Q_{ST}–F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity

Tuomas Leinonen[1], R. J. Scott McCairns[1], Robert B. O'Hara[2] and Juha Merilä[1]

Abstract | Comparative studies of the divergence of quantitative traits and neutral molecular markers, known as Q_{ST}–F_{ST} comparisons, provide a means for researchers to distinguish between natural selection and genetic drift as causes of population differentiation in complex polygenic traits. The use of Q_{ST}–F_{ST} comparisons has increased rapidly in the last few years, highlighting the utility of this approach for addressing a wide range of questions that are relevant to evolutionary and ecological genetics. These studies have also provided lessons for the design of future Q_{ST}–F_{ST} comparisons. Methods based on the Q_{ST}–F_{ST} approach could also be used to analyse various types of 'omics' data in new and revealing ways.

Genetic drift
Random change in allele frequencies due to stochastic factors.

Deme
A group of individuals that actively interbreed and share a common gene pool.

Directional selection
Selection that favours the extreme phenotypes at one end of the distribution but disfavours those at the opposite end.

[1]*Ecological Genetics Research Unit, Department of Biosciences, PO Box 65, FI-00014 University of Helsinki, Finland*
[2]*Biodiversity and Climate Research Centre, D-60325 Frankfurt, Germany*
Correspondence to J.M.
e-mail: juha.merila@helsinki.fi
doi:10.1038/nrg3395
Published online
5 February 2013

Most plant and animal species, including humans, are subdivided into many partially isolated subpopulations. Depending on the relative strengths of natural selection, genetic drift, migration and mutation, these subpopulations become differentiated — both genetically and phenotypically — over time[1,2]. Understanding the causes and consequences of this differentiation is of broad interest in different disciplines of biological sciences, including both fundamental research (for example, evolutionary biology, ecology and genetics) and applied realms (for example, forestry and fishery management, medicine and conservation biology). Of particular interest is determining to what degree population differentiation is caused by selective (that is, adaptive) versus neutral (that is, stochastic) processes.

At the genetic level, there is a well-developed theory and body of empirical evidence explaining population differentiation. The degree of this differentiation can be measured by F_{ST} (BOX 1), which is a standardized measure of genetic differentiation among populations for a genetic locus[2]. For neutral loci that are not influenced by natural selection, the degree of differentiation among subpopulations depends largely on their effective size and the amount of migration between them: small, isolated populations tend to become more differentiated from each other than large populations that are connected by gene flow (for example, REF. 3). However, the degree of genetic differentiation among subpopulations also depends on the strength and nature (for example, diversifying or balancing selection) of the predominant selective pressures experienced by the populations or demes under study. In the case of adaptive population divergence, directional selection is expected to increase F_{ST} of selected or linked loci, relative to that of neutral loci[4,5]. Yet, because most quantitative traits of evolutionary, ecological, economic and even of medical interest — such as body size and intelligence quotient — are known or thought to have a polygenic basis[6,7], distinguishing neutral and selective patterns of population differentiation at the phenotypic level is not easily accomplished with standard F_{ST} estimates. Trait-based inference, however, can be accomplished under a related analytical framework.

Q_{ST} is a quantitative genetic analogue of F_{ST} that measures, similarly to F_{ST}, the amount of genetic variance among populations relative to the total genetic variance in the trait (rather than at a specific locus in the case of F_{ST}; BOX 1). The value of Q_{ST} for a neutral quantitative trait that has an additive genetic basis is expected to be equal to the F_{ST} for a neutral genetic locus (BOX 1). This finding — which is based on the work of Sewall Wright[8] (see REF. 9 for a historical account of the development of the method) — provides a basis for evolutionary inference: given a set of assumptions (see below), F_{ST} measured from neutral molecular markers can be used as a null expectation for the degree of population divergence due to drift and migration[10,11]. In cases in which $Q_{ST} \approx F_{ST}$, the inference is that trait divergence among subpopulations could have been achieved by genetic drift alone. If $Q_{ST} > F_{ST}$, trait divergence exceeds neutral expectation,

Box 1 | **Deriving measures of population differentiation in molecular markers and phenotypic traits**

Wright's F_{ST} and related estimators[2] allow the partitioning of the total genetic variation (σ^2_{GT}) in single genes (for example, neutral marker loci) into within-population (σ^2_{GW}) and between-population (σ^2_{GB}) components, such that a standardized measure of the degree of among-population allelic differentiation is obtained as:

$$F_{ST} = \sigma^2_{GB}/(\sigma^2_{GB} + \sigma^2_{GW}) \quad (1)$$

Similar partitioning of genetic variance in a quantitative polygenic trait for populations diverging owing to genetic drift can be achieved by relating the components of allelic variation from equation 1 to those of genetic variance in the polygenic trait as[8]:

$$\sigma^2_{GB} = 2F_{ST}\, V_A \quad (2)$$

$$\sigma^2_{GW} = (1 - F_{ST})\, V_A \quad (3)$$

$$\sigma^2_{GT} = (1 + F_{ST})\, V_A \quad (4)$$

where V_A refers to additive genetic variance in a common ancestral population. Therefore, a quantitative trait analogue of F_{ST} — coined Q_{ST}[10] — can be obtained as:

$$Q_{ST} = \sigma^2_{GB}/\sigma^2_{GT} = 2F_{ST}/(1 + F_{ST}) = \sigma^2_{GB}/(\sigma^2_{GB} + 2\sigma^2_{GW}) \quad (5)$$

Hence, for neutral traits in diploid organisms, the quantities defined by equations 1 and 5 are expected to be the same. Building on the framework of variance partitioning first outlined by Wright, both Lande[13] and Whitlock[52] have independently confirmed the expectation of equivalency under neutrality. Thus, barring technical challenges in estimating the quantities of interest, this expectation ($F_{ST} = Q_{ST}$) provides a theoretically sound basis for inferring deviations from neutrality ($Q_{ST} > F_{ST}$ or $Q_{ST} < F_{ST}$). As long as the markers used to estimate F_{ST} are neutral, and the within- and among-population components of variance that are used to estimate Q_{ST} are based on genetic rather than phenotypic data, the neutral expectation ($F_{ST} = Q_{ST}$) is shown to be robust under a variety of demographic scenarios[13,25,52].

and is likely to have been caused by directional selection. If $Q_{ST} < F_{ST}$, trait divergence among populations is less than expected by genetic drift alone; this pattern is suggestive of uniform selection or stabilizing selection across the populations.

Q_{ST}–F_{ST} comparisons have been used in an increasing number of studies to infer the action of natural selection on complex phenotypic traits, as well as to quantify the degree of spatial genetic structuring in quantitative traits among populations (see REFS 9,11,12 for earlier reviews). In this Review, we first introduce the concepts and issues related to estimating the parameters of interest. We then summarize the insights that have been accumulating from the rapidly increasing number of empirical and theoretical studies focused on Q_{ST}–F_{ST} comparisons and discuss generalizations that are emerging from the empirical data. We will also cover some of the recent methodological and conceptual developments and challenges relating to the use of Q_{ST}–F_{ST} comparisons, and we explore the utility, applicability and promise of the Q_{ST}–F_{ST} approach in relation to quickly evolving genomic methods and rapidly accumulating genomic data.

Estimating and comparing F_{ST} and Q_{ST}

Estimating Q_{ST}. Q_{ST} is defined as the proportion of variation in a given trait that is attributable to genetically based differences among populations, as scaled to the total genetic variation in the trait[10,13]. This can be represented as $Q_{ST} = \sigma^2_{GB}/(\sigma^2_{GB} + 2\sigma^2_{GW})$, where σ^2_{GB} is the among-population (additive) genetic component of variance, and σ^2_{GW} is the average within-population component of additive genetic variance (V_A) (BOX 1). From this, it is obvious that the estimation of Q_{ST} requires quantitative genetic data from multiple populations. This is also the Achilles heel of the Q_{ST}–F_{ST} approach: logistic demands to obtain such data can be formidable, as the estimation of the parameters of interest requires breeding experiments that are conducted under standardized environmental conditions (that is, 'common-garden' experiments). Although labour-intensive, this is the most reliable method of ensuring that among-population variance components reflect genetic differences, and are not inflated by direct environmental effects. Moreover, breeding experiments can be designed such that within-population components of variance truly reflect additive genetic variance. A basic text on quantitative genetics will provide numerous design options, but strategies that use half-sib crosses will generally provide unbiased estimates of V_A. If a half-sib design is not tractable, a reasonable approximation of σ^2_{GW} can be achieved from a sufficient number of full-sib families. Alternatively, by relaxing assumptions underlying standard quantitative genetic inference, purely phenotypic data can be used to estimate P_{ST}, which is a Q_{ST} analogue that lacks some of the genetic rigour of Q_{ST} (BOX 2). The P_{ST} approach is the only option when species are not suitable for the breeding designs that allow the estimation of V_A. In these cases, which often involve species of conservation interest, P_{ST} can provide a reasonable proxy for Q_{ST}. However, the impact of its underlying assumptions regarding the magnitude of environmental effects on within- and among-population components of variance should always be evaluated with sensitivity analyses[14] (BOX 2).

Uniform selection
Selection that favours similar phenotypes in different populations.

Stabilizing selection
Selection that eliminates both extremes and favours the intermediate phenotypes.

Additive genetic variance
(V_A). The part of total genetic variance that determines the response to selection in quantitative traits. It can be modelled as allelic effects that have an additive effect on the phenotype in heterozygotes.

Common garden
An experimental setting in which individuals from different populations are reared under identical environmental conditions to standardize environmental influences on phenotypes.

Box 2 | P_{ST}: inference without common-garden data

Given the difficulty of estimating within- and among-population components of genetic differentiation in quantitative traits, it has become a popular practice to replace Q_{ST} with its phenotypic analogue, P_{ST}, a term coined by Leinonen *et al.*[94]. The main challenges involved with use of P_{ST} are that both within- and among-population components of variance (equation 5 in BOX 1) can be confounded by environmental effects[95]. Although the inclusion of environmental variance in the within-population component of variance is likely to render P_{ST} estimates conservative, the opposite is true in the case of the among-population component of variance. Namely, environmental differences experienced by different populations are a common source of phenotypic divergence in a wide range of taxa and traits. For example, Q_{ST} values estimated using wild-caught copepods (that is, P_{ST}) were 1.8 times larger than those estimated using animals reared in a common-garden experiment from the same populations[96]. However, when judiciously applied, P_{ST} estimates can still be informative: sensitivity analyses can be carried out to evaluate the impact of assumptions regarding the magnitude of environmental effects on within- and among-population components of variance[14,49,97]. Meta-analytical results are also reassuring, as they show that P_{ST} estimates are not generally higher than Q_{ST} estimates[9]. That said, given the widespread occurrence of counter-gradient variation[98], this similarity could be coincidental. Nevertheless, being measured on the same scale as Q_{ST} and F_{ST}, P_{ST} estimates provide a useful yardstick to compare the relative influence of genetic adaptation, phenotypic plasticity and genetic drift as causes of population differentiation. Hence, although P_{ST} estimates cannot provide hard evidence proving the action of natural selection in the past, they are informative regarding the degree of phenotypic differentiation among populations and different traits.

When quantitative genetic data from multiple populations have been obtained, Q_{ST} and associated dispersion estimates (that is, standard errors or confidence limits) can be obtained using linear mixed-model approaches. However, because estimating Q_{ST} is essentially based on estimating variance components and their ratios (see the above equation for Q_{ST}), the precision of these estimates is likely to be poor unless many populations and families are used in the estimation[15]. Although this may represent a challenge for studies of organisms that are not easily cultured in common-garden settings, studies based on small numbers of populations and families can still be informative provided that selection has been strong (that is, $Q_{ST} >> F_{ST}$). Various statistical approaches have been used to estimate confidence intervals and standard errors of Q_{ST}, many of which have been shown to yield erroneous or even biased estimates[15]. In this respect, Bayesian methods and the parametric bootstrap seem to work best[15].

Numerous technical refinements that help to deal with some additional issues in Q_{ST} estimation have been developed in recent years. For example, progress in multivariate Q_{ST}–F_{ST} methods[16–19] has been a welcome development as a means of dealing with the causes and consequences of correlated selection. Because of pleiotropy and linkage, different traits can be genetically correlated and hence not free to evolve independently[20]; consequently, inference based on traditional univariate analyses of Q_{ST} cannot distinguish between selection acting directly on a trait, or a correlated signature that arises through selection acting on covarying traits. Although they have been under-used to date, such multivariate approaches represent an important step towards the goal of studying evolution of the phenome, given hierarchical — and sometimes conflicting — genome-wide influences on trait expression. The Q_{ST}–F_{ST} approach can also be used to study hierarchical partitioning of genetic variance in quantitative traits across varying levels of spatial structure[21,22]. The challenge here is to correct for different variance parameters at different levels of the hierarchy to make them analogous to F_{ST}. A method to do this has been introduced only recently[23].

Estimating F_{ST}. Accurately defining a neutral baseline for the degree of differentiation that is expected under genetic drift is equally crucial for the successful implementation of this analytical framework. This can be obtained by estimating F_{ST} for neutral marker loci, such as microsatellites or appropriate single-nucleotide polymorphisms (SNPs), in the classical Weir and Cockerham[24] framework. The methods and issues associated with estimation of F_{ST} have been recently covered in many excellent reviews[2,25,26]. Here, it suffices to say that the crucial issues from the perspective of Q_{ST}–F_{ST} comparisons are whether the markers used are indeed neutral, and whether their mutation rates are not too high relative to rates of migration[27]. Marker neutrality may be less of an issue if inference is based on the distribution of mean F_{ST} derived from a large number of loci; however, the use of mean F_{ST} is probably inappropriate. Whitlock[25] has demonstrated that Q_{ST} tends to behave similarly to a single-locus F_{ST}, and as such, the correct approach is to compare Q_{ST} against the distribution range of single-locus F_{ST} estimates, not their mean. In this instance, outlier loci have the potential to bias the threshold values defining neutrality, as loci under selection will inflate the variance, thereby setting an artificially inflated upper value of F_{ST} and potentially leading to a failure to detect traits under selection. In these cases, we would argue that testing assumptions of marker neutrality should be considered an essential first step in any Q_{ST}–F_{ST} analysis.

However, care must also be taken not to artificially deflate neutral thresholds. Although there are numerous methods currently available to detect outlier loci that are under selection[28,29], such tests can be sensitive to false positives[30–32]. The practical consequence of this would be removal of high-but-neutral F_{ST} values (false-positive outliers) when calculating thresholds of neutral divergence, thereby potentially leading to false positives in the evaluation of Q_{ST}. This seems to be less problematic for Bayesian-based outlier tests[33–35], but it is important to be aware of this potential source of inferential bias.

Recently, the choice of marker that is used to define F_{ST} has also become a cause for concern. Microsatellites in particular have been criticized as being too variable, and their use as markers may lead to situations in which within-group heterozygosity is sufficiently high to significantly deflate F_{ST}, even when calculated for theoretically maximally divergent populations. The practical result is the setting of too liberal a neutral baseline expectation against which Q_{ST} is evaluated[36]. This has led some authors to advocate the use of SNP data[37], based on the assumption of lower mutation rates. However, it is currently unclear whether the enthusiasm for SNP data is well founded. Nucleotide mutation rates are highly variable, depending on their

Bayesian methods
Statistical methods in which the probability of a hypothesis is tested using a prior probability, which is updated whenever new data are obtained. Estimated parameters are derived from a posterior distribution.

Parametric bootstrap
A method of estimating confidence intervals from simulated data sets that are constructed from a fitted statistical model. This contrasts non-parametric bootstrapping in which estimates are derived by resampling data with replacement.

Neutral marker loci
Loci (for example, microsatellites or SNPs) that are inherited in a Mendelian manner and not influenced by selection.

Microsatellites
Short repeated sequences of DNA.

location in the genome and the type of mutation[38,39]. Moreover, the significance and degree of potential bias from mutation relative to other sources of variance in F_{ST} and Q_{ST} estimates is as yet unclear, although simulation studies suggest that bias is unlikely to be severe unless migration rates are low relative to mutation rates in marker loci[27,40].

SNP data also have other characteristics that complicate their use in Q_{ST}–F_{ST} comparisons. For example, the use of SNP panels has been shown to upwardly bias F_{ST} estimates[41], probably owing to the fact that most SNP panels contain a mixture of both neutral and selected markers[42]. Yet, an advantage of SNPs that gives them the potential to surpass the utility of microsatellites is the relative ease with which large numbers of markers can be scored, thereby permitting a more reliable estimate of the distribution of single-locus F_{ST}. Thus, there is probably no simple solution to the question of marker selection at this time: careful pre-planning and marker screening — with an emphasis on minimizing false positives in outlier detection — before inference are highly recommended.

Comparing Q_{ST} to F_{ST}. The theory and conditions for the expectation $Q_{ST} = F_{ST}$ under neutrality rest on firm theoretical foundations (BOX 1), and the empirical results from a few experimental tests give at least qualitative support for these theoretical expectations[43,44]. Moreover, as the metric (that is, Q_{ST}) is scaled identically to the empirical distribution of null (that is, neutral) expectation against which it is to be compared, it represents both a theoretical and practical advantage over other statistical tests for natural selection for which the neutral distributions are dependent on multiple assumptions (BOX 3). However, the correct application of Q_{ST}–F_{ST} comparisons depends on the recognition that both parameters are estimated from data, and as such, accurate inference is dependent on the dispersion of those estimates. Regrettably, this issue was frequently ignored in early applications of the technique, which tended to focus on point estimates of the parameters. Although this error is less common in recent studies, examples can still be found in the literature; thus, the message bears repeating.

Various direct numerical approaches have been used to compare Q_{ST} and F_{ST} estimates. However, as Whitlock has pointed out[25], it is important to make the distinction between two types of Q_{ST}–F_{ST} comparisons. The first refers to comparisons of mean Q_{ST} estimated across several traits with the mean F_{ST} estimated over several loci. The second involves the comparison of Q_{ST} for a single trait to that of a distribution of F_{ST} estimated for several loci. Although the two types of comparisons are deceptively similar, different statistical approaches are needed for each[25,45]. Comparing the mean Q_{ST} across a number of traits might be useful for example in conservation planning, when trying to assess the overall importance of local adaptation in a species, but there are various problems with using the mean Q_{ST}, such as non-independence of the traits measured[25]. Thus, in most cases the Q_{ST} of a single trait and its comparison with a distribution of F_{ST} across a number of loci is the appropriate method of choice. Yet even in the case of single-trait comparisons, meaningful inference also requires an estimate of the statistical error around Q_{ST}, which is a demand that can be met through bootstrapping or sampling from the posterior distribution of Bayesian-based estimates.

As previously noted, the usual comparison of Q_{ST} to mean F_{ST} is incorrect, as under neutrality Q_{ST} is expected to behave similarly to a single-locus F_{ST}[25], and there can be appreciable variation in F_{ST} across neutral loci. Whitlock & Guillaume[45] have recently developed a simulation approach to test whether the Q_{ST} of a given trait is consistent with the null hypothesis of selective neutrality. This is achieved by first simulating a range of neutral Q_{ST} values (Q^{n}_{ST}) — which can be derived using F_{ST} values from even a fairly small number of loci[45] — and then testing whether Q_{ST} for the trait of interest falls within

Box 3 | Alternatives to Q_{ST}–F_{ST} comparisons

There are various alternative formal model-based approaches to detect or infer the action of natural selection, each of which is accompanied by specific limitations and/or assumptions.

Phenotype-based inference

Direct measurements of natural selection are possible[20], but challenging, especially in a multiple-population context. Even when possible, current and past selection pressures could be different. Conversely, direct measures of neutral trait divergence can be inferred from experimental lines, such as from mutation-accumulation experiments[57], thus yielding an alternative baseline against which among-population differences can be compared. However, such an approach is necessarily limited to organisms that have extremely rapid generation times. Lande's[99,100] rate test and equivalents (for example, REFS 101,102) are ultimately quantitative genetic approaches, but they require information and/or assumptions on quantities (for example, mutation rate and time since divergence) that are not usually available. Consequently, a relatively small number of studies have used these approaches[72,83,103,104].

Genotype-based inference

Tests based on patterns of DNA polymorphisms, such as the McDonald–Kreitman test[105], are restricted to coding DNA or protein sequences, and the link to complex phenotypes is not usually traceable. Genome scans and outlier tests[28,29,106] provide yet another way to detect the action of past natural selection, but the link between selected loci and phenotypes is usually not easy to establish, except perhaps for oligogenic or monogenic traits. Conversely, a $Q_{ST} \neq F_{ST}$ result implies heterogeneity in the patterns of genomic differentiation. As such, it becomes natural to ask whether outlier detection methods can be used to identify quantitative trait loci (QTLs) under selection. The general answer to this question is, unfortunately, negative[12,40]. There are two main reasons for this. First, most quantitative traits are coded by many genes[6], and small shifts in allele frequencies at individual QTLs are hard to detect using genome scans[40]. Second, apart from shifts in allele frequencies at individual QTLs, covariance among allelic effects across QTLs also contributes to population differentiation in quantitative traits (that is, Q_{ST}[12]). Therefore, pronounced adaptive differentiation among populations can take place without detectable differentiation in underlying QTLs[12,92,107].

The advantage of Q_{ST}

From a practical perspective, the contrast between Q_{ST}–F_{ST} and genome scans for detecting adaptive differentiation in quantitative traits is one of an unfortunate trade-off. Although genome scans are logistically easier to conduct than Q_{ST}–F_{ST} studies, the latter are inferentially superior when focal traits are strongly polygenic. That said, as shown by both simulations and empirical data, under certain conditions (such as high gene flow, strong divergent selection and fairly simple genetic architecture), genome scans can detect adaptive differentiation[40]. Nevertheless, the Q_{ST}–F_{ST} approach exhibits a distinctive advantage of practical importance over genome scans: the rate at which phenotypic differentiation as measured by Q_{ST} reaches its equilibrium in response to local selection is substantially faster (tens rather than hundreds of generations) than allele frequencies at QTLs[40], thereby permitting the detection of recent selective events (see also REF. 108).

Whitlock says: simulate Q_{ST} values from the SNPs used for F_{ST}. Then compare the observed Q_{ST} to the null distribution.

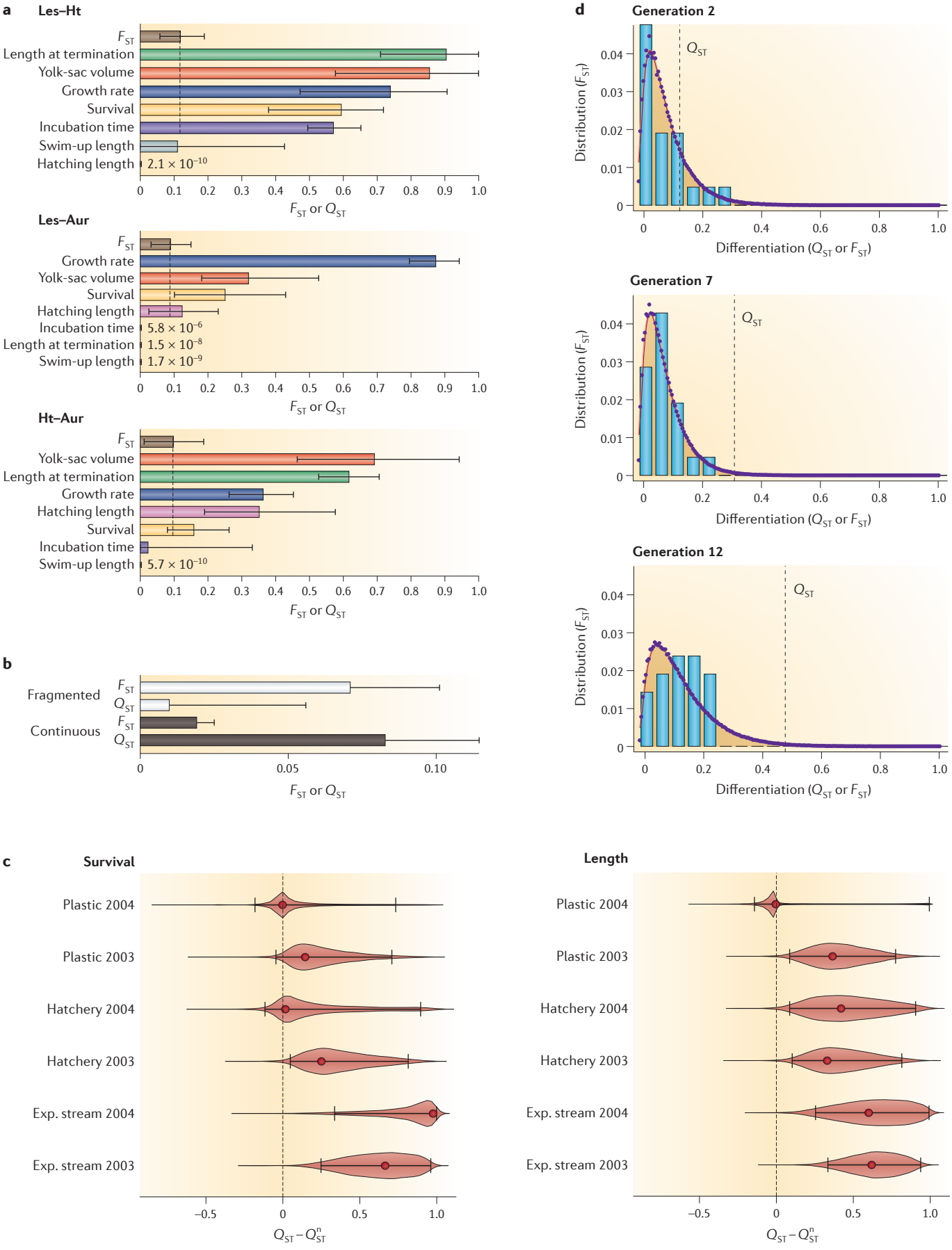
this neutral distribution. This method has been shown to have better statistical power and a lower type 1 error rate than the traditional method of comparing Q_{ST} and F_{ST}[45]. The multivariate method of Ovaskainen *et al.*[19] also makes use of this approach, and is able to disentangle genetic drift from selection even when data are available for only a few populations. This is an important point, given that quantitative genetic data from multiple populations is difficult to obtain, and that there is an inherent lack of precision in Q_{ST} estimates that are obtained using traditional methods from a small number of populations[15]. Another virtue of this method is that it allows genetic drift and directional selection to be distinguished between as causes of population differentiation, even in cases in which the traditional Q_{ST}–F_{ST} approach loses it power, that is, when levels of neutral differentiation are very high[46]. Although the multivariate approach of Ovaskainen *et al.*[19] has not yet been used in any empirical study, a few recent studies have followed the recommendation by Whitlock & Guillaume[45] and compared observed Q_{ST} estimates with the distributions of Q_{ST} that are expected under neutrality[47–51] (FIG. 1c).

Caution should be exercised when interpreting any empirical result, as theory assumes that both within- and among-population components of variance reflect pure additive genetic variance: the presence of non-additive variance can cause Q_{ST} to be greater or smaller than F_{ST}, even for neutral traits[52–56]. Non-additive variance may complicate inference from Q_{ST}–F_{ST} comparisons, especially when highly differentiated groups (for example, subspecies) are compared[23]. The effects of epistasis on Q_{ST} have not been explored in detail, but Whitlock[52] found that simple additive-by-additive epistasis in a neutral trait is likely to bias Q_{ST} estimates downwards, thus rendering tests for directional selection conservative. Likewise, the effect of dominance seems to lower Q_{ST} with respect to neutral expectation[53,54]. However, under certain conditions, dominance can also inflate Q_{ST} over its neutral expectation[55,56], but it seems that such inflation is unlikely for traits that involve many loci[54]. Control over non-additive sources of variance can be achieved through specialized breeding designs and explicit statistical modelling and may be crucial for certain Q_{ST}–F_{ST} comparisons (for example, traits with a monogenic architecture). However, their likely effect is to reduce the power of the method to detect selection, so that the inference of directional selection will be conservative. When it comes to detecting and testing for stabilizing selection, this downward bias represents a major challenge, as tests will tend to be liberal if dominance is present.

Users of the technique should also be aware of an unresolved source of potential bias: the unknown effect of mutational variance. In this respect there may be two main points to consider: first, what are the frequencies of mutations at coding genes relative to neutral regions of the genome? Second, what are the relative phenotypic effects of such mutations, and how are they likely to affect the estimation of Q_{ST}? Just as mutational differences can affect precision in estimates of F_{ST}, so too might differential mutation rates in causative loci influence precision in Q_{ST} estimates. However, results of mutation-accumulation experiments hint at a similarity in rates of mutation: empirically based observations show that the rate of neutral mutation in phenotypes varies widely across traits and species, much to the same extent as the variability in the rate of mutation that is reported in neutral genetic markers[57]. Results of theoretical modelling also suggest that the rate of mutation accumulation — when environmental variance (V_E) is moderate ($0.001 \leq V_E \leq 1$)[58] — may be similar to that reported for microsatellite markers[59]. A few experiments have also directly compared molecular and phenotypic mutation rates, again with results that vary across species. For example, in *Caenorhabditis elegans*, rates are approximately equal, whereas in *Drosophila melanogaster*, the rate of molecular mutation may be 2–3 times greater than that of neutral, phenotypic change[57]. The reassuring conclusion we might draw from these observations is that even when differences exist, they are not orders of magnitude apart. Thus, rate-based discrepancies between genetic and phenotypic indices of divergence may prove to be a limited cause for concern, although this awaits formal and detailed investigation.

Potentially more problematic may be differences in the magnitude of mutational effects. Recent simulations suggest that for mutations occurring incrementally with a constant variance, Q_{ST} seems to be surprisingly immune to the effects of mutation rate[27]. As quantitative traits may accumulate mutational variance over many causative loci[60], under a constant mutation rate per locus, the value of Q_{ST} under neutrality might be expected to decrease inversely with the number of loci. Coincidentally, if the loci influencing variability in quantitative traits are also subject to high rates of mutation, this could also compensate for related bias in F_{ST}[26,27]. However, whether a high number of underlying loci can compensate for a low per-locus mutation rate depends on the genetic architecture. For purely additive inheritance, the number of loci might cancel out[27]. Extrapolating interpretations of previous work on the phenotypic effects of mutations is also encouraging with regard to the validity of Q_{ST}-based inference. For example, work by Caballero *et al.*[61] suggests that varying mutational effects are more likely to influence the dominance component of phenotypic variance. As previously discussed, this is likely to bias Q_{ST} estimates downwards, making inference more conservative, rather than over-estimating potentially 'false' signals of selection.

In summary, there are various issues to be considered when interpreting results from Q_{ST}–F_{ST} comparisons, some of which still await further investigation, and others for which results of recent studies provide yet largely unused solutions. The outstanding challenges aside, Q_{ST}–F_{ST} comparisons provide a well-founded and tractable inferential metric that implicitly encompasses contemporary and historical phenotypic changes as reflected in quantitative genetic parameters. However, the onus remains on the researcher to ensure that the technique is properly applied. Ideally, experiments should be designed with the same rigour as any breeding



◀ Figure 1 | **Q_{ST} is a highly flexible index that is useful for addressing diverse biological questions.** Q_{ST}–F_{ST} comparisons can be applied to a diverse range of taxa to address myriad questions across disciplines. **a** |The traditional application of Q_{ST}–F_{ST} comparisons is to investigate the relative roles of natural selection and genetic drift in population differentiation. The example here shows that natural selection has been the driving force for population divergence in various life-history traits in grayling (*Thymallus thymallus*). Trait differentiation is shown between different pairs of three Norwegian grayling populations (Les, Ht and Aur), which originated from a common source 80–120 years ago. Dashed vertical lines indicate the effect of genetic drift (F_{ST}), horizontal lines indicate confidence intervals and the horizontal bars indicate the effect of selection (Q_{ST}). **b** | The technique also has practical applications in the context of conservation; for example, for comparing relative roles of genetic drift (F_{ST}) and selection (Q_{ST}) in common frog (*Rana temporaria*) populations living in continuous and fragmented habitats. The patterns suggest that genetic drift may be constraining adaption in fragmented landscapes. **c** | Best practices dictate that index distributions are considered. For Q_{ST} alone, this can include the comparison of simulated distributions of Q_{ST} with the distribution expected under neutrality (Q^n_{ST}). Here, the difference between Q_{ST} and Q^n_{ST} values (values >0 indicate directional selection) are shown for the brown trout (*Salmo trutta*) for mean trait values for survival and body length, and for the plasticity of these traits. Results for hatchery-reared and wild (experimental (exp.) stream) populations provide evidence for divergent selection in both settings. **d** | However, it is more typical that Q_{ST} is compared to a neutral distribution that is inferred from molecular markers. For example, flowering time has been shown to respond to selection over 12 generations of divergent selection. Here, Q_{ST} values (dashed vertical lines) are compared to a simulated distribution of F_{ST} (red lines and purple dots) and the actual F_{ST} values from 21 microsatellite markers (blue bars). In **a**, **b** and **d**, F_{ST} and Q_{ST} are shown on the same scale, and the horizontal axes depict either the value of F_{ST} or Q_{ST}. Part **a** is reproduced, with permission, from REF. 72 © (2002) Macmillan Publishers Ltd. All rights reserved; part **b** is reproduced, with permission, from REF. 69 © (2007) Wiley; part **c** is reproduced, with permission, from REF. 51 © (2012) Wiley; and part **d** is modified, with permission, from REF. 74 © (2010) Wiley.

plan used in quantitative genetics. Equal care and consideration should go into the choice of markers used to infer the neutral baseline against which Q_{ST} is ultimately compared. This should include a routine screening for outlier loci, and as computational solutions for modelling and estimating mutation rate become tractable, a screen to ensure that mutational variance will not bias estimation.

Evolutionary and ecological insights

There has been an exponential increase in the number of studies using the Q_{ST}–F_{ST} approach: whereas two previous meta-analyses listed 18 (REF. 11) and 62 (REF. 9) studies, our literature search retrieved 148 studies (Supplementary information S1 (table)). A steady increase can also be seen in the number of theoretical studies, which now number 36 (8 were published by 2001, whereas 22 had been published by 2008). Empirical studies cover a wide range of taxa (from pathogenic fungi[62,63] to humans[64,65]), traits and research problems not only in various subdisciplines of evolutionary biology, but also in evolutionary genetics, plant and animal breeding sciences, forestry and conservation biology. Some representative studies, illustrating the applicability and diversity of issues that have been addressed using Q_{ST}–F_{ST} comparisons, are listed in TABLE 1.

Detecting selection. Typically, Q_{ST}–F_{ST} comparisons are used as an exploratory tool to detect traits that are under selection, especially in cases in which background information on the traits is limited[9]. However, applying the Q_{ST}–F_{ST} approach at different taxonomic levels (from demes to species) and at different spatial scales (such as populations originating from different areas or habitats) opens up possibilities to address an even wider range of questions.

Usually the objects of study are populations that are known to differ in morphological or life-history characters, and the aim is to find out the extent to which natural selection explains the differentiation. For example, in the first study to explicitly use the term Q_{ST}, Spitze[10] showed that in *Daphnia obtusa* populations that were known to have diverged in quantitative traits, Q_{ST} of body size exceeded the corresponding F_{ST} from neutral allozyme markers. This provided evidence that natural selection had been the driving force behind the observed differentiation in body size[10]. A common inference from Q_{ST}–F_{ST} comparisons is that local adaptation can take place despite high gene flow (that is, low F_{ST}; FIG. 1a). In one example of this, despite high gene flow between sympatric rainbow smelt (*Osmerus mordax*) ecotypes, they were found to maintain adaptive differentiation through divergent selection on feeding-related traits[66].

Although the earliest studies were done on model invertebrates, Q_{ST}–F_{ST} comparisons have since been used to detect selection in myriad taxa. For example, a review of studies applying Q_{ST}–F_{ST} comparisons to forest trees found that differentiation in most of the twelve species across a range of life-history and morphological traits exceeded neutral differentiation, indicating that these traits have been subjected to diversifying selection[67]. The magnitude of adaptive differentiation as reflected in Q_{ST} was dependent on the geographical range and number of populations used[67].

Addressing spatial and temporal questions. Although the geographic distance separating populations can influence Q_{ST} and F_{ST}, their comparison can also shed light on the relative importance of population history and natural selection in explaining population differentiation. This opens up a wide range of applications for which Q_{ST}–F_{ST} comparisons can be useful. For example, Q_{ST}–F_{ST} comparisons have been used to shed light on biological invasions by comparing indices of divergence in the invasive species' native and invasive ranges, thus providing information on the evolution of invasiveness and the adaptive potential of invasive species[68]. Q_{ST}–F_{ST} comparisons have also been used to study possible constraints on adaptive differentiation that are imposed by habitat fragmentation. For example, a comparison of Q_{ST} and F_{ST} in continuous and fragmented habitats found that habitat fragmentation was associated with increased genetic drift and a lower degree of adaptive differentiation in common frogs (*Rana temporaria*)[69], thus suggesting a reduced adaptive potential in fragmented versus continuous habitats (FIG. 1b). The breadth of possible applications and inferences made from Q_{ST}–F_{ST} comparisons widens even further when a temporal dimension is added. For example, Q_{ST}–F_{ST} comparisons across generations provided experimental evidence

Allozyme
One of two or more enzymes that are encoded by different alleles at the same locus.

Sympatric
Species or populations that exist in the same geographical area.

Table 1 | **Examples of applications of Q_{ST}–F_{ST} comparisons**

Context	Species	Inference
Local adaptation	*Rana temporaria*[109], *Tyto alba*[110], *Helianthus maximiliani*[111], various tree species[67]	Identification of natural selection as a cause of broad-scale clinal variation in morphological and life-history traits
Sexual selection	*Silene latifolia*[112]	Identification of sex-specific selection as the cause of evolution of sexual dimorphism
Speciation	*Pundamilla* spp[113], *Larus* spp[114]	Adaptive divergence maintains species integrity despite high gene flow
Evolutionary stasis	*Antichiropus variabilis*[115], *Pinus pinastris*[50]	Identification of selective constraints explaining phenotypic uniformity across species ranges
Human-induced evolution	*Thlaspi caerulescens*[116], *Rana temporaria*[69], *Arabidopsis halleri*[117]	Demonstrations of how human-induced habitat changes can either cause or impair adaptation
Artificial selection	*Oryza sativa*[118], *Zea mays*[119]	Demonstrations of how selective breeding shapes diversification and population structuring of crop species
Conservation	*Arabis fecunda*[120], *Araucaria araucana*[121]	Demonstrations that setting conservation priorities should not be based only on neutral marker diversity, and that Q_{ST}–F_{ST} comparisons can be used to identify populations that are suitable for translocation
Management	*Liatris scariosa*[121], *Salmo trutta*[51]	Identification of units or populations that are suitable for translocation or stocking
Transcriptomics	*Salmo salar*[83]	Identification of genes under selection using the distribution of Q_{ST} values of transcription levels
Human evolution	*Homo sapiens*[64,122]	Identification of adaptive phenotypic differentiation among human populations

for contemporary adaptive evolution of phototactic behaviour in *Daphnia magna*[70].

Although adding a temporal dimension to Q_{ST}–F_{ST} comparisons increases the number of possible applications, across-generation comparisons of Q_{ST}–F_{ST} are quite rare. In natural populations these have been mainly limited to demonstrating adaptive genetic evolution in anadromous fish that have recently colonized freshwaters, and for which the history of colonization is well known (for example, REF. 71). Human-assisted introductions have also been used to demonstrate adaptive genetic evolution (for example, REF. 72) (FIG. 1a). Combining data from natural and captive populations has the potential to provide information on the rates of evolution in different environments. It can also inform conservation and management when captive breeding is done for the purpose of future re-introduction into the wild[73]. As an illustration, important life-history traits of hatchery-reared fish can evolve in a direction such that they are adaptive in the hatchery environment, but harmful in the wild (for example, REF. 51) (FIG. 1c). This has obvious and important implications for management of fish populations.

Studying the genetic basis of evolutionary transitions. Combining Q_{ST}–F_{ST} comparisons with genomic investigations can provide deeper insights into the genetic underpinnings of evolutionary divergence. A good example of how the Q_{ST}–F_{ST} approach can be used first to demonstrate an adaptive genetic response to selection, and then to provide a platform for a more detailed investigation of the genetic architecture of the focal traits, is a study of bread wheat[74]. The authors combined spatial and temporal (across 12 generations) estimates of Q_{ST} and F_{ST} to establish that flowering time has evolved in response to selection, and they subsequently tested for divergent selection on candidate loci to identify the causal genes that underlie the evolution of flowering time (FIG. 1d). For many species, experimental manipulations are not possible and pedigree information is not available. In these cases, the increasing availability of genomic data could provide a way of estimating the quantitative genetic parameters that are needed for inferring the relative roles of natural selection and genetic drift in the observed divergence (see below).

Q_{ST}–F_{ST} trends. In line with the results of previous meta-analyses[9,11], the pattern emerging from a compilation of data from 143 published studies (Supplementary information S1 (table)) is that Q_{ST} generally exceeds F_{ST} (FIG. 2). Thus, although the variance in the data is large, directional natural selection seems to be the most common cause for divergence in many studied traits. The degree to which population differentiation in neutral marker genes is predictive of the degree of genetic differentiation in quantitative traits has been subject to debate[11,12,75]. Such a correlation would be expected on theoretical grounds: divergent selection that causes differentiation in quantitative trait loci (QTLs) can also lead to differentiation in neutral markers by restricting gene flow (which is termed 'isolation by adaptation'[76]). In line with earlier results based on fewer studies, there seems to be a positive relationship between Q_{ST} and F_{ST} (Spearman rank correlation = 0.24; $P < 0.001$), although at low values of F_{ST}, Q_{ST} tends to be generally higher than F_{ST}, and at high values of F_{ST}, Q_{ST} is generally lower than F_{ST} (FIG. 2). However, the relationship between the expected 'true'

Anadromous
Fish that spend most of their lives in the sea and migrate to fresh water to breed.

Quantitative trait loci
(QTLs). Segments of a chromosome affecting or linked to a quantitative trait.

Isolation by adaptation
A positive correlation between the degree of adaptive phenotypic and molecular genetic divergence among populations that is independent of the geographical distance separating the populations.

values of Q_{ST} and F_{ST} (Supplementary information S1 (table)) cannot be statistically differentiated from a 1:1 linear relationship (FIG. 2). Therefore, the degree of neutral marker differentiation may be a better predictor than previously thought of the degree of differentiation in genomic regions that underlie differentiation in adaptive traits[77,78], which is an observation that warrants further investigation.

Applying Q_{ST}–F_{ST} to 'omics' data

Measures of transcript abundance are essentially phenotypic data reflecting variability in levels of gene expression. As such, the Q_{ST}–F_{ST} framework lends itself to the analysis of expression data. Indeed, much of the impetus for early 'population transcriptomics' — in this context, the analyses of transcriptome-wide patterns of expression across multiple populations — was founded on the hypothesis that different levels of transcriptional variation within and among populations could be important determinants of local adaptation[79,80]. This was perhaps best demonstrated by Whitehead and Crawford[81], who showed that selection on gene expression could be inferred by comparing the ratio of among- and within-population variance in transcript abundance (FIG. 3). Although not strictly a Q_{ST}–F_{ST} analysis, this ratio of variance is related to the Q_{ST} index. Thus, it is somewhat surprising that to date, relatively few studies have subjected transcriptomic data to formal Q_{ST}–F_{ST} analyses. One likely reason for this is that Q_{ST}–F_{ST} analyses require a quantitative genetic approach, and therefore, transcriptomic data are needed from many individuals. In our review of the literature we only found three such studies[71,82,83], and even these suffer to some extent from the common problems that plague many Q_{ST}–F_{ST} comparisons. Nevertheless, their findings have revealed evidence of directional selection in the transcriptome, and perhaps more importantly, point to a means of studying how gene expression has evolved in a broader range of taxa. Thus, Q_{ST}–F_{ST} analyses may be a particularly expedient tool in the immediate future as the use of other high-throughput 'omics' data (such as proteomics, metabolomics and lipidomics) become more common.

The future of Q_{ST}–F_{ST}

Quantitative genetics was instrumental in the development of the 'modern synthesis' and continues to provide the basic framework for comprehending the evolution of complex traits, such as understanding inheritance and the genetic underpinnings of trait variability within populations[6,7,84]. Like other quantitative genetic approaches, Q_{ST} — which is itself a quantitative genetic parameter in essence — will continue to provide a means of understanding the causes and extent of genetic differentiation in complex polygenic traits. As such, potential applications of Q_{ST}–F_{ST} comparisons are varied and cover all areas of study in which genetic differentiation in polygenic traits is of interest. The recent methodological and analytical developments in quantitative genetics — spurred on by access to large amounts of genomic information — have the potential to bring about important refinements to Q_{ST}–F_{ST} comparisons. For example, access to large numbers of SNP markers opens the possibility of estimating quantitative genetic parameters without experimental crosses or access to recorded pedigrees[85–87]. Likewise, marker data can now be used to improve quantitative genetic parameter estimation through 'weighting' of the relationship matrix[88], which would improve the accuracy of Q_{ST} estimates. However, in the context of Q_{ST} estimation, the challenge with both of these approaches will be in obtaining unbiased estimates of the among-population genetic components of variance.

Recent advances in sequencing and related technologies are likely to revolutionize evolutionary and genetic research in many ways, but the pace of data acquisition risks outstripping that of theoretical developments. From the perspective of the evolutionary biologist, missing or incomplete null or neutral models for many omics data (for example, transcriptome and metabolome data) limit our understanding of how selection has shaped their evolution. Analyses that use Q_{ST}–F_{ST} comparisons provide a useful means of bridging this gap, although so far they have rarely been used in this context. One particularly interesting avenue might be the application of complementary analyses to RNA-seq data. With sufficient coverage of the transcriptome, such data can also be used to infer expression differences. Genes for which *cis*-regulatory elements are found within the untranslated regions (UTRs) could be analysed both by Q_{ST}–F_{ST} comparisons of transcript abundance and by metrics of

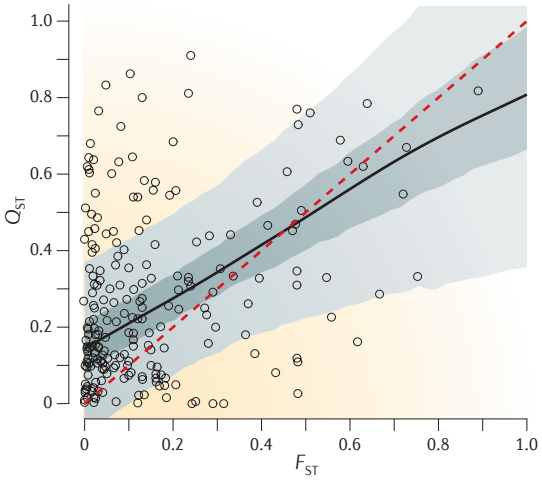


Figure 2 | **Comparison of mean Q_{ST} and F_{ST} estimates across published studies.** There is a significant non-parametric correlation (Spearman rank correlation coefficient = 0.24, $n = 218$, $P < 0.001$) between average Q_{ST} and F_{ST} estimates across all studies published to date. Moreover, the fitted relationship between F_{ST} and the expected 'true' value of Q_{ST} (see Supplementary information S1 (table)) does not significantly differ from a 1:1 relationship (dashed red line). The solid line denotes the posterior mode of predicted Q_{ST} estimated as a function of its relationship with F_{ST}, whereas the light grey and dark grey shaded areas denote the 50% and 95% posterior density intervals, respectively. Note that the 95% posterior density limits include the 1:1 line over the full range of possible F_{ST} values.

RNA-seq
High-throughput sequencing of cDNA.

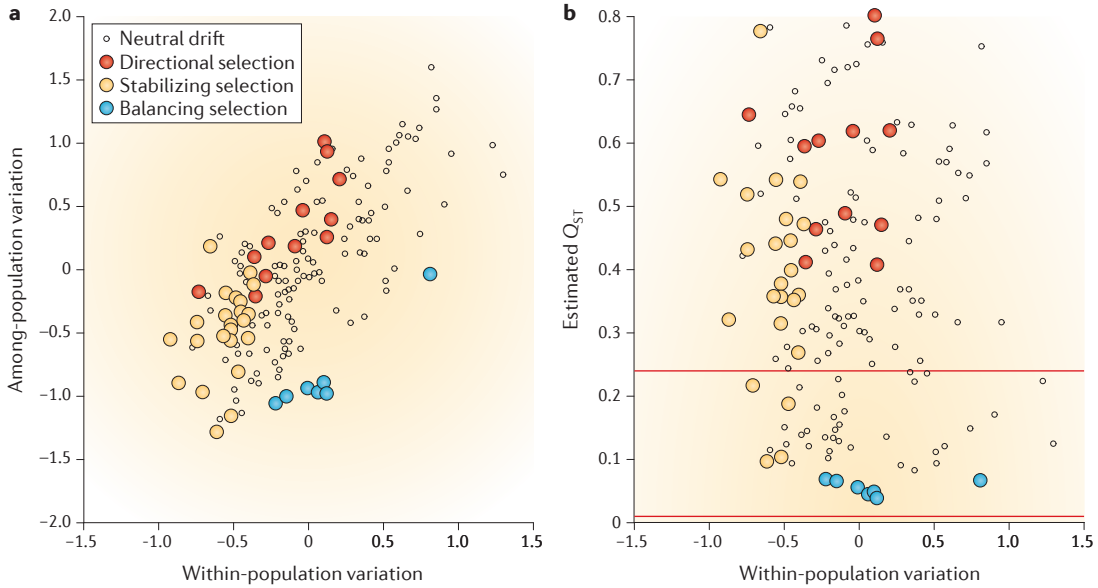


Figure 3 | **Inferring selection on gene expression.** Early comparisons of the ratio of among-population to within-population variance in transcription (**a**) were instrumental for inferring the adaptive importance of variation in gene expression. Although analogous to Q_{ST}, inference based on such an *F*-ratio comparison may be less robust. For example, many transcripts inferred to be neutrally divergent on the basis of overlapping variance ratios clearly exceed the range of neutral expectation that is defined by the reported range in F_{ST} (this range is indicated by the red lines in part **b**). Additionally, such an analysis may over-estimate the number of transcripts under stabilizing selection, as evidenced by associated Q_{ST} estimates being significantly greater than neutrality. It should be noted that the region defining neutral expectation is probably overly conservative given that it is based on the full range of reported F_{ST} values, rather than first screening for outlier loci and/or establishing a mean estimate bounded by confidence limits. As such, the correct identification of transcripts under stabilizing selection is probably obscured in part **b**. Data is derived, and figure is modified with permission, from REF. 81 © (2006) US National Academy of Sciences.

sequence divergence. In a narrow sense, such an analysis would be a considerable step towards overcoming one of the limitations of Q_{ST}–F_{ST} studies, namely that they cannot identify specific genomic regions that are under selection. In more general terms, such a connection between mechanistic and phenomenological aspects may yield fundamental insights into the proximate–ultimate distinction, which continues to influence biological thought[89].

The developing field of 'phenomics'[90,91] is also likely to benefit from a Q_{ST}-based analytical framework. Q_{ST}–F_{ST} comparisons could provide an expedient means to filter and classify traits that have been under different modes or strengths of selection. Multivariate Q_{ST}–F_{ST} methods are likely to be particularly useful in this respect, although the computational challenges might turn out to be formidable. In particular, numerical methods that are able to handle increasingly large variance–covariance matrices must be optimized: as the dimensionality of a matrix scales quadratically with the number of focal traits, some form of dimension reduction will be needed if inherently high-throughput endeavours such as phenome-wide analyses are to be tractable.

With sufficient attention to experimental design, Q_{ST} estimation can be fairly precise. However, as pointed out in earlier papers[9,14,15], the published Q_{ST} estimates and their standard errors suffer from considerable heterogeneity in quality and many inaccuracies. These problems can partly be traced back to a lack of ready-to-use software to estimate the parameters of interest. Given the enormous interest in Q_{ST}–F_{ST} comparisons, as reflected in an exponentially increasing body of work, there is an obvious need for reliable, publicly available applications. To some extent this is being addressed as authors begin to provide analytical scripts; these are typically codes for user-defined functions in the R computing language[15,18,45,47]. However, a fully integrated R package or standalone application would be helpful.

Finally, there is also a need for further theoretical work, particularly in two areas. The first is an investigation of the effects of possible negative bias in F_{ST} caused by high mutation rates on the Q_{ST}–F_{ST} comparison. The second is in understanding how Q_{ST} behaves under selection. Thus far there has been surprisingly little work on this (see REFS 92,93 for some rare examples), and we have little understanding of how different patterns of divergent selection and migration affect Q_{ST}–F_{ST} divergence.

In summary, although Q_{ST}–F_{ST} comparisons can be a reliable means of testing for adaptive population divergence, the onus remains on the researcher to ensure that the technique is properly applied. Ideally, experiments should be designed with the same rigour as any breeding plan used in quantitative genetics. Equal care and consideration should go into the choice of markers that are used to infer the neutral baseline against which Q_{ST} is ultimately compared. This should include a routine

Proximate–ultimate distinction
Proximate causation refers to biological functions in terms of physiological factors, whereas ultimate causation explains traits in terms of the evolutionary forces they are subjected to.

Phenomics
Large-scale phenotyping of the full set of phenotypes of individuals.

screening for outlier loci, and as computational solutions for modelling and estimating mutation rate become tractable, a screen to ensure that mutational variance will not bias estimation. Overall, it is clear that although expanding the theoretical underpinnings of Q_{ST} remains an open area for future research, the operational flexibility and documented successes of Q_{ST}–F_{ST} comparisons suggest that the application of this analytical framework can (and will) continue in the interim. This is likely to provide important insights into the selective processes shaping data types for which theoretical evolutionary models are unavailable and/or ambiguous.

1. Wright, S. Evolution in Mendelian populations. *Genetics* **16**, 97–159 (1931).
2. Holsinger, K. E. & Weir, B. S. Genetics in geographically structured populations: defining, estimating and interpreting F_{ST}. *Nature Rev. Genet.* **10**, 639–650 (2009).
3. Waples, R. S. & Gaggiotti, O. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol. Ecol.* **15**, 1419–1439 (2006).
4. Lewontin, R. C. & Krakauer, J. Distribution of gene frequency as a test of the theory of selective neutrality of polymorphisms. *Genetics* **74**, 175–195 (1973).
5. Beaumont, M. A. Adaptation and speciation: what can F_{st} tell us? *Trends Ecol. Evol.* **20**, 435–440 (2005).
6. Mackay, T. F. C., Stone, E. A. & Ayroles, J. F. The genetics of quantitative traits: challenges and prospects. *Nature Rev. Genet.* **10**, 565–577 (2009).
7. Hill, W. G. & Kirkpatrick, M. What animal breeding has taught us about evolution. *Annu. Rev. Ecol. Evol. Systemat.* **41**, 1–19 (2010).
8. Wright, S. The genetic structure of populations. *Ann. Eugen.* **15**, 323–354 (1951).
9. Leinonen, T., O'Hara, R. B., Cano, J. M. & Merilä, J. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *J. Evol. Biol.* **21**, 1–17 (2008).
A second meta-analysis of Q_{ST}–F_{ST} studies with an extended historical narrative of the development of comparative marker and quantitative genetic studies.
10. Spitze, K. Population-structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics* **135**, 367–374 (1993).
The study in which the term Q_{ST} was coined to denote the proportion of variance in quantitative traits residing among populations.
11. Merilä, J. & Crnokrak, P. Comparison of genetic differentiation at marker loci and quantitative traits. *J. Evol. Biol.* **14**, 892–903 (2001).
The first meta-analysis and review of Q_{ST}–F_{ST} studies.
12. McKay, J. K. & Latta, R. G. Adaptive population divergence: markers, QTL and traits. *Trends Ecol. Evol.* **17**, 285–291 (2002).
13. Lande, R. Neutral theory of quantitative genetic variance in an island model with local extinction and colonization. *Evolution* **46**, 381–389 (1992).
14. Brommer, J. E. Whither P_{ST}? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. *J. Evol. Biol.* **24**, 1160–1168 (2011).
15. O'Hara, R. B. & Merilä, J. Bias and precision in Q_{ST} estimates: problems and some solutions. *Genetics* **171**, 1331–1339 (2005).
16. Kremer, A., Zanetto, A. & Ducousso, A. Multilocus and multitrait measures of differentiation for gene markers and phenotypic traits. *Genetics* **145**, 1229–1241 (1997).
17. Chenoweth, S. F. & Blows, M. W. Q_{ST} meets the G matrix: the dimensionality of adaptive divergence in multiple correlated quantitative traits. *Evolution* **62**, 1437–1449 (2008).
18. Martin, G., Chapuis, E. & Goudet, J. Multivariate Q_{ST}-F_{ST} comparisons: a neutrality test for the evolution of the G matrix in structured populations. *Genetics* **180**, 2135–2149 (2008).
19. Ovaskainen, O., Karhunen, M., Zheng, C. Z., Arias, J. M. C. & Merilä, J. A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. *Genetics* **189**, 621–632 (2011).
20. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226 (1983).
21. Volis, S., Yakubov, B., Shulgina, I., Ward, D. & Mendlinger, S. Distinguishing adaptive from nonadaptive genetic differentiation: comparison of Q_{ST} and F_{ST} at two spatial scales. *Heredity* **95**, 466–475 (2005).
22. Manier, M. K., Seyler, C. M. & Arnold, S. J. Adaptive divergence within and between ecotypes of the terrestrial garter snake, *Thamnophis elegans*, assessed with F_{ST}-Q_{ST} comparisons. *J. Evol. Biol.* **20**, 1705–1719 (2007).
23. Whitlock, M. C. & Gilbert, K. J. Q_{ST} in a hierarchically structured population. *Mol. Ecol. Resources* **12**, 481–483 (2012).
24. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population-structure. *Evolution* **38**, 1358–1370 (1984).
25. Whitlock, M. C. Evolutionary inference from Q_{ST}. *Mol. Ecol.* **17**, 1885–1896 (2008).
26. Meirmans, P. G. & Hedrick, P. W. Assessing population structure: F_{ST} and related measures. *Mol. Ecol. Resources* **11**, 5–18 (2011).
27. Kronholm, I., Loudet, O. & de Meaux, J. Influence of mutation rate on estimators of genetic differentiation - lessons from *Arabidopsis thaliana*. *BMC Genet.* **11**, 33 (2010).
28. Storz, J. F. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.* **14**, 671–688 (2005).
29. Foll, M. & Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977–993 (2008).
30. Excoffier, L., Hofer, T. & Foll, M. Detecting loci under selection in a hierarchically structured population. *Heredity* **103**, 285–298 (2009).
31. Pavlidis, P., Jensen, J. D., Stephan, W. & Stamatakis, A. A critical assessment of storytelling: Gene Ontology categories and the importance of validating genomic scans. *Mol. Biol. Evol.* **29**, 3237–3248 (2012).
32. Thornton, K. R. & Jensen, J. D. Controlling the false-positive rate in multilocus genome scans for selection. *Genetics* **175**, 737–750 (2007).
33. Narum, S. R. & Hess, J. E. Comparison of F_{ST} outlier tests for SNP loci under selection. *Mol. Ecol. Resources* **11**, 184–194 (2011).
34. Pérez-Figueroa, A., García-Pereira, M. J., Saura, M., Rolán-Alvarez, E. & Caballero, A. Comparing three different methods to detect selective loci using dominant markers. *J. Evol. Biol.* **23**, 2267–2276 (2010).
35. Vilas, A., Pérez-Figueroa, A. & Caballero, A. A simulation study on the performance of differentiation-based methods to detect selected loci using linked neutral markers. *J. Evol. Biol.* **25**, 1364–1376 (2012).
36. Edelaar, P., Burraco, P. & Gomez-Mestre, I. Comparisons between Q_{ST} and F_{ST} - how wrong have we been? *Mol. Ecol.* **20**, 4830–4839 (2011).
37. Edelaar, P. & Björklund, M. If F_{ST} does not measure neutral genetic differentiation, then comparing it with Q_{ST} is misleading. Or is it? *Mol. Ecol.* **20**, 1805–1812 (2011).
38. Lercher, M. J. & Hurst, L. D. Human SNP variability and mutation rate are higher in regions of high recombination. *Trends Genet.* **18**, 337–340 (2002).
39. Hodgkinson, A., Ladoukakis, E. & Eyre-Walker, A. Cryptic variation in the human mutation rate. *PLoS Biol.* **7**, 226–232 (2009).
40. Le Corre, V. & Kremer, A. The genetic differentiation at quantitative trait loci under local adaptation. *Mol. Ecol.* **21**, 1548–1566 (2012).
An excellent theoretical treatment and review of factors influencing the utility of Q_{ST}–F_{ST} comparisons.
41. Albrechtsen, A., Nielsen, F. C. & Nielsen, R. Ascertainment biases in SNP chips affect measures of population divergence. *Mol. Biol. Evol.* **27**, 2534–2547 (2010).
42. Helyar, S. J. *et al.* Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Mol. Ecol. Resources* **11**, 123–136 (2011).
43. Porcher, E., Giraud, T., Goldringer, I. & Lavigne, C. Experimental demonstration of a causal relationship between heterogeneity of selection and genetic differentiation in quantitative traits. *Evolution* **58**, 1434–1445 (2004).
44. Morgan, T. J., Evans, M. A., Garland, T., Swallow, J. G. & Carter, P. A. Molecular and quantitative genetic divergence among populations of house mice with known evolutionary histories. *Heredity* **94**, 518–525 (2005).
45. Whitlock, M. C. & Guillaume, F. Testing for spatially divergent selection: comparing Q_{ST} to F_{ST}. *Genetics* **183**, 1055–1063 (2009).
46. Hendry, A. P. $Q_{ST} > = \neq < F_{ST}$? *Trends Ecol. Evol.* **17**, 502 (2002).
47. Lind, M. I., Ingvarsson, P. K., Johansson, H., Hall, D. & Johansson, F. Gene flow and selection on phenotypic plasticity in an island system of *Rana temporaria*. *Evolution* **65**, 684–697 (2011).
48. Alberto, F. *et al.* Adaptive responses for seed and leaf phenology in natural populations of sessile oak along an altitudinal gradient. *J. Evol. Biol.* **24**, 1442–1454 (2011).
49. Holand, A. M., Jensen, H., Tufto, J. & Moe, R. Does selection or genetic drift explain geographic differentiation of morphological characters in house sparrows *Passer domesticus*? *Genet. Res.* **93**, 367–379 (2011).
50. Lamy, J.-B. *et al.* Uniform selection as a primary force reducing population genetic differentiation of cavitation resistance across a species range. *PLoS ONE* **6**, e23476 (2011).
51. Rogell, B. *et al.* Strong divergence in trait means but not in plasticity across hatchery and wild populations of sea-run brown trout *Salmo trutta*. *Mol. Ecol.* **21**, 2963–2976 (2012).
52. Whitlock, M. C. Neutral additive genetic variance in a metapopulation. *Genet. Res.* **74**, 215–221 (1999).
53. Goudet, J. & Büchi, L. The effects of dominance, regular inbreeding and sampling design on Q_{ST}, an estimator of population differentiation for quantitative traits. *Genetics* **172**, 1337–1347 (2006).
54. Goudet, J. & Martin, G. Under neutrality, $Q_{ST} \leq F_{ST}$ when there is dominance in an island model. *Genetics* **176**, 1371–1374 (2007).
55. López-Fanjul, C., Fernández, A. & Toro, M. A. The effect of neutral nonadditive gene action on the quantitative index of population divergence. *Genetics* **164**, 1627–1633 (2003).
56. López-Fanjul, C., Fernández, A. & Toro, M. A. The effect of dominance on the use of the Q_{ST}-F_{ST} contrast to detect natural selection on quantitative traits. *Genetics* **176**, 725–727 (2007).
57. Halligan, D. L. & Keightley, P. D. Spontaneous mutation accumulation studies in evolutionary genetics. *Annu. Rev. Ecol. Evol. Systemat.* **40**, 151–172 (2009).
58. Johnson, T. & Barton, N. Theoretical models of selection and mutation on quantitative traits. *Phil. Trans. R. Soc. B* **360**, 1411–1425 (2005).
59. Ellegren, H. Microsatellites: simple sequences with complex evolution. *Nature Rev. Genet.* **5**, 435–445 (2004).
60. Houle, D., Morikawa, B. & Lynch, M. Comparing mutational variabilities. *Genetics* **143**, 1467–1483 (1996).
61. Caballero, A., Keightley, P. D. & Turelli, M. Average dominance for polygenes: drawbacks of regression estimates. *Genetics* **147**, 1487–1490 (1997).
62. Zhan, J. *et al.* Variation for neutral markers is correlated with variation for quantitative traits in the plant pathogenic fungus *Mycosphaerella graminicola*. *Mol. Ecol.* **14**, 2683–2693 (2005).
63. Zhan, J., Stefanato, F. L. & McDonald, B. A. Selection for increased cyproconazole tolerance in *Mycosphaerella graminicola* through local adaptation and in response to host resistance. *Mol. Plant Pathol.* **7**, 259–268 (2006).

64. Rogers, A. R. & Harpending, H. C. Population structure and quantitative characters. *Genetics* **105**, 985–1002 (1983).
65. Rogers, S. M., Gagnon, V. & Bernatchez, L. Genetically based phenotype-environment association for swimming behavior in lake whitefish ecotypes (*Coregonus clupeaformis* Mitchill). *Evolution* **56**, 2322–2329 (2002).
66. Saint-Laurent, R., Legault, M. & Bernatchez, L. Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchill). *Mol. Ecol.* **12**, 315–330 (2003).
67. Savolainen, O., Pyhäjärvi, T. & Knürr, T. Gene flow and local adaptation in trees. *Annu. Rev. Ecol. Evol. Systemat.* **38**, 595–619 (2007).
Contains an excellent review of Q_{ST}–F_{ST} studies carried out in forest trees.
68. Keller, S. R. & Taylor, D. R. History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecol. Lett.* **11**, 852–866 (2008).
69. Johansson, M., Primmer, C. R. & Merilä, J. Does habitat fragmentation reduce fitness and adaptability? A case study of the common frog (*Rana temporaria*). *Mol. Ecol.* **16**, 2693–2700 (2007).
70. Cousyn, C. *et al.* Rapid, local adaptation of zooplankton behavior to changes in predation pressure in the absence of neutral genetic changes. *Proc. Natl Acad. Sci. USA* **98**, 6256–6260 (2001).
71. Aykanat, T., Thrower, F. P. & Heath, D. D. Rapid evolution of osmoregulatory function by modification of gene transcription in steelhead trout. *Genetica* **139**, 233–242 (2011).
72. Koskinen, M. T., Haugen, T. O. & Primmer, C. R. Contemporary Fisherian life-history evolution in small salmonid populations. *Nature* **419**, 826–830 (2002).
73. Pelletier, F., Réale, D., Watters, J., Boakes, E. H. & Garant, D. Value of captive populations for quantitative genetics research. *Trends Ecol. Evol.* **24**, 263–270 (2009).
74. Rhoné, B., Vitalis, R., Goldringer, I. & Bonnin, I. Evolution of flowering time in experimental wheat populations: a comprehensive approach to detect genetic signatures of natural selection. *Evolution* **64**, 2110–2125 (2010).
75. Crnokrak, P. & Merilä, J. Genetic population divergence: markers and traits. *Trends Ecol. Evol.* **17**, 501–501 (2002).
76. Nosil, P., Funk, D. J. & Ortiz-Barrientos, D. Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375–402 (2009).
77. Crnokrak, P. & Roff, D. A. Dominance variance - associations with selection and fitness. *Heredity* **75**, 530–540 (1995).
78. Latta, R. G. & McKay, J. K. Genetic population divergence: markers and traits - response. *Trends Ecol. Evol.* **17**, 501–502 (2002).
79. Oleksiak, M. F., Churchill, G. A. & Crawford, D. L. Variation in gene expression within and among natural populations. *Nature Genet.* **32**, 261–266 (2002).
80. Whitehead, A. & Crawford, D. L. Variation within and among species in gene expression: raw material for evolution. *Mol. Ecol.* **15**, 1197–1211 (2006).
81. Whitehead, A. & Crawford, D. L. Neutral and adaptive variation in gene expression. *Proc. Natl Acad. Sci. USA* **103**, 5425–5430 (2006).
82. Kohn, M. H., Shapiro, J. & Wu, C. I. Decoupled differentiation of gene expression and coding sequence among *Drosophila* populations. *Genes Genet. Systems* **83**, 265–273 (2008).
83. Roberge, C., Guderley, H. & Bernatchez, L. Genomewide identification of genes under directional selection: gene transcription Q_{ST} scan in diverging Atlantic salmon subpopulations. *Genetics* **177**, 1011–1022 (2007).
The first formal application of Q_{ST}-based inference on transcriptome-wide data.
84. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era - concepts and misconceptions. *Nature Rev. Genet.* **9**, 255–266 (2008).
85. Visscher, P. M. *et al.* Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet.* **2**, e41 (2006).
86. Visscher, P. M. Whole genome approaches to quantitative genetics. *Genetica* **136**, 351–358 (2009).
87. Deary, I. J. *et al.* Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* **482**, 212–215 (2012).
88. Lee, S. H., Goddard, M. E., Visscher, P. M. & van der Werf, J. H. J. Using the realized relationship matrix to disentangle confounding factors for the estimation of genetic variance components of complex traits. *Genet. Selection Evol.* **42**, 22 (2010).
89. Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W. & Uller, T. Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? *Science* **334**, 1512–1516 (2011).
90. Houle, D. Numbering the hairs on our heads: the shared challenge and promise of phenomics. *Proc. Natl Acad. Sci. USA* **107**, 1793–1799 (2010).
91. Houle, D., Govindaraju, D. R. & Omholt, S. Phenomics: the next challenge. *Nature Rev. Genet.* **11**, 855–866 (2010).
92. Le Corre, V. & Kremer, A. Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics* **164**, 1205–1219 (2003).
93. Santure, A. W. & Wang, J. L. The joint effects of selection and dominance on the Q_{ST}-F_{ST} contrast. *Genetics* **181**, 259–276 (2009).
94. Leinonen, T., Cano, J. M., Mäkinen, H. & Merilä, J. Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *J. Evol. Biol.* **19**, 1803–1812 (2006).
The first study to use the term P_{ST} to refer to the phenotypic equivalent of Q_{ST}.
95. Pujol, B., Wilson, A. J., Ross, R. I. C. & Pannell, J. R. Are Q_{ST}-F_{ST} comparisons for natural populations meaningful? *Mol. Ecol.* **17**, 4782–4785 (2008).
96. Lee, C. E. & Frost, B. W. Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). *Hydrobiologia* **480**, 111–128 (2002).
97. Merilä, J. Quantitative trait and allozyme divergence in the greenfinch (*Carduelis chloris*, Aves: Fringillidae). *Biol. J. Linnean Soc.* **61**, 243–266 (1997).
98. Conover, D. O., Duffy, T. A. & Hice, L. A. The covariance between genetic and environmental influences across ecological gradients reassessing the evolutionary significance of countergradient and cogradient variation. *Ann. NY Acad. Sci.* **1168**, 100–129 (2009).
99. Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334 (1976).
100. Lande, R. Statistical tests for natural selection on quantitative characters. *Evolution* **31**, 442–444 (1977).
101. Lynch, M. & Hill, W. G. Phenotypic evolution by neutral mutation. *Evolution* **40**, 915–935 (1986).
102. Lynch, M. Neutral models of phenotypic evolution. in *Ecological Genetics* (ed. Read, A. L.) 86–108 (Princeton Univ. Press, 1994).
103. Lynch, M. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Naturalist* **136**, 727–741 (1990).
104. Baker, A. J. Genetic and morphometric divergence in ancestral European and descendent New Zealand populations of chaffinches (*Fringilla coelebs*) *Evolution* **46**, 1784–1800 (1992).
105. Egea, R., Casillas, S. & Barbadilla, A. Standard and generalized McDonald-Kreitman test: a website to detect selection by comparing different classes of DNA sites. *Nucleic Acids Res.* **36**, W157–W162 (2008).
106. Luikart, G., England, P. R., Tallmon, D., Jordan, S. & Taberlet, P. The power and promise of population genomics: from genotyping to genome typing. *Nature Rev. Genet.* **4**, 981–994 (2003).
107. Latta, R. G. Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *Am. Naturalist* **151**, 283–292 (1998).
108. Pritchard, J. K. & Di Rienzo, A. Adaptation - not by sweeps alone. *Nature Rev. Genet.* **11**, 665–667 (2010).
109. Alho, J. S. *et al.* Allen's rule revisited: quantitative genetics of extremity length in the common frog along a latitudinal gradient. *J. Evol. Biol.* **24**, 59–70 (2011).
110. Antoniazza, S., Burri, R., Fumagalli, L., Goudet, J. & Roulin, A. Local adaptation maintains clinal variation in melanin-based coloration of European barn owls (*Tyto alba*). *Evolution* **64**, 1944–1954 (2010).
111. Kawakami, T. *et al.* Natural selection drives clinal life history patterns in the perennial sunflower species, *Helianthus maximiliani*. *Mol. Ecol.* **20**, 2318–2328 (2011).
112. Yu, Q., Ellen, E. D., Wade, M. J. & Delph, L. F. Genetic differences among populations in sexual dimorphism: evidence for selection on males in a dioecious plant. *J. Evol. Biol.* **24**, 1120–1127 (2011).
113. Magalhaes, I. S., Mwaiko, S., Schneider, M. V. & Seehausen, O. Divergent selection and phenotypic plasticity during incipient speciation in Lake Victoria cichlid fish. *J. Evol. Biol.* **22**, 260–274 (2009).
114. Gay, L. *et al.* Speciation with gene flow in the large white-headed gulls: does selection counterbalance introgression? *Heredity* **102**, 133–146 (2009).
115. Wojcieszek, J. M. & Simmons, L. W. Evidence for stabilizing selection and slow divergent evolution of male genitalia in a millipede (*Acatichiropus variabilis*). *Evolution* **66**, 1138–1153 (2012).
116. Jiménez-Ambriz, G. *et al.* Life history variation in the heavy metal tolerant plant *Thlaspi caerulescens* growing in a network of contaminated and noncontaminated sites in southern France: role of gene flow, selection and phenotypic plasticity. *New Phytol.* **173**, 199–215 (2007).
117. Meyer, C.-L. *et al.* Variability of zinc tolerance among and within populations of the pseudometallophyte species *Arabidopsis halleri* and possible role of directional selection. *New Phytol.* **185**, 130–142 (2010).
118. Sreejayan, Kumar, U. S., Varghese, G., Jacob, T. M. & Thomas, G. Stratification and population structure of the genetic resources of ancient medicinal rice (*Oryza sativa* L.) landrace Njavara. *Genet. Resources Crop Evol.* **58**, 697–711 (2011).
119. Pressoir, G. & Berthaud, J. Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity* **92**, 95–101 (2004).
120. McKay, J. K. *et al.* Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proc. R. Soc. Lond. B.* **268**, 1715–1721 (2001).
121. Gravuer, K., von Wettberg, E. & Schmitt, J. Population differentiation and genetic variation inform translocation decisions for *Liatris scariosa* var. *novae-angliae*, a rare New England grassland perennial. *Biol. Conserv.* **124**, 155–167 (2005).
122. Price, A. L. *et al.* Effects of *cis* and *trans* genetic ancestry on gene expression in African Americans. *PLoS Genet.* **4**, e1000294 (2008).

Acknowledgements
We thank M. Karhunen for helpful comments on the manuscript. Financial support was provided by the Academy of Finland (grants 250435 to J.M., 252597 to T.L. and 259944 to R.J.S.M.), the Finnish Cultural Foundation (grant to R.J.S.M.) and the Landesoffensive zur Entwicklung wissenschaftlich-ökonomischer Exzellenz" (LOEWE) of the state Hesse in Germany through the Biodiversity and Climate Research Centre (Bik-F) in Frankfurt am Main (to R.B.O.).

Competing interests statement
The authors declare no competing financial interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (table)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF