Nei 1973). Except for some details of the estimation process, G_{ST} is equivalent to F_{ST} (although not necessarily to $F_{ST,coal}$, unless the mutation rate is low). G_{ST} is defined in terms of H_S (the heterozygosity within populations) and H_T (the heterozygosity of the entire set of populations if forced to Hardy–Weinberg genotype frequencies with the observed total allele frequencies): $G_{ST} = (H_T - H_S)/H_T$. Many authors have noticed that when H_S is high, G_{ST} is constrained to a maximum value less than 1. For example, imagine the case when each population is completely isolated from all others, sharing no alleles with any other population. In this case, isolation is complete. However, as long as mutation continues, new alleles will appear within each local population, causing H_S to be greater than 0. As long as H_S is greater than 0, G_{ST} will be less than 1, sometimes much less. These new alleles will share coancestry with other alleles in the same population, but by mutation they will be genetically distinct.

G_{ST} (or F_{ST}) is decreased by high mutation rates (Wright 1943; Nagylaki 1998; Hedrick 1999; Balloux & Lugon-Moulin 2002). We can quantify the effect of mutation on G_{ST} from known theory in at least one special case, the finite island model with infinite-allele mutation. Let there be d populations in the system, all of equal effective size N, and each contributes and receives m migrants each generation, evenly divided to and from each of the other $d - 1$ populations. Mutation occurs at rate μ, and every new mutation is distinguishable from all previously existing alleles (the infinite-alleles model). Under this model, at equilibrium, assuming that both μ and m are small, G_{ST} for the island model is given by (Takahata 1983)

$$G_{ST} \approx \frac{1}{\left(\frac{d}{d-1}\right)^2 4Nm + \left(\frac{d}{d-1}\right)4N\mu + 1}.$$

With a large number of demes, this becomes (Maynard Smith 1970)

$$G_{ST} \approx \frac{1}{4Nm + 4N\mu + 1}.$$

So long as $m \gg \mu$, G_{ST} is not affected much by mutation, because the $N\mu$ term is much smaller than the Nm term. But if μ is greater than m, mutation dominates the equation and migration becomes unimportant for G_{ST}. The same issue occurs for Weir & Cockerham's (1984) θ. For the biallelic case of the island model, Wright (1943) showed that

$$F_{ST} \cong \frac{1}{4N(m + u + v) + 1},$$

where u and v are the forward and backward mutation rates. For the biallelic case, mutation in either direction reduces F_{ST}.

This effect of mutation on G_{ST} is unfortunate if we want a measure of differentiation that describes all neutral loci or if we want to estimate migration parameters. It is not a bias in estimating G_{ST} *per se*, because G_{ST} is truly affected by mutation as well as by other evolutionary processes. If one is interested in the variance among populations in allele frequency *including* the effects of mutation, G_{ST} and θ will both give an accurate portrayal of the standardized variance among populations for that locus. However, this is rarely if ever the goal of such measures. It would be useful to have a measure of the coalescent F_{ST} not affected by mutation, so that we could accomplish the goals of most F_{ST} studies. G_{ST} is an extremely bad measure of the evolutionary isolation of a population when the mutation rate is close to or greater than the migration rate. G_{ST} should certainly not be used when using microsatellite loci to study the relationship between subspecies, for example.

G'_{ST} and D

To adjust for the lower G_{ST} when H_S is high, Hedrick (2005) introduced G'_{ST}. G'_{ST} is defined by analogy to Lewontin's D' index of linkage disequilibrium. G'_{ST} is defined as G_{ST} divided by the maximum possible G_{ST} with the same overall allele frequencies, so that G'_{ST} always has a range from 0 to 1 independent of H_S:

$$G'_{ST} = \frac{G_{ST}}{\left(\frac{(d-1)(1-H_S)}{d-1+H_S}\right)}.$$

If H_S is high, then G'_{ST} can be much greater than G_{ST}. G'_{ST} is intended to be a standardized measure of G_{ST}, which accounts for different levels of total genetic variation at different loci. G'_{ST} has been used increasingly often in the literature.

A similar measure in increasing use is D (Jost 2008). Jost's D is an explicit measure of relative differentiation between populations. It is defined as

$$D = \frac{H_T - H_S}{1 - H_S}\left(\frac{d}{d-1}\right).$$

This D differs from F_{ST} in a fundamental definitional way: F_{ST} measures deviations from panmixia, while D measures deviations from total differentiation. As a result, their denominators differ, and thus, the two indices can behave quite differently. D indicates the proportion of allelic diversity that lies among populations, while F_{ST} is proportional to the variance of allele frequency among populations. D is more related to the genetic distance between populations than to the variance in allele frequencies; it may be preferable to call D a genetic distance measure.

D has several admirable properties: it goes to zero when all populations are identical and it goes to one as different populations become completely distinct. It monotonically increases with increasing divergence between populations. Unfortunately, D also has no direct relevance to the current evolution literature, in the sense that this quantity does not appear in any evolutionary theory in this form. As such we should, like Jost, consider D purely as a measure of differentiation, without the other intended purposes of F_{ST}. Jost (2009) shows that D is a measure of differentiation, but also warns that it ought not to be used for measuring migration.

Methods—the Finite Island Model

This paper will develop no new mathematics, but it will be useful to make some calculations based on theory to make the main points. Like all other papers that have discussed G'_{ST} and D (e.g. Jost 2008; Ryman & Leimar 2008, 2009), the calculations are based on the finite island model with infinite-alleles mutation. This mutation model is tractable (and has been used often in recent discussions about these various measures), but it is an incomplete choice for studying microsatellite alleles, which often mutate to previously existing alleles creating homoplasy.

Define f_0 and f_1 as the probabilities of identity in state of two alleles chosen at random from the same population or different populations, respectively. Recurrence equations for f_0 and f_1 can be found, similar to several previous authors (e.g. Rousset 2004), as

$$f'_0 = (1-\mu)^2\left[\frac{1}{2N} + \left(1-\frac{1}{2N}\right)\left((1-m)^2 + \frac{m^2}{(d-1)}\right)f_0 + \left(1-\frac{1}{2N}\right)\left(2m(1-m) + \frac{m^2(d-2)}{(d-1)}\right)f_1\right]$$

$$f'_1 = (1-\mu)^2\left[\left(\frac{2m(1-m)}{d-1} + \frac{m^2(d-2)}{(d-1)^2}\right)f_0 + \left(1 - \left(\frac{2m(1-m)}{d-1} + \frac{m^2(d-2)}{(d-1)^2}\right)\right)f_1\right]$$

assuming that the populations are censused immediately after drift and reproduction. These equations are easily solved for their equilibrium values, and the exact solutions at equilibrium are used in the figures. From these quantities, we can calculate $H_S = 1 - f_0$, the heterozygosity within a population and $H_T = 1 - \bar{f}$ as the heterozygosity of the pooled subpopulations, where $\bar{f} = (1/d)f_0 + (1 - 1/d)f_1$. Note that J_S from Jost (2008) is equal to f_0 and J_T is the same as $\bar{f}$. From these, G_{ST}, G'_{ST} and D can all be calculated from their usual formulas. $F_{ST,coal}$ can be calculated from the equations for G_{ST} by taking the limit as μ approaches 0. Throughout this paper, measurement error is ignored; we assume that the indices are estimated from an exhaustive sample of all individuals. The changes in various indices with changing parameters described in this paper do not reflect sampling problems, but instead reflect true differences between the indices. Meirmans & Hedrick (2011) discuss the sampling properties of these measures.

G_{ST} and F_{ST} when $\mu \ll m$

Figures 1 and 2 show the effects of varying either migration rate (Fig. 1) or mutation rate (Figs 1 and 2) on $F_{ST,coal}$, G_{ST}, G'_{ST} and D. These figures demonstrate many of the main points of this paper.

Let us first bring our attention to bear on the case when the mutation rate is much smaller than the migration rate (see the right-hand sides of the figures in Fig. 1 or the left-hand side of the figures in Fig. 2). As expected, when $\mu \ll m$, G_{ST} is an accurate measure of $F_{ST,coal}$. G_{ST} is not measurably affected by mutation when mutation rate is low relative to migration, as has been pointed out by several authors (Hedrick 1999, 2005; Ryman & Leimar 2008). When migration is common and mutation rare, there is no reason to make any adjustment to the usual measure of F_{ST} or G_{ST} for evolutionary inference. This result holds even for some high diversity systems; H_S can be high even with low $F_{ST,coal}$ if the local population size or the number of demes is high, yet G_{ST} will be affected by mutation only if $\mu \geq m$.

When $\mu \ll m$, both G'_{ST} and D can be quite high even when the relative differentiation of populations as

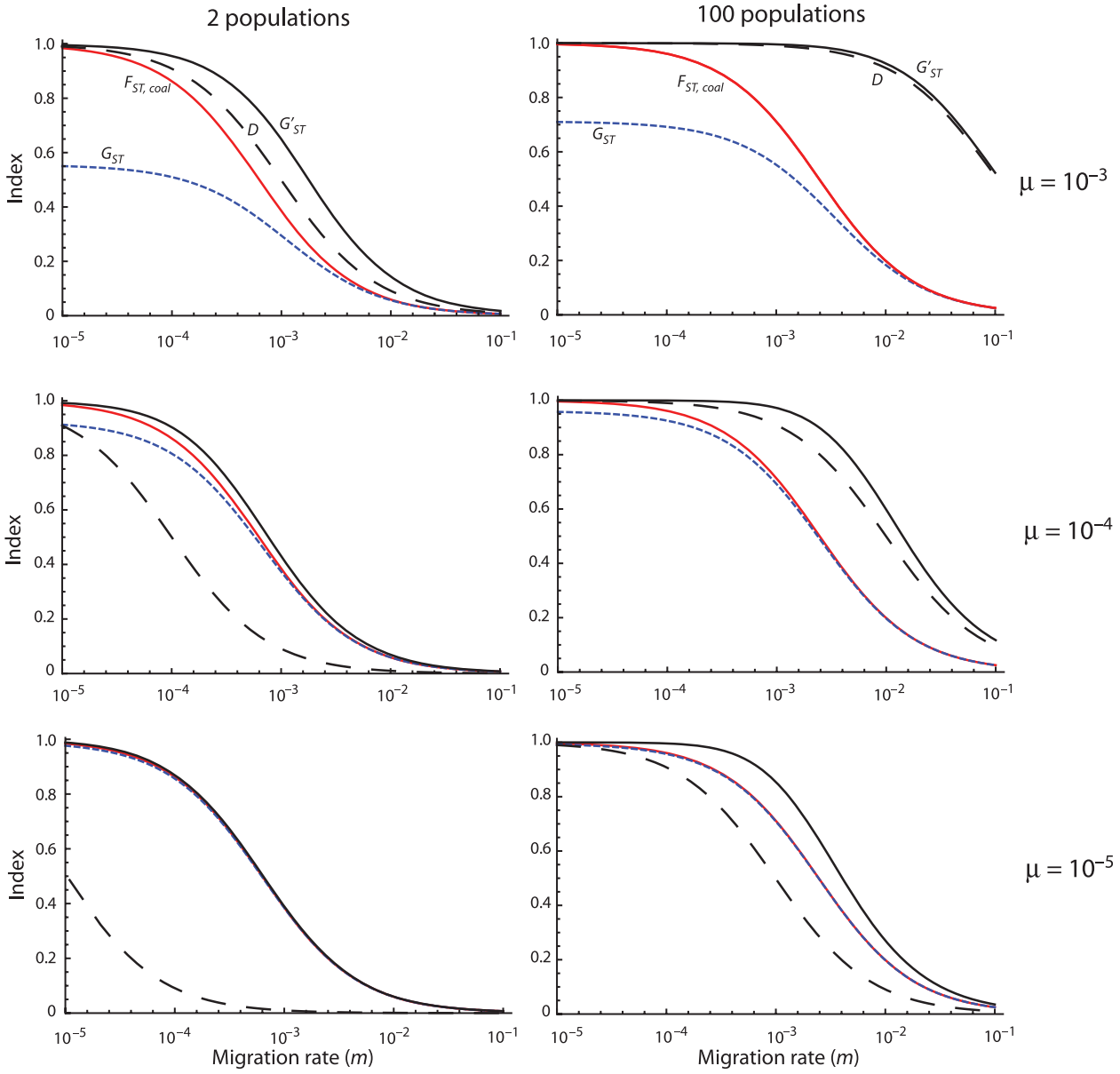

Fig. 1 $F_{ST,coal}$, G_{ST}, G'_{ST} and Jost's D over a range of migration rates. For all parameter values, G'_{ST} is equal to or larger than $F_{ST,coal}$ or G_{ST}; for high migration rates, G'_{ST} is much higher and presents a biased view of the population structure at a locus. For migration rates lower than the mutation rate, G_{ST} can be much smaller than $F_{ST,coal}$, but for migration rates greater than the mutation rate, G_{ST} and F_{ST} are approximately the same. Both G'_{ST} and D are insensitive to other parameters when the mutation rate and number of demes are high. For these graphs, the mutation rate was set to $\mu = 0.001$, 0.0001 or 0.00001 from the top row to the bottom row, with $N = 100$ individuals per population in an equilibrium finite island model. The left and right graphs show results for 2 and 100 populations per species, respectively. The line for G_{ST} is often obscured by the line for $F_{ST,coal}$, because they are nearly identical for low mutation rates. [Correction added after online publication 16 February 2011: in the second sentence 'migration rate' was corrected to 'mutation rate']

measured by $F_{ST,coal}$ is low (Figs 1 and 2, right). In fact, both G'_{ST} and D can approach one even with reasonably high migration and low relative genetic differentiation (Fig. 1, right top panel). Certainly, neither of these indices can be relied upon as a 'correction' for G_{ST} with high local diversity, and high G'_{ST} or D values may be difficult to interpret in terms of the evolutionary isolation of the populations.

G_{ST} and F_{ST} when $\mu > m$

When mutation is common relative to migration, as will be the case for microsatellite loci and relatively isolated populations, G_{ST} is much lower than $F_{ST,coal}$, as has been frequently noted earlier (see the left-hand sides of the top panels in Fig. 1 or the right-hand side of both panels in Fig. 2; see also Ryman & Leimar 2009). However,

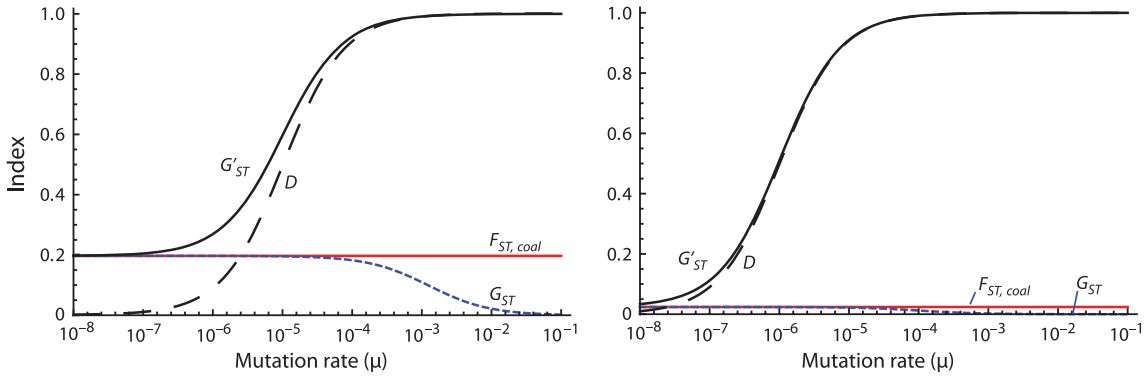

Fig. 2 The effect of varying mutation rate on the differentiation indices. Both G'_{ST} and D are heavily dependent on the mutation rate and asymptote to high values when the mutation rates approach those observed at microsatellite loci, for cases when the number of individuals in the species is large. G_{ST} and $F_{ST,coal}$ are also shown for comparison. These calculations are based on a finite island model with 100 populations at equilibrium, with (left panel) 1000 diploid individuals per deme and a migration rate of 0.001 or (right panel) 100 000 diploid individuals per deme and a migration rate of 0.0001.

neither G'_{ST} nor D 'corrects' for this effect: neither is an estimate of $F_{ST,coal}$, and both are likely to be quite high over many possible parameter combinations. As the right panel of Fig. 2 shows, both G'_{ST} and D can be quite large even when the coalescent F_{ST} is relatively small, when mutation is stronger than migration and the local population size is large.

When $\mu > m$, it would be better to estimate $F_{ST,coal}$ using more complete information about the mutation process. For example, R_{ST} and ϕ_{ST} can correctly estimate F_{ST} for microsatellites (Kronholm *et al.* 2010; Meirmans & Hedrick 2011), although such estimation becomes somewhat biased if a high fraction of mutations have large effects (Balloux *et al.* 2000). Perhaps even better, markers with low mutation rates would need no adjustment. Coalescent-based models can also be used to estimate N_em and $N_e\mu$ from DNA sequence or microsatellite data (e.g. migrate, Beerli & Felsenstein 2001).

In contrast, how to interpret G'_{ST} and D is unclear in this parameter range. Contra Jost (2009), migration parameters cannot be estimated from G_{ST} or G'_{ST} when $\mu > m$, because the value of G_{ST} is determined by mutation much more than by migration. In principle, we could estimate μ to correct for its effects, but in practice when $\mu > m$, the measurement uncertainty of μ far outweighs the signal from m. Therefore, any resulting 'corrected' estimate would have little value. High G'_{ST} or D carry little information other than there are a large number of alleles in the system. With high mutation rates, private alleles appear rapidly before migration can move them to other populations, even in a reasonably well-mixed species.

Problems with D and G'_{ST}

Both D and G'_{ST} were originally defined to address the errors of interpretation of G_{ST} caused by high diversity within populations. D in particular was also derived as a multiplicative diversity measure (Jost 2007, 2008). However, both D and G'_{ST} share a basic difficulty, which is that neither estimate quantities that otherwise appear in the evolutionary biology literature. As Ricotta points out with respect to Jost's (2007) ecological diversity index, 'the key question we should ask of a beta measure is: does it measure the thing we are biologically interested in?' (Ricotta 2010, p. 1982). In contrast, F_{ST} (and by extension the other quantities that are meant to approximate it) appears naturally in the theoretical literature of evolutionary genetics (e.g. Whitlock 2002, 2003; Rousset 2004); F_{ST} is a quantity that we have reason to need to know, in large part because it is expected to be similar from locus to locus. We may have imperfect statistical estimators of this quantity (such as G_{ST} for loci with high mutation rates), but the underlying quantity is inherently interesting, and these biases can be addressed with other techniques. Without such a literature showing the value and means of interpreting G'_{ST} and D, there is little reason to measure them. D is a good measure of genetic differentiation among populations on its own terms; the issue is whether that measure can be interpreted to tell us something we want to know about the state or evolution of a set of populations. This remains to be proven.

D, at least in the case of the finite island model, turns out to be relatively insensitive to important evolutionary processes affecting population differentiation. For the finite island model at equilibrium as defined in the methods section, it can be shown that the value of Jost's D is

$$D = \frac{(d-1)^2\mu(2-\mu)}{2(d-1)m(1-\mu)^2 - dm^2(1-\mu)^2 + (d-1)^2\mu(2-\mu)}.$$

When $\mu \ll 1$ and $m \ll 1$, this simplifies to (Jost 2008)

$$D \cong \frac{1}{1+\frac{m}{\mu(d-1)}}.$$

Note that D is not a function of the local population size N and so does not describe the effects of local drift, the primary cause of among-population variance of neutral allele frequencies. D is, to a good approximation, a function only of the ratio m/μ $(d-1)$, for the finite island case. In other words, for the finite island model, D is a measure of the relative likelihood of mutation or migration between a specific pair of demes.

One stated impetus for the derivation of D is the difficulty of interpreting G_{ST} when H_S is high. As we have seen, this occurs when $m < \mu$; therefore, it is interesting to look at the behaviour of D in this part of the parameter space. When $m \ll \mu$, however, D is always approximately 1. (In reality, with the homoplasy caused by repeated mutation to the same allele in different populations, D would be lower than one, even with completely isolated populations.) In this important case, D is actually not sensitive to any aspect of the population demographic parameters (Fig. 3). However, a high D does tell us that the mutation rate is as high or higher than the pairwise migration rate. As such, D could be used to tell us when a particular molecular marker has a mutation rate that is too high to be used to infer the general properties of most of the genome. If $D/(d-1)$ is > 0.1 for a locus, then the G_{ST} or θ of that locus cannot be relied on to indicate the demographic patterns in the metapopulation.

When $\mu\, d \ll m$, D is approximately

$$D \cong \frac{(d-1)\mu}{m},$$

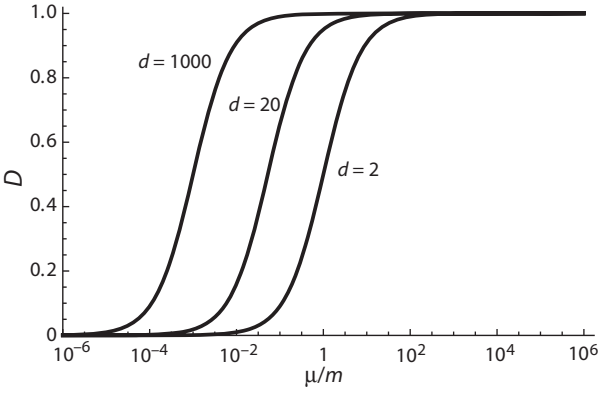

Fig. 3 Jost's D varies as a function of mutation rate divided by the migration rate. If the mutation rate is greater than the migration rate (so $\mu/m > 1$), D is approximately 1, independent of any other information, with the infinite-alleles model of mutation. For lower mutation rates, D varies in proportion to μ, so that D will vary greatly among loci. [Correction added after online publication 16 February 2011: in the second sentence 'mutation rate' was corrected to 'migration rate'.]

as first derived by Jost (2008). Therefore, when migration rates are relatively high, D will vary in proportion to μ. Thus, D is not expected to be the same for all loci, but it will vary as greatly as the mutation rate does from locus to locus. Even among microsatellite loci, there is great variation in μ, from 10^{-6} to 10^{-2} (Ellegren 2000; Whittaker *et al.* 2003); the range is even greater over all types of molecular markers. D therefore does not provide information about populations but only about specific loci (see also Ryman & Leimar 2009). Without a specific reason to care about the D of a particular locus, it is unclear what use this index might have. One cannot apply D to a convenient neutral marker locus to learn much about the population as a whole, and it is yet to be shown that D is a useful measure even for loci of specific interest. D measured even from such an important locus would be informative only about that locus, not the state of the genome as a whole.

D is derived in the context of Jost's (2007) ecological measure of 'true' diversity. While many of the principles carry over between ecology and population genetics, there are certain crucial differences. In measuring ecological diversity, the units are species, and there is direct intellectual and practical value in measuring the distribution of species within and among communities. While there is unnecessary danger in throwing away the knowledge that can be gained from other ecological diversity measures, the type of beta diversity proposed by Jost (2007) is clearly valuable. In contrast, with population genetic markers, the loci and the alleles that are used in the studies are normally not of direct interest. Understanding the number of alleles at a particular locus may be interesting as an intellectual exercise, but it is rarely the goal of such studies. Indeed, it is commonly the case that the microsatellite loci chosen for inclusion in a study have been selected by the researcher precisely because they have high allelic diversity, meaning that any measure that depends on the absolute diversity is even more unlikely to generalize to other parts of the genome. The principles behind D are sound, but it has not been proven to measure a quantity that is useful for evolutionary or conservation studies when applied to the types of data for which it has been proposed. Such uses may yet be proposed and defended, but until such uses are confirmed we ought not give much prominence to D.

Conclusions

Measuring genetic differentiation between populations can be problematic. For most common purposes, we would like a measure that gives similar results for all neutral loci, because the results would then reflect the properties of the populations rather than of the loci themselves. Fortunately F_{ST} and G_{ST} are, at least for the

finite island model considered here, good measures of the coalescent F_{ST} when the mutation rate is small relative to the migration rate. (We ignore here the long-standing discussion about the statistical properties of estimators of these quantities, focusing on their parametric values.) For F_{ST} or G_{ST} calculated from low mutation markers, we will get a much better estimate of the intended quantity without any adjustment like G'_{ST} or D.

Unfortunately, G_{ST} can be sensitive to mutation rate variation among loci, at least when the mutation rate approaches or exceeds the migration rate. This leads to measures of G_{ST} for some loci, particularly microsatellite loci, with very low values compared to other markers. Thus, G_{ST} will be confusing if used for high mutation rate markers taken from populations with low migration rates. However, even for loci with high mutation rates, the coalescent F_{ST} can be estimated if the mutation process is accounted for explicitly, such as with R_{ST} or ϕ_{ST}. Measures of this sort that explicitly account for the microsatellite mutation process should be preferred to analyse these sorts of data.

G'_{ST} and D have also been used increasingly often since their recent derivation as an explicit response to the behaviour of G_{ST} measured with loci with high mutation rates. The results in this paper cast doubt on the usefulness of these measures. Neither G'_{ST} nor D is an estimate of the coalescent F_{ST} except in the low mutation case when G_{ST} provides an excellent measure for $F_{ST,coal}$ anyway; neither is a 'correction' that can be interpreted like $F_{ST,coal}$. (It should be noted that Jost does not intend D to be a corrected estimate of F_{ST}, but others have interpreted it in this way.) Moreover, both G'_{ST} and D are very sensitive to cases when the mutation rate is greater than the migration rate; both are likely to be close to one for any locus with a high mutation rate, more or less regardless of the values of the other parameters. Thus, in the very parameter range that motivates their use, these measures are largely insensitive to variation in the evolutionary process.

D is intended to measure the genetic differentiation among populations (Jost 2008) in way that increases with greater numbers of alleles unique to local populations. For this purpose, it performs well—for the specific locus being measured. Without precise knowledge of mutation rates, however, D conveys little information about the population as a whole.

It is important to remember that the calculations in this paper (or indeed in other recent explorations of these indices) are based on the infinite-allele model, which is not a perfect description of the evolution of microsatellites. Microsatellites are likely to show repeated mutation to the same allele, causing a confusing homoplasy when these mutations occur in different populations. The fact that real D values rarely reach one is likely due to the effects of this homoplasy. More importantly, homoplasy provides an alternate reason for lower than expected values of genetic differentiation between isolated populations.

Neither G'_{ST} nor D is explicitly motivated by a particular evolutionary process; for neither is there a body of work defining any particular known use of the measure. In this absence, the value of these new measures is unclear. D is, for the finite island model at least, simply a function of the ratio of mutation rate and migration rate. When H_S is high because of high mutation rates, there is almost no power of D or G'_{ST} to distinguish relatively well-mixed populations from those with very low migration rates, especially when there are a large number of populations.

It is certainly possible that uses for D can be derived in the future. Such proposals, of course, should be evaluated on their ability to generate useful information. For example, with the island model at equilibrium, D provides an index of whether a particular locus is more affected by mutation than by migration. Inference from such loci should be treated with caution. For many purposes, microsatellite loci with high repeat numbers may prove to be challenging to interpret. SNPs may prove more reliable for these reasons.

Another possible use is that D could be used to roughly provide bounds in the order of the migration rate. If D is large and an equilibrium finite model can be assumed, then we might infer that the migration rate between a pair of populations $[m/(d-1)]$ is lower than the mutation rate. We have a rough idea of the probable mutation rates at many loci, so we may be able in such cases to infer whether migration is greater or less than the estimated mutation rate. For example, a high D value with a neutral microsatellite locus would suggest that migration between pairs of populations is <1 in a thousand (although it would convey little about the total migration unless the number of populations was well known). Such interpretations rely heavily on equilibrium island model assumptions, however, and D is slow to reach equilibrium (Ryman & Leimar 2009).

D has been explicitly derived as an aid to making conservation decisions, but as yet no algorithm has been proposed to rank populations according to D. Such an algorithm would presumably require an as-yet-underived version of D that was population specific. More importantly, if such an algorithm was applied to a set of loci, it would probably only reflect the effects of population structure on diversity for those loci alone. D based on microsatellite loci (the common domain of G_{ST} currently) is very unlikely to have any useful conservation value. Although D was defined initially in the context of discussion of microsatellite loci (Jost 2008), almost no value can be derived from D of microsatellite

data. Until such uses are made explicit and shown to provide important information for conservation purposes, D should be viewed with caution by the conservation community.

Empirical papers should continue to report F_{ST} and G_{ST}. If the work uses microsatellites and the likely migration rate is less than 0.01, it seems wise to focus the analysis on the coalescent approach by Slatkin (1995) or Michalakis & Excoffier (1996). More reliable inference may be possible by using markers with lower mutation rates, such as SNPs or microsatellites with low repeat numbers [which are likely to have much lower mutation rates (Whittaker *et al.* 2003)]. Until a useful evolutionary interpretation of G'_{ST} or D is provided, their use should be limited, although it is always a good idea to report useful quantities like H_S that can be used for various ways (e.g. allowing calculation of G'_{ST} or D if a reader so desires). More than one summary measure of differentiation will have value; we should endeavour to use the summaries that best convey the answers to the biological questions that motivate our studies. For most within-species purposes, this is likely to be F_{ST}.

Acknowledgements

This contribution has benefited greatly from the comments of Phil Hedrick, Lou Jost, Nolan Kane, Ilkka Kronholm, Patrick Meirmans, Sally Otto, Loren Rieseberg, Mark Vellend and anonymous reviewers. This work was funded by a Discovery Grant from the Natural Science and Engineering Research Council (Canada).

References

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155–168.

Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J (2000) Microsatellites can be misleading: an empirical and simulation study. *Evolution*, **54**, 1414–1422.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, Series B*, **263**, 1619–1626.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 4563–4568.

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*, **16**, 551–558.

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.

Jost L (2007) Partitioning diversity into independent alpha and beta components. *Ecology*, **88**, 2427–2439.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Jost L (2009) Reply: D vs G'_{ST}: response to Heller and Siegismund (2009) and Ryman and Leimar (2009). *Molecular Ecology*, **18**, 2088–2091.

Kronholm I, Loudet O, de Meaux J (2010) Influence of mutation rate on estimators of genetic differentiation—lessons from *Arabidopsis thaliana*. *BMC Genetics*, **11**, 33.

Maynard Smith J (1970) Population size, polymorphism, and the rate of non-Darwinian evolution. *American Naturalist*, **104**, 231–237.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.

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

Nagylaki T (1998) Fixation indices in subdivided populations. *Genetics*, **148**, 1325–1332.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, **70**, 3321–3323.

Ricotta C (2010) On beta diversity decomposition: trouble shared is not trouble halved. *Ecology*, **91**, 1981–1983.

Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*, **142**, 1357–1362.

Rousset F (2004) *Genetic Structure and Selection in Subdivided Populations*. Princeton, Princeton, NJ.

Ryman N, Leimar O (2008) Effect of mutation in genetic differentiation among nonequilibrium populations. *Evolution*, **62**, 2250–2259.

Ryman N, Leimar O (2009) G_{ST} is still a useful measure of genetic differentiation—a comment on Jost's D. *Molecular Ecology*, **18**, 2084–2087.

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

Takahata N (1983) Gene identity and genetic differentiation of population in the finite island model. *Genetics*, **104**, 487–512.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Whitlock MC (2002) Selection, load, and inbreeding depression in a large metapopulation. *Genetics*, **160**, 1191–1202.

Whitlock MC (2003) Fixation probability and time in a metapopulation. *Genetics*, **164**, 767–779.

Whitlock MC (2008) Evolutionary inference from Q_{ST}. *Molecular Ecology*, **17**, 1885–1896.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: FST doesn't equal 1/(4Nm+1). *Heredity*, **82**, 117–125.

Whittaker JC, Harbord RM, Boxall N, Mackay I, Dawson G, Sibly RM (2003) Likelihood-based estimation of microsatellite mutation rates. *Genetics*, **164**, 781–787.

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