

OPINION

If F_{ST} does not measure neutral genetic differentiation, then comparing it with Q_{ST} is misleading. Or is it?

PIM EDELAAR* and MATS BJÖRKLUND†

*Department of Conservation Biology, Estación Biológica de Doñana—CSIC, Av. Americo Vespucio s/n, E-41092 Sevilla, Spain,

†Department of Animal Ecology, Evolutionary Biology Centre, University of Uppsala, Norbyvägen 18D, SE-75236 Uppsala, Sweden

Abstract

The comparison between neutral genetic differentiation (F_{ST}) and quantitative genetic differentiation (Q_{ST}) is commonly used to test for signatures of selection in population divergence. However, there is an ongoing discussion about what F_{ST} actually measures, even resulting in some alternative metrics to express neutral genetic differentiation. If there is a problem with F_{ST} , this could have repercussions for its comparison with Q_{ST} as well. We show that as the mutation rate of the neutral marker increases, F_{ST} decreases: a higher within-population heterozygosity (H_e) yields a lower F_{ST} value. However, the same is true for Q_{ST} : a higher mutation rate for the underlying QTL also results in a lower Q_{ST} estimate. The effect of mutation rate is equivalent in Q_{ST} and F_{ST} . Hence, the comparison between Q_{ST} and F_{ST} remains valid, if one uses neutral markers whose mutation rates are not too high compared to those of quantitative traits. Usage of highly variable neutral markers such as hypervariable microsatellites can lead to serious biases and the incorrect inference that divergent selection has acted on populations. Much of the discussion on F_{ST} seems to stem from the misunderstanding that it measures the differentiation of populations, whereas it actually measures the fixation of alleles. In their capacity as measures of population differentiation, Hedrick's G'_{ST} and Jost's D reach their maximum value of 1 when populations do not share alleles even when there remains variation within populations, which invalidates them for comparisons with Q_{ST} .

Keywords: adaptive genetic differentiation, F_{ST} , Hedrick's G'_{ST} , Jost's D , neutral genetic markers, Q_{ST}

Received 29 September 2010; revision received 18 January 2011; accepted 31 January 2011

Introduction

A major topic in evolutionary biology and conservation genetics is to determine the genetic structure of populations and to elucidate which processes have led to the observed structures. When the focus is on neutral genetic markers, population differentiation is commonly expressed as F_{ST} , G_{ST} or related measures (from now on we use F_{ST} and G_{ST} interchangeably, because for our purposes here, they are equivalent). F_{ST} is a parameter that expresses the probability that two alleles drawn at random in a subpopulation are identical by descent compared to two alleles drawn at random in the whole

population. Thus, a positive F_{ST} means that the average relatedness among individuals in the same subpopulation is higher than among individuals from different subpopulations, i.e. that there is a population genetic structure (Holsinger & Weir 2009).

The same principle applies to the loci that influence the expression of quantitative traits. Hence, if these quantitative trait loci (QTL—here to mean the functional loci, not tightly linked noncoding loci) are exposed only to the same neutral evolutionary forces as noncoding neutral genetic markers, then their expected level of differentiation among populations should be equal to that of neutral loci. Under the assumption that the effects of these QTL on a quantitative trait are additive and equal, then quantitative genetic divergence (expressed as Q_{ST}) of traits only exposed to neutral

Correspondence: Pim Edelaar, Fax: (+34) 954 621 125; E-mail: edelaar@ebd.csic.es

evolution equals F_{ST} for neutral genetic markers (Lande 1992; Spitze 1993; Whitlock 1999; McKay & Latta 2002). However, if divergent or convergent selection also acts on populations, divergence in quantitative genetic traits under selection will differ from divergence in neutral genetic markers. Thus, the comparison of quantitative genetic divergence (Q_{ST}) with neutral genetic divergence (F_{ST}) can inform us on the relative roles of selection and neutral evolution in driving patterns of population differentiation. When $Q_{ST} > F_{ST}$, divergent selection is inferred to have acted on the trait of interest; when $Q_{ST} < F_{ST}$, convergent (stabilizing) selection is inferred; and when $Q_{ST} = F_{ST}$, selection may have operated but no other process than divergence due to genetic drift needs to be invoked to explain the population genetic structure (Spitze 1993; Leinonen *et al.* 2008; Whitlock 2008).

Compared with other methodological approaches to test for divergent or convergent selection, the comparison of Q_{ST} and F_{ST} can be applied to a large number of populations and traits, does not require the collection of fitness measures in the wild and can detect the cumulative effect of past selection even in the absence of current selection. It therefore is a good method for screening for signatures of historical selection in a biological system (Whitlock 2008). Since its inception in the early 1990s, it seems to have become firmly established in the toolbox of evolutionary biologists (Leinonen *et al.* 2008).

Recently, it has been argued that F_{ST} might not be an accurate measure of neutral genetic differentiation (Hedrick 2005; Meirmans 2006; Jost 2008; Kronholm *et al.* 2010; Meirmans & Hedrick 2010). Later we discuss these criticisms in the context of comparisons between Q_{ST} and F_{ST} .

A brief overview of Q_{ST} and F_{ST}

Q_{ST} is calculated as the ratio of additive genetic variance among populations to total additive genetic variation: $Q_{ST} = V_{AMONG}/V_{TOTAL}$ (see Whitlock 2008 and Leinonen *et al.* 2008 for more details and an overview of how to estimate Q_{ST} properly). As $V_{AMONG} = V_{TOTAL} - V_{WITHIN}$, we can rewrite this as $Q_{ST} = (V_{TOTAL} - V_{WITHIN})/V_{TOTAL}$, which can range from 0 (no differentiation, $V_{TOTAL} = V_{WITHIN}$) towards 1 (complete differentiation, $V_{TOTAL} \gg V_{WITHIN}$). These variance components can be derived from results of controlled breeding designs, or using animal model methodology when interbreeding between populations is uncontrolled, such as in the wild (Pujol *et al.* 2008). In practice, when two populations overlap completely, Q_{ST} equals zero. When Q_{ST} is 0.50, overlap is moderate. Yet when population means are far apart and overlap is in practice (with limited sample sizes) nonexistent, Q_{ST} is

still smaller than 1. This is because the within-population variance component still has a significant weight in the ratio of $(V_{TOTAL} - V_{WITHIN})/V_{TOTAL}$. A Q_{ST} of exactly 1 could only be reached when populations are fixed for alternative genotypes, so when $V_{WITHIN} = 0$.

F_{ST} can be calculated from the distribution of variation in neutral markers within and among populations (estimation techniques reviewed in Holsinger & Weir 2009). In its simplest form, F_{ST} is estimated using heterozygosity: $F_{ST} = (H_{TOTAL} - H_{WITHIN})/H_{TOTAL}$ (Nei 1973), where H_{TOTAL} is the total expected heterozygosity based on the observed allele frequencies of the pooled populations, and H_{WITHIN} is the average expected within-population heterozygosity based on the allele frequencies of each population. With this notation, the parallel between the formulas for Q_{ST} and F_{ST} in comparing among-population variation to total variation is obvious. Note that H_{WITHIN} is normally referred to as H_S , the *subpopulation* heterozygosity, but from now on we use the subscript *within* and the term *population* for the lowest level to facilitate comparison with Q_{ST} notation. In theory, the range of F_{ST} is from 0 (no differentiation) to 1 (complete differentiation), just like Q_{ST} .

Issues with F_{ST} and its comparison with Q_{ST}

Recently, it has been argued that for highly variable markers, such as microsatellites, F_{ST} might not be an accurate measure of neutral genetic differentiation (Hedrick 2005; Meirmans 2006; Jost 2008; Kronholm *et al.* 2010; Meirmans & Hedrick 2010). A more variable marker (in the sense of having more relatively common alleles) will have a higher H_{WITHIN} , because homozygotes will be more rare. As both H_{WITHIN} and H_{TOTAL} have a theoretical maximum of 1, $H_{TOTAL} - H_{WITHIN}$ will tend to zero as a marker is more and more variable. This in turn implies that F_{ST} will necessarily approach zero for more variable markers. Jost (2008), Heller & Siegismund (2009), Gerlach *et al.* (2010), Meirmans & Hedrick (2010) and Whitlock (2011) give numerical examples and empirical data that provide compelling evidence for this relationship. Hence, a highly variable microsatellite can never yield a high F_{ST} if many alleles are common, even when populations do not share a single allele and are in that sense completely genetically different. In fact, F_{ST} has a theoretical maximum of $1 - H_{WITHIN}$ (Hedrick 2005; Meirmans & Hedrick 2010). (Note that the lower values of F_{ST} for more variable markers are independent of the method used to estimate F_{ST} , be it method-of moments, maximum likelihood or Bayesian: Holsinger & Weir 2009; Meirmans & Hedrick 2010).

This has a problematic effect on the comparison between Q_{ST} and F_{ST} . If the estimate of F_{ST} depends on

the variability of the neutral genetic marker used, and some types of markers are typically more variable than others, then the difference between Q_{ST} and F_{ST} also depends on the type of marker used (Hendry 2002). In fact, interpretations are biased toward inferring divergent selection on the trait of interest when more variable markers are used, as is commonly performed nowadays.

On the other hand, it follows from the definition of F_{ST} that we *should* obtain lower values for F_{ST} for more variable markers: if there are many alleles in each population, then two randomly drawn alleles are unlikely to be identical by descent. This creates the apparent paradoxical situation that the expectation that $Q_{ST} = F_{ST}$ under neutrality cannot be true for all types of markers if F_{ST} decreases with marker variability, but on the other hand, F_{ST} should decrease for more variable markers.

Comparing Q_{ST} with F_{ST} : which markers to use?

Marker variability is related to mutation rate, and both F_{ST} and Q_{ST} have been suggested to depend on mutation rate (e.g. Kronholm *et al.* 2010). Therefore, a solution to this apparent paradox might be to compare Q_{ST} and F_{ST} only when neutral markers have the same mutation rates as the underlying QTL of quantitative traits. In that case neutral loci and QTL will have the same marker variability and will show the same rate of neutral divergence when populations are not exposed to selection, and $Q_{ST} = F_{ST}$ (Hendry 2002). To test this claim, we modelled the neutral divergence of a quantitative trait composed of 10 QTL versus the divergence of 10 neutral loci in exactly the same way (contra Kronholm *et al.* 2010), using the program quantiNemo (Neuenschwander *et al.* 2008). The meta-population was composed of 25 populations consisting of 1000 individuals each, which were connected by a migration rate of 10^{-3} according to the island model. We systematically varied the mutation rate of the QTL and neutral loci from 10^{-8} to 10^{-3} . Each locus consisted of a maximum of 10 alleles, and simulations were started with all loci maximally polymorphic to reach equilibrium sooner. We recorded Q_{ST} and F_{ST} after 5000 generations (longer runs showed that populations had reached equilibrium for Q_{ST} and F_{ST} under all mutation rates) and ran 10 replicates for each mutation rate. The results (Fig. 1) clearly show that F_{ST} decreases when mutation rates of neutral markers go up. However, they also show that Q_{ST} equally decreases when mutation rates of QTL go up. When mutation rates for QTL and neutral markers are equal, Q_{ST} and F_{ST} yield similar values. Hence, when neutral loci and QTL have the same mutation rate (i.e. the same marker variability)

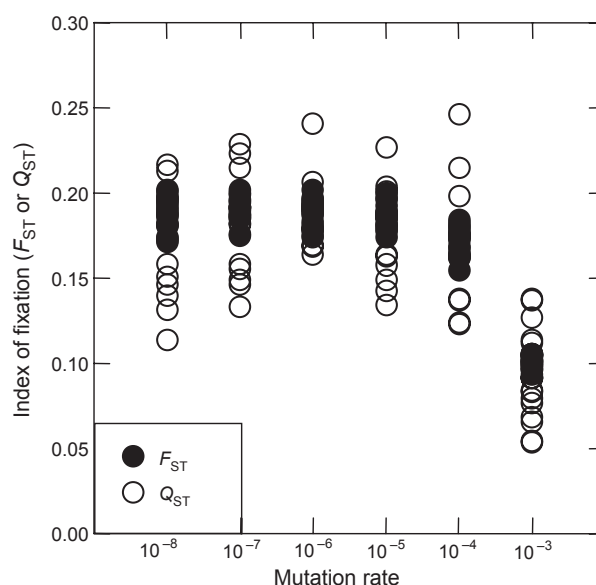


Fig. 1 F_{ST} and Q_{ST} decrease in the same way as the mutation rates of underlying loci become high. We varied the mutation rate of 10 neutral loci and 10 QTL underlying a quantitative trait, within a meta-population connected by a low level of migration according to the island model, and recorded F_{ST} and Q_{ST} at equilibrium in 10 replicates. See text for more details.

they will show the same rate of neutral divergence when populations are not exposed to selection, and so there is no paradox. Under all mutation rates $Q_{ST} = F_{ST}$, but the absolute values will be lower when mutation rate is high. (Note that Q_{ST} varies much more over replicates than F_{ST} , as discussed before by O'Hara & Merilä 2005 and Whitlock 2008).

Hence, the neutral marker used in the comparison of Q_{ST} and F_{ST} should preferably have a mutation rate similar to that of the investigated quantitative trait. Unfortunately, the mutational dynamics of QTL are not well known. The overall mutation rate in eukaryotes (including humans) is approximately 10^{-8} per base pair per generation (Drake *et al.* 1998; Nachman & Crowell 2000; Roach *et al.* 2010), but it may vary across the genome (Ellegren *et al.* 2003). Apparently, such variation occurs on different scales, including sequence context effects (the influence of neighbouring nucleotides on the rate of mutation at individual sites), variation within chromosomes (on the scales of kilobases as well as megabases) and between chromosomes (among autosomes as well as between autosomes and sex chromosomes). In addition, part of these mutations will not have functional effects because they are synonymous in their coding for amino acids, so this low mutation rate may even be an overestimate of the phenotypic mutation rate for QTL. Besides point mutations of base pairs, functional mutations may also be attributed to other

kinds of mutation mechanisms such as insertions and deletions (and any corresponding frame shifts), trinucleotide microsatellites coding for certain amino acids, gene duplications, etc. The evaluation of the different mutation mechanisms and their corresponding mutation rates is an active research field, and a full review is beyond this note. However, if we do not know the mutation rate of the quantitative genetic trait we are interested in, then how do we choose which marker to use to estimate F_{ST} ?

Fortunately, we may not need to know the exact mutation rate of the quantitative genetic trait under study. In our simulations (Fig. 1), Q_{ST} and F_{ST} seem to be rather constant at mutation rates lower than 10^{-5} . This result is in accordance with the famous equation by Wright (1943) for expected F_{ST} at equilibrium under the island model of population structure:

$$F_{ST} = 1/[1 + 4N(m + \mu)], \quad (1)$$

where N is the population size, m is the migration rate, and μ is the mutation rate. Normally, this formula is simplified by ignoring the mutation rate, because mutation rate is typically so small relative to the migration rate. In Fig. 1, with a migration rate of 10^{-3} , mutation has no noticeable effect on F_{ST} if its rate is $\leq 10^{-5}$. However, for mutation rates that are high relative to migration rates, the effect of mutation on F_{ST} should not be ignored. To generalize this result for other migration rates, we calculated how mutation rate influences F_{ST} using Equation 1 (Fig. 2). Just as in Fig. 1, F_{ST} is hardly affected by mutation when the mutation rate is at least 2 orders of a magnitude smaller than the migration rate, but F_{ST} is increasingly affected as mutation rate increases (see also Whitlock 2011 for similar results). In effect, the curves in Fig. 2 show by which amount F_{ST} is reduced relative to Q_{ST} for a given mutation rate of the neutral marker, assuming a low mutation rate of 10^{-6} or less for the QTL. This result holds for different population sizes, except that for larger population sizes mutation already decreases F_{ST} at lower mutation rates (not shown). As long-term (equilibrium) values of both migration rate and population size are rarely known, it therefore seems safest to use neutral markers with very low mutation rates of 10^{-6} or less, because these give similar F_{ST} values as unselected QTL with mutation rates that also seem to be that low. More variable neutral markers (with higher mutation rates) can only be used when migration rate is at least 2 orders of magnitude larger (Fig. 2). This finding thus generally disqualifies the use of highly polymorphic microsatellites, which can have mutation rates in the order of 10^{-3} (Buschiazzo & Gemmell 2006), for Q_{ST} - F_{ST} comparisons.

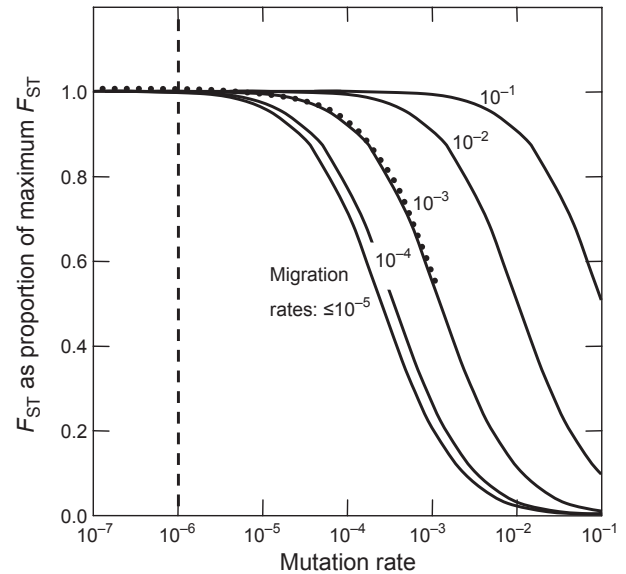


Fig. 2 Mutation decreases F_{ST} when mutation rate is relatively high compared to migration rate. For a fixed average population size of 1000, and variable mutation and migration rates, we calculated equilibrium F_{ST} using Equation 1. Each obtained F_{ST} was divided by the maximum obtained F_{ST} for a given migration rate (i.e. when mutation rate was negligibly low), and this proportion plotted against mutation rate. Only at a rate of 10^{-6} or lower, mutation does not have a noticeable effect on F_{ST} for any level of migration. Note that the dotted part of the curve for a migration rate of 10^{-3} coincides with the results in Fig. 1.

Instead, we would like to promote the use of neutral single-nucleotide polymorphisms (SNPs; reviewed in Brumfield *et al.* 2003; Morin *et al.* 2004). We feel that these provide us with a closer approximation to the genomic mechanisms driving quantitative trait evolution and to the presumably low mutation rates of underlying QTL than do highly variable markers such as microsatellites. However, one should take into account that we might obtain an unrepresentative estimate of F_{ST} if we only focus on few SNPs, because neutral markers are exposed to evolutionary stochasticity. In addition, Q_{ST} of single quantitative traits should not be compared with the average F_{ST} , but with the distribution of F_{ST} (Whitlock 2008). By estimating F_{ST} for a large number of independent SNPs (i.e. without linkage disequilibrium within populations), we can approximate the genome-wide distribution of F_{ST} for independent neutral QTL-like loci. This then provides the expectation for the distribution of differentiation in a neutral quantitative trait, with which we can statistically compare our estimate of the observed Q_{ST} to test for the effect of selection on the trait (Whitlock 2008; Whitlock & Guillaume 2009). Scoring large numbers of SNPs is now becoming increasingly feasible for any kind of organism (Ouborg *et al.*

2010; Tautz *et al.* 2010). It is even possible to take into account that a small percentage of SNPs might be under divergent selection (Brumfield *et al.* 2003; Glover *et al.* 2010), which otherwise tends to yield higher F_{ST} values (but a more conservative test for divergent selection on the quantitative trait of interest). Another issue is the occurrence of multi-allelic SNPs. Assuming these are not scoring errors, their presence reflects the underlying mutation rate of DNA and may reflect the phenotypic effects of a QTL with the same multi-allelic states, so for these reasons multi-allelic SNPs should not be dismissed or corrected. However, as the frequencies of third (or even fourth) alleles for SNPs are low and multi-allelic SNPs are generally rare, the average mutation rate of SNPs is hardly changed by multi-allelic SNPs and their effect on average F_{ST} will be negligible.

Repercussions for past Q_{ST} – F_{ST} comparisons based on multi-allelic markers

Previously, we argued that more variable markers will yield lower F_{ST} values than those of neutral QTLs, and that this could bias comparisons of Q_{ST} with F_{ST} . But are F_{ST} values really much reduced in practice? To test this, Heller & Siegmund (2009) analysed 34 recent studies dealing with 43 species and calculated F_{ST} as well as the alternative population differentiation measures G'_{ST} (Hedrick 2005) and Jost's D (Jost 2008). These alternative measures of differentiation are specifically derived to avoid the effect of within-population heterozygosity on F_{ST} , so that we can compare values across markers or studies. Their results showed that G'_{ST} was on average 85% larger than F_{ST} , and D 60% larger, but could be up to 4.5 times larger than F_{ST} . Meirmans & Hedrick (2010) expanded their search to include another 41 species from 36 studies, which confirmed these findings and showed G'_{ST} and D could be up to 10 times larger than F_{ST} . Both studies clearly showed that F_{ST} decreases when H_{WITHIN} increases, confirming the theoretical expectation that the more variable the marker used, the lower the estimate of F_{ST} .

This has implications for empirical studies that compared Q_{ST} with F_{ST} . The most recent meta-analysis of such studies (Leinonen *et al.* 2008; see also earlier reviews in Merilä & Crnokrak 2001 and McKay & Latta 2002) concluded that the Q_{ST} values are on average higher than F_{ST} values (mean difference 0.12) suggesting a predominant role for natural selection as a cause of differentiation in quantitative traits. An important proportion of these studies are based on markers that likely have higher mutation rates than QTL. Given the theoretical and empirical evidence that in that case F_{ST} provides us with a considerably lowered value, we argue that a considerable part of the studies this meta-

analysis is based upon are likely flawed (Edelaar *et al.* in preparation). If so, the overall conclusion of the meta-analysis is also compromised.

When using markers with high mutation rates, could we not compare Q_{ST} with modern alternatives for F_{ST} ?

Leinonen *et al.* (2008) were in fact aware of the reduction in F_{ST} by high within-population heterozygosity and re-analysed a subset of studies for which they could calculate Hedrick's G'_{ST} (Jost's D was not yet available at the time). While initially this subset of studies showed the same trends as the total set of studies, after this correction for within-population marker variability the average difference between F_{ST} and Q_{ST} was not significantly different from zero for any type of trait. They acknowledged that this result gives support to the contention that the estimates of neutral genetic divergence may be biased downwards. However, they questioned whether G'_{ST} is 'actually the correct measure of neutral genetic differentiation that can serve as the expectation for neutral divergence in Q_{ST} '.

We agree with Leinonen *et al.* (2008) when they suggest that G'_{ST} should not be used as the neutral expectation for Q_{ST} . It was developed as a standardized version of G_{ST} , where the observed G_{ST} is rescaled to the maximum it could reach given the level of within-population heterozygosity: $G'_{ST} = G_{ST}/G_{STmax}$, where $G_{STmax} = (k - 1) * (1 - H_{WITHIN}) / (k - 1 + H_{WITHIN})$, with k equaling the number of populations (Hedrick 2005; a nearly unbiased estimator for G'_{ST} is given in Meirmans & Hedrick 2010). This implies that when comparing two populations for a highly variable marker, G'_{ST} gives a value of 1 as soon as these populations do not share any alleles (Meirmans & Hedrick 2010). In contrast, the putatively equivalent Q_{ST} would be <1 if these alleles were replaced by genotypic breeding values, because we would still have within-population variance.

Jost (2008) attacked the same problem as Hedrick (2005), but applied a mathematical partitioning of genetic diversity (heterozygosity) into pure, independent within- and between-population components, which yielded his $D = [(H_{TOTAL} - H_{WITHIN}) / (1 - H_{WITHIN})] * (k/k - 1)$. Just like G'_{ST} , D takes a value of 1 when comparing two populations that do not share any alleles of a highly variable marker, whereas Q_{ST} would not. This means that in their function as an analogue of Q_{ST} , Jost's D and G'_{ST} overestimate neutral genetic differentiation.

Another way to understand why these modern measures of population differentiation for multi-allelic markers cannot replace F_{ST} in its comparison with Q_{ST} is to realize that there are actually two different perspectives of genetic differentiation (Gregorius 2010).

The first is the classical perspective of differentiation between *populations* for their allelic composition: how far are populations from having nonoverlapping allele frequencies. Jost's D (and to a certain extent G'_{ST}) refers to this kind of genetic differentiation. The second is the opposite perspective of differentiation between *alleles* for their population affiliations: how far are alleles from belonging to nonoverlapping populations (cf. differential fixation between alleles). F_{ST} (and to a certain extent G'_{ST}) refers to this kind of genetic differentiation. Gregorius (2010) concludes that the recent criticism against F_{ST} stems from the fact that it is normally interpreted as the former, but actually should be interpreted as its opposite, i.e. as a measure of differential fixation. Historically, F_{ST} was first applied to bi-allelic, Mendelian genetic markers. Here, the distance and fixation perspectives of differentiation are theoretically identical (Gregorius 2010): when two populations are fixed for an alternative allele, they also have completely nonoverlapping allele frequencies and vice-versa. The distinction between the two types of differentiation therefore never became ingrained into our thinking, and F_{ST} became wrongly equated with population differentiation in either sense.

Just like F_{ST} , Q_{ST} is also a measure of differential fixation, so population differentiation in the second sense. Remember that Q_{ST} only reaches its theoretical maximum of 1 if $V_{WITHIN} = 0$ but $V_{TOTAL} > 0$, which means complete differential fixation. Of course a comparison between two measures of population differentiation should be measuring the same *kind* of population differentiation. Jost's D (and to a certain extent G'_{ST}) does not depend on within-group variability because it is a measure that refers to the first perspective of differentiation, and hence is theoretically not useful to compare with Q_{ST} to test for signatures of selection. Its relationship with F_{ST} is nonlinear as it is not just a simple transformation of F_{ST} (see also fig. 2 in Meirmans & Hedrick 2010; and Whitlock 2011). In addition, as Jost's D captures the difference among populations in allele composition and F_{ST} the differential fixation of alleles across populations, Jost's D can actually be smaller than F_{ST} (see also fig. 2 in Meirmans & Hedrick 2010). For example, when 99 of 100 populations are fixed for allele A and the final population is fixed for allele B, Jost's D is near to zero but F_{ST} (and G'_{ST}) is one (Gregorius 2010; Meirmans & Hedrick 2010). If these alleles were replaced by genotypic breeding values, Q_{ST} would also be 1 as $V_{WITHIN} = 0$ but $V_{TOTAL} > 0$. Finally, Whitlock (2011) argued that Jost's D does not describe the effects of local genetic drift, which would make it unsuitable as an estimator for population divergence by neutral evolution.

It has also been suggested that if F_{ST} can be corrected for within-population variability perhaps Q_{ST} should

also be corrected, as traits that are more variable within populations would have a lower Q_{ST} value (Hendry 2002; Leinonen *et al.* 2008; Meirmans & Hedrick 2010). We do not think this is desirable or even necessary. As Q_{ST} is a fixation index, it *should* reflect the divergence among populations relative to variability within: if a trait is more variable within populations, then populations that have the same absolute divergence will show more overlap and hence are in effect less divergent, as correctly represented by a lower Q_{ST} value. Also, it is true that a higher mutation rate of QTL results in larger values of V_{WITHIN} , as V_{WITHIN} (the genetic variance of the sum of QTL effects) equals the sum of the genetic variances of each QTL, if the effects are neutral, equal, additive and independent (no covariance). In turn, a larger V_{WITHIN} because of a higher mutation rate of QTL means a lower value for Q_{ST} (Fig. 1). However, the same is true for neutral loci: a higher mutation rate means a lower value for F_{ST} (Fig. 1). Hence, if both types of loci have an equal mutation rate, then the comparison between Q_{ST} and F_{ST} is not changed and remains valid despite lower absolute values of both (Fig. 1), and no correction is necessary for Q_{ST} or F_{ST} .

Final remarks

In summary, when comparing Q_{ST} with F_{ST} , we should use neutral markers with a mutation rate that gives an estimate for F_{ST} that equals the estimate for a neutral Q_{ST} given the mutation rate of the underlying QTL. Neutral markers that are more variable than neutral QTL because of higher mutation rates should not be used: F_{ST} estimates for these will be biased downward, and alternative metrics that are designed to deal with differences in marker variability such as Jost's D or G'_{ST} just express something different about the data.

There have been many criticisms against the Q_{ST} method (see e.g. Merilä & Crnokrak 2001; Leinonen *et al.* 2008; Whitlock 2008; Miller *et al.* 2008). For example, theory assumes purely additive and equal effects of QTL to quantitative traits and no epistasis or dominance, which is probably never true in real situations, and this could bias estimates of Q_{ST} . However, even greater issues may be how to correctly estimate Q_{ST} , and the bias introduced by the likely nonrandom choice of traits in empirical studies (Goudet & Martin 2007; Miller *et al.* 2008; Whitlock 2008; Santure & Wang 2009). Following Whitlock (2008), we therefore suggest that empiricists construct a priori hypotheses for traits they expect to be under selection for mechanistic reasons and use the Q_{ST} approach as an initial screening. Such Q_{ST} - F_{ST} comparisons should then be accompanied or followed-up by other approaches to evaluate whether, how and why selection acts on these traits (Whitlock 2008).

Previously, we only discussed F_{ST} -like metrics that evaluate alleles as being identical or nonidentical (i.e. ignoring genetic distance among alleles). However, for sequence data, one can calculate genetic distances among alleles, and under certain assumptions compare within- and among-population variation in these genetic distances to estimate F_{ST} (using AMOVA). It appears that for these kind of data F_{ST} is independent of mutation rate (Kronholm *et al.* 2010; see also Whitlock 2011). In addition, the phenotypic effects of a given QTL may be caused by mutational changes at different positions within the QTL, so in that sense may perhaps also be better represented by a haplotypic point-of-view. The direct use of haplotypes (instead of the SNPs that cause such haplotype variation) for comparisons of Q_{ST} and F_{ST} should be further explored theoretically.

While here we tried to reconcile the recent criticisms against F_{ST} with the usefulness of Q_{ST} - F_{ST} comparisons, this paper at the same time adds to the criticisms against Q_{ST} - F_{ST} comparisons as currently employed. We stress that the neutral loci used to estimate F_{ST} should not have such high mutation rates that they decrease F_{ST} compared to a neutral Q_{ST} . Uncritical use of neutral markers renders Q_{ST} - F_{ST} comparisons to those between apples and oranges.

Acknowledgements

We thank Lou Jost, John Nason, Laurent Excoffier, Michael Whitlock and four anonymous reviewers for feedback and reprints, all contributors to the online Nature Networks Population Genetics forum for added insight, and José Antonio Godoy, Francisco Xavier Picó and Maria Méndez for comments on a previous version.

References

- Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution*, **18**, 249–256.
- Buschiazzo E, Gemmill NJ (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. *Bioessays*, **28**, 1040–1050.
- Drake JW, Charlesworth B, Charlesworth D, Crow JF (1998) Rates of spontaneous mutation. *Genetics*, **148**, 1667–1686.
- Ellegren H, Smith NGC, Webster MT (2003) Mutation rate variation in the mammalian genome. *Current Opinion in Genetics & Development*, **13**, 562–568.
- Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010) Calculations of G_{ST} and D : forget G_{ST} but not all of statistics! *Molecular Ecology*, **19**, 3845–3852.
- Glover KA, Hanse MM, Lien S, Als TD, Høyheim B, Skaala Ø (2010) A comparison of SNP and STR loci for delineating population structure and performing individual assignment. *BMC Genetics*, **11**, 2.
- Goudet J, Martin G (2007) Under neutrality, $Q_{ST} \leq F_{ST}$ when there is dominance in an island model. *Genetics*, **176**, 1371–1374.
- Gregorius H-R (2010) Linking diversity and differentiation. *Diversity*, **2**, 370–394.
- Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.
- Heller R, Siegmund R (2009) Relationship between three measures of genetic differentiation G_{ST} , D_{ST} and G'_{ST} : how wrong have we been? *Molecular Ecology*, **18**, 2080–2083.
- Hendry AP (2002) $Q_{ST} > = < F_{ST}$? *Trends in Ecology and Evolution*, **17**, 502.
- Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting F_{ST} . *Nature Reviews Genetics*, **10**, 639–650.
- Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.
- 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.
- Lande R (1992) Neutral theory of quantitative genetic variance in an island model with local extinction and recolonization. *Evolution*, **46**, 381–389.
- Leinonen T, O'Hara RB, Cano JM, Merilä J (2008) Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, **21**, 1–17.
- McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution*, **17**, 285–291.
- Meirmans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, **60**, 2399–2402.
- Meirmans PG, Hedrick PW (2010) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.
- Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, **14**, 892–903.
- Miller JR, Wood BP, Hamilton MB (2008) F_{ST} and Q_{ST} under neutrality. *Genetics*, **180**, 1023–1037.
- Morin PA, Luikart G, Wayne RK, the SNP workshop group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, **19**, 208–216.
- Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. *Genetics*, **156**, 297–304.
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, **70**, 3321–3323.
- Neuenschwander S, Hospital F, Guillaume F, Goudet J (2008) quantiNemo: an individual-based program to simulate quantitative traits with explicit genetic architecture in a dynamic metapopulation. *Bioinformatics*, **24**, 1552–1553.
- O'Hara RB, Merilä J (2005) Bias and precision in Q_{ST} estimate: problems and some solutions. *Genetics*, **171**, 1331–1339.
- Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma RK, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. *Trends in Genetics*, **26**, 177–187.
- Pujol B, Wilson AJ, Ross RIC, Pannell JR (2008) Are Q_{ST} - F_{ST} comparisons for natural populations meaningful? *Molecular Ecology*, **17**, 4782–4785.

- Roach JC, Glusman G, Smit AFA *et al.* (2010) Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science*, **328**, 636–639.
- Santure AW, Wang J (2009) The joint effects of selection and dominance on the QST–FST contrast. *Genetics*, **181**, 259–276.
- Spitze K (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.
- Tautz D, Ellegren H, Weigel D (2010) Next generation molecular ecology. *Molecular Ecology*, **19**, S1.
- Whitlock MC (1999) Neutral additive genetic variance in a metapopulation. *Genetical Research*, **74**, 215–221.
- Whitlock MC (2008) Evolutionary inference from Q_{ST} . *Molecular Ecology*, **17**, 1885–1896.
- Whitlock MC (2011) $G'st$ and D do not replace F_{st} . *Molecular Ecology*. DOI: 10.1111/j.1365-294X.2010.04996.x.
- Whitlock MC, Guillaume F (2009) Testing for spatially divergent selection: comparing Q_{st} to F_{st} . *Genetics*, **183**, 1055–1063.
- Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.
-
- P.E.'s interests are centered on the processes that drive or limit adaptation, population differentiation and speciation (currently in the context of biological invasions), with a special focus on the role of the individual in this. M.B.'s interests are speciation, the ecological and evolutionary effects of climate change on host-parasite interactions, and the constancy of the G-matrix on ecological time scales.
-