

## INVITED REVIEWS AND META-ANALYSES

# The genetic differentiation at quantitative trait loci under local adaptation

VALÉRIE LE CORRE\* and ANTOINE KREMER†‡

\*INRA, UMR1347 Agroécologie, F-21000 Dijon, France, †INRA, UMR1202 Biodiversité Gènes et Communautés, F-33610 Cestas, France, ‡UMR1202 Biodiversité Gènes et Communautés, Université de Bordeaux, F-33405 Talence, France

## Abstract

Most adaptive traits are controlled by large number of genes that may all together be the targets of selection. Adaptation may thus involve multiple but not necessarily substantial allele frequency changes. This has important consequences for the detection of selected loci and implies that a quantitative genetics framework may be more appropriate than the classical 'selective sweep' paradigm. Preferred methods to detect loci involved in local adaptation are based on the detection of 'outlier' values of the allelic differentiation  $F_{ST}$ . A quantitative genetics framework is adopted here to review theoretical expectations for how allelic differentiation at quantitative trait loci ( $F_{STQ}$ ) relates to (i), neutral genetic differentiation ( $F_{ST}$ ) and (ii), phenotypic differentiation ( $Q_{ST}$ ). We identify cases where results of outlier-based methods are likely to be poor and where differentiation at selected loci conveys little information regarding local adaptation. A first case is when neutral differentiation is high, so that local adaptation does not necessitate increased differentiation. A second case is when local adaptation is reached via an increased covariance of allelic effects rather than via allele frequency changes, which is more likely under high gene flow when the number of loci is high and selection is recent. The comparison of theoretical predictions with observed data from the literature suggests that polygenic local adaptation involving only faint allele frequency changes are very likely in some species such as forest trees and for climate-related traits. Recent methodological improvements that may alleviate the weakness of  $F_{ST}$ -based detection methods are presented.

**Keywords:**  $F_{ST}$ ,  $F_{STQ}$ , local adaptation,  $Q_{ST}$ , quantitative trait, selected loci

Received 29 August 2011; revision received 19 December 2011; accepted 21 December 2011

## Introduction

There is local adaptation when natural selection causes local populations to evolve divergent phenotypic trait values, thus allowing an organism to better suit the heterogeneity of biotic and abiotic conditions. The survey of published reciprocal transplant studies on plants and animals suggests that local adaptation is common, although it may be prevented by gene flow and genetic drift or constrained by lack of genetic variation or large phenotypic plasticity (Savolainen *et al.* 2007; Hereford 2009; Fraser *et al.* 2011). Numerous traits and combinations of traits may be involved in local adaptation. However, as argued by Pritchard & Di Rienzo (2010), it

seems likely that many, or most, adaptive events in natural populations occur by the evolution of polygenic traits, rather than via the fixation of single beneficial mutations (also known as 'selective sweep'). Recent genome-wide association studies in humans (Yang *et al.* 2010; Stranger *et al.* 2011) and in various model organisms (e.g. in *Arabidopsis thaliana*: Atwell *et al.* 2010; in *Drosophila melanogaster*: Jumbo-Lucioni *et al.* 2010) have indeed confirmed that variation at many important traits is controlled by a large number of loci dispersed throughout the genome.

Polygenic adaptation typically involves small allele frequency changes at many loci (Mackay *et al.* 2009). Therefore, quantitative trait loci (QTL) involved in polygenic adaptation will go undetected using methods for detecting the molecular signature of selective sweeps at individual loci (Pritchard & Di Rienzo 2010). Several

Correspondence: Valérie Le Corre, Fax: +33 (0) 380693222; E-mail: lecorre@dijon.inra.fr

methods have been devised that aim at specifically detecting loci where allele frequency shifts have occurred in response to local selection. These methods are based on the identification of 'outlier' loci that exhibit higher genetic differentiation among populations than expected under a model of neutral evolution (Luikart *et al.* 2003; Storz 2005). In most cases, the rationale of these outlier detection methods is the existence of large phenotypic differentiation of the selected trait as observed in common garden experiments, although they were sometimes conducted in cases where the adaptive syndrome was unknown or cryptic (Bonin *et al.* 2006). Most studies use  $F_{ST}$  as a measure of allelic differentiation (Beaumont 2005) and employ a 'genome scan' approach where a large set of molecular markers is screened for the presence of outliers (Luikart *et al.* 2003). The detected outliers can be the loci under selection themselves, or, more probably, linked neutral loci (Nosil *et al.* 2009). As an alternative or complement to genome scan when genetic determinants for the locally selected trait are known, differentiation statistics at candidate genes can be compared with differentiation at a random set of putatively neutral markers (Stinchcombe & Hoekstra 2008).

The expected benefits and caveats of genome scans have been well covered in previous review papers (Luikart *et al.* 2003; Beaumont 2005; Storz 2005; Holderegger *et al.* 2008; Stinchcombe & Hoekstra 2008). The efficiency of different methods based on  $F_{ST}$  outlier tests has been evaluated using simulated and experimental data sets (Beaumont & Balding 2004; Foll & Gaggiotti 2008; Excoffier *et al.* 2009; Pérez-Figueroa *et al.* 2010; Narum & Hess 2011; Nunes *et al.* 2011). In brief, it is largely recognized that the main drawback of outlier detection methods stems from the large expected variability in  $F_{ST}$  among neutral loci. Complex demographic histories and population structures not accounted for by the model used for deriving expected neutral  $F_{ST}$  values may lead to a high rate of false positive. The rate of false negatives, that is, the power of outlier detection methods, has been less often considered than the rate of false positives. It has been stressed that selected loci will not be easily detected if selection is weak in comparison with migration (Beaumont & Balding 2004). On the other hand, Pérez-Figueroa *et al.* (2010) cautioned that the power to discriminate high outlier  $F_{ST}$  values is low when the neutral  $F_{ST}$  is itself high.

It is important to emphasize that the above studies did not explicitly consider selected loci involved in polygenic adaptation. Further, the power of detecting selected loci was never linked with the level of phenotypic divergence of the selected trait. This is certainly because there is no simple explicit population genetics framework to predict how local selection shapes pat-

terns of allelic variation at QTL that underlie quantitative traits (Turelli & Barton 1990). The existing body of theory suggests that there can be a large discrepancy between the phenotypic differentiation at a selected trait and the allelic differentiation at the QTL underlying the trait. As shown by Latta (1998, 2004) and McKay & Latta (2002), this is because part of the adaptive divergence in trait values is attributable to covariance in allele frequencies among individual QTL. The contribution of covariance to the overall phenotypic differentiation has further been explored by Le Corre & Kremer (2003) at the theoretical level. They confirmed that adaptive divergence at a selected trait can be observed in the absence of marked allelic differentiation at the underlying QTL. As emphasized by Storz (2005), this result has important implications for the prospect of using genome scans to identify selected loci and needs further examination. There has been a recent renewal of interest in the determinants of adaptive divergence in the context of climate change (Reusch & Wood 2007; Hoffmann & Willi 2008; Rasanen & Hendry 2008; Kavanagh *et al.* 2010). It is therefore timely to revisit the consequences of local selection in a comparative way between traits and their underlying genes.

The aim of this paper is to provide a review of the consequences of local selection at the level of genes underlying adaptive quantitative traits: What is the expected level of allelic differentiation at QTL in comparison with the neutral differentiation? How informative is the value of differentiation at QTL regarding the amount of adaptive trait divergence? We will first present a general relationship between allelic differentiation at QTL and phenotypic differentiation. The interacting effects of local selection and migration on differentiation will be addressed using some classical quantitative genetics models and simulation studies that consider the divergence of a locally adapted trait under a subdivided population framework. The main characteristics and assumptions of the models used are presented in Box 1. Departures from these models such as varying genetic architecture (number of QTL and their non-additive effects) and non-equilibrium situations will be dealt. Finally, we will confront theoretical expectations with experimental data available in the literature.

### **A general relationship between phenotypic differentiation and genetic differentiation at selected loci**

We compare the impact of divergent selection on phenotypic traits and their underlying genes by considering different populations connected by gene flow and undergoing stabilizing selection towards an

**Box 1. Genetic models of population divergence at a quantitative trait**

Consider a set of  $n$  populations, each of  $N$  diploid individuals, connected by migration (e.g. according to an individual-based finite island model). Mating occurs at random within each population and generations are discrete and non-overlapping. The genotypic value  $G$  of an individual is determined by its alleles at a number of loci. If additivity within and between loci is assumed, genotypic values are simply the sum over loci of allelic effects on the trait. The phenotypic value  $Z$  of an individual is the sum of its genotypic value and an independent random environmental effect distributed normally with mean 0 and variance  $V_E$ . Stabilizing selection within each population is modelled using a Gaussian-like fitness function (Haldane 1954):

$$W_i(Z) = \exp \left[ \frac{-(Z - Z_{\text{opt}i})^2}{2\omega^2} \right]$$

where  $Z_{\text{OPT}i}$  is the optimal value for the trait in population  $i$  and  $\omega^2$  is the intensity of selection.  $\omega^2$  values are assumed to be the same for all populations. The  $Z_{\text{OPT}i}$  values may either be identical for all populations, which corresponds to uniform selection, or vary among populations, which corresponds to divergent selection.

*The single-locus, two-allele model*

In that simplistic model, the trait is encoded by a single locus with two alleles having equal opposite effects, so that, in the absence of dominance, the three genotypic values are  $-a$ , 0 and  $a$ . The genetic variance for the trait is the variance of genotypic values:  $V_g = 2pq a^2$ . There is no analytical expression for allele frequencies at migration–selection equilibrium, but these can be derived using recurrence equations (Santure & Wang 2009).

*Multilocus models*

The genetic value of an additive trait is the sum of genotypic values over  $N_L$  loci. Hence, its genetic variance is a function of both the variance of genotypic values and their covariance among loci. A particular case is the diallelic symmetrical model, which is an extension of the single-locus, two-allele model presented above to  $N_L$  loci. Allele frequencies and genetic variances under migration and selection for multilocus models cannot be solved analytically and have been studied using simulations (Latta 1998; Le Corre & Kremer 2003; Lopez *et al.* 2008; Santure & Wang 2009).

*The infinitesimal model*

In the Fisher–Bulmer infinitesimal model (Fisher 1918; Bulmer 1980), the quantitative trait is determined by an infinite number of unlinked loci, with the allelic effects being additive and infinitesimally small. Under this model, selection at each individual locus is so weak that it causes only an infinitesimal change in allele frequencies. Therefore, selection induces no change in the variance of genotypic values. Response to selection is entirely because of the generation of covariance of genotypic values among loci. Assumptions of the infinitesimal model, and, specifically, the normality of genotypic values, allow solving for the values of genetic variances at migration–selection equilibrium using recurrence equations. Migration may however induce some deviation from normality, as it mixes individuals with different mean phenotypic values. Therefore, assumptions of the infinitesimal model may be violated when migration is strong and the phenotypic deviation of migrants is large (Tufto 2000; Hendry *et al.* 2001; Lopez *et al.* 2008).

optimal value within each population. Local selection is uniform when all populations share the same optimal phenotypic value, or divergent when phenotypic optima vary among populations. Selection is thus characterized by (i), its intensity at the within-population level and (ii), divergence of optima at the between-

population level. Stabilizing selection is typically modelled as a Gaussian-type curve, where selection intensity is measured by the variance parameter  $\omega^2$ , selection being weaker as  $\omega^2$  increases (see Box 1). Divergence of optima is simply measured as the variance of optimal phenotypic values among populations,

$V_{\text{OPT}}$ . As compared with neutral migration–drift equilibrium, local divergent selection is expected to decrease genetic variance for the trait at the within-population level ( $V_W$ ) and increase genetic variance among populations ( $V_B$ ).

Genetic differentiation at the trait is measured by  $Q_{\text{ST}}$ , a parameter analogous to the single-locus genetic differentiation measure  $F_{\text{ST}}$  (Spitze 1993; Whitlock 2008):

$$Q_{\text{ST}} = \frac{V_B}{V_B + 2V_W},$$

where  $V_W$  and  $V_B$  are the additive genetic variance for the trait within a population and among populations, respectively.

The genetic variance for a multilocus trait is made of two distinct contributions: first, contribution from the variance of allelic effects at each locus  $i$ , called the genic variance; second, a contribution arising from the covariance of allelic effects among pairs of loci:

$$V = \sum_i \sigma_i^2 + \sum_i \sum_j \text{Cov}_{ij}$$

Following Gavrilets & Hastings (1995), Le Corre & Kremer (2003) introduced the parameter  $\theta$  as the ratio of these two contributions:

$$\theta = \frac{\sum_i \sum_j \text{Cov}_{ij}}{\sum_i \sigma_i^2}$$

Covariance of allelic effects arises from two causes (Bulmer 1980, 1989). First, linkage disequilibrium *sensu stricto* may contribute to a covariance between effects of alleles carried on a same gamete. Second, under non-random gametic association at reproduction (Hardy–Weinberg disequilibrium), covariance may arise between alleles of different gametic origins at a same locus or at different loci. In a single population after random mating, this latter contribution is zero. However, it is non-zero at the level of a subdivided population. For simplicity, the two covariance contributions will not be considered independently. Rather, the parameter  $\theta$ , which sums up both contributions, will be considered throughout the paper.

The following relationship between differentiation for a quantitative trait ( $Q_{\text{ST}}$ ) and the mean genetic differentiation at the  $n$  loci that additively control the trait ( $F_{\text{STQ}}$ ) was established by Le Corre & Kremer (2003):

$$Q_{\text{ST}} = \frac{(1 + \theta_B) F_{\text{STQ}}}{(\theta_B - \theta_W) F_{\text{STQ}} + 1 + \theta_W} \quad (1)$$

where  $\theta_W$  and  $\theta_B$  represent the relative contribution of covariance of allelic effects among loci at the within-

and between-population levels in comparison with the genic variance. Note that throughout the text, we will use  $F_{\text{STQ}}$  to refer to the mean genetic differentiation at QTL, whereas  $F_{\text{ST}}$  will refer to the neutral genetic differentiation.

From (eqn 1), it can be concluded that (Kremer & Le Corre 2011):

When  $\theta_B = \theta_W$ , then  $Q_{\text{ST}} = F_{\text{STQ}}$

When  $\theta_B > \theta_W$ , then  $Q_{\text{ST}}$  will always be larger than  $F_{\text{STQ}}$

When  $\theta_B < \theta_W$ , then  $Q_{\text{ST}}$  will always be lower than  $F_{\text{STQ}}$

Equation (1) was obtained within the frame of diallelic loci having equal additive contributions (the diallelic symmetrical model presented in Box 1). There is no simple writing in more complex cases with unequal contributions of loci and multiallelism. From simulations (Le Corre & Kremer 2003; Fig. 3), it was shown that the decoupling between  $Q_{\text{ST}}$  and  $F_{\text{STQ}}$  was larger under more realistic situations, suggesting that relationship (1) is conservative and corresponds to the minimum observable discrepancy between  $Q_{\text{ST}}$  and  $F_{\text{STQ}}$ . Relationship (1) holds even under non-equilibrium conditions between gene flow and selection (Kremer & Le Corre 2011).

### Consequences of local selection on covariance among selected loci

Box 2 illustrates how  $\theta_W$  and  $\theta_B$  are built up in simplified examples of two populations undergoing local stabilizing selection.  $\theta_W$  is negative in a population undergoing directional or stabilizing selection. This build-up of negative covariance under selection is known as the Bulmer effect (Bulmer 1980, 1989). On the opposite,  $\theta_B$  will be positive when selection is divergent among populations. The example given in Box 2 corresponds to a simplified case where the variance between populations ( $V_B$ ) is initially zero. In a more general context, the sign of  $\theta_B$  depends on the difference between the value of the genetic variance among populations before the onset of selection and the variance of phenotypic optima  $V_{\text{OPT}}$  (Latta 1998).

Under divergent local selection, the genetic variance among populations increases towards  $V_{\text{OPT}}$ . This response is permitted by two processes: first, the build-up of covariance of additive effects among loci; second, a change in allele frequencies that results in a change in the genic variance. The build-up of positive covariance among loci ( $\theta_B$ ) contributes to a larger decoupling between differentiation at QTL ( $F_{\text{STQ}}$ ) and phenotypic differentiation  $Q_{\text{ST}}$ , as predicted by (eqn 1). Conversely, a change in allele frequencies will affect both  $F_{\text{STQ}}$  and

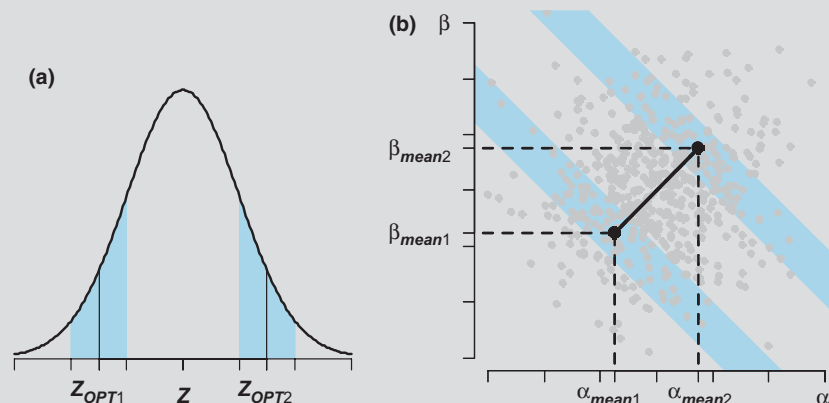
### Box 2. Expected values of $\theta_W$ and $\theta_B$ under local stabilizing selection

We consider an additive trait controlled by two equivalent loci and for simplicity with no environmental contribution:  $Z = \alpha + \beta$ , where  $Z$  is the trait's value, and  $\alpha$  and  $\beta$  are the genotypic values at each locus. We further consider an infinite number of alleles at each locus and large population sizes, so that genotypic values at each locus are normally distributed. In this particular example, and because of the equal contribution of the two loci,  $\theta_W$  and  $\theta_B$  take the form of correlation coefficients:

$$\theta_W = \frac{\text{COV}(\alpha, \beta)}{\sigma_\alpha \sigma_\beta} \text{ and } \theta_B = \frac{\text{COV}(\alpha_{\text{mean}}, \beta_{\text{mean}})}{\sigma_{\alpha_{\text{mean}}} \sigma_{\beta_{\text{mean}}}}$$

where  $\alpha$  and  $\beta$  are genotypic values within populations, and  $\alpha_{\text{mean}}$  and  $\beta_{\text{mean}}$  are mean within-population genotypic values.

Under stabilizing selection, selection will drive phenotypic values  $Z$  of individuals within each population to the local optimum  $Z_{\text{OPT}}$ . The result of the first generation of selection for two initially undifferentiated populations with different optimal values  $Z_{\text{OPT1}}$  and  $Z_{\text{OPT2}}$  is depicted on graphs a and b. Within each population, stabilizing selection will screen individuals having allelic values corresponding to a trade-off between  $\alpha$  and  $\beta$  so that their sum will approach the optimal phenotypic value  $Z_{\text{OPT}i}$  (graph a). This leads to negative values of  $\theta_W$  in each population (graph b). Overall, the mean values of  $\alpha$  and  $\beta$  of the selected individuals show a covariation opposite in sign to that observed at the within-population level (graph b). Divergent selection will create positive values of  $\theta_B$  regardless of the strength of stabilizing selection. In the case of uniform selection, that is, when  $Z_{\text{OPT1}} = Z_{\text{OPT2}}$ , the first generation of selection will still create negative values of  $\theta_W$  within each population, but no covariance at the between-population level ( $\theta_B = 0$ ).



$Q_{\text{ST}}$  in parallel, thus reducing the discrepancy between differentiation at QTL and phenotypic differentiation.

### Local adaptation under two simple quantitative genetics models

Simplified quantitative genetics models allow examining in more detail the build-up of genetic covariance and allele frequency changes under local stabilizing selection and migration. We present results from the single-locus model and the infinitesimal model, two models that represent the two most extreme genetic determinisms conceivable for a trait. Results were derived

using the same set of parameters for both models, as described in Table 1. These parameters were similar to those used to perform individual-based simulations in two previous papers (Le Corre & Kremer 2003; Kremer & Le Corre 2011). Briefly, we considered 25 populations connected via an islands' model of migration with a fixed initial total genetic variance  $V_T(0) = 5$  and an environmental variance set to 1. There was local Gaussian stabilizing selection, with local optima varying according to a linear gradient on a  $5 \times 5$  grid. Migration-drift equilibrium was assumed to take place at generation 0. Models were implemented until a migration-drift-selection equilibrium was attained.



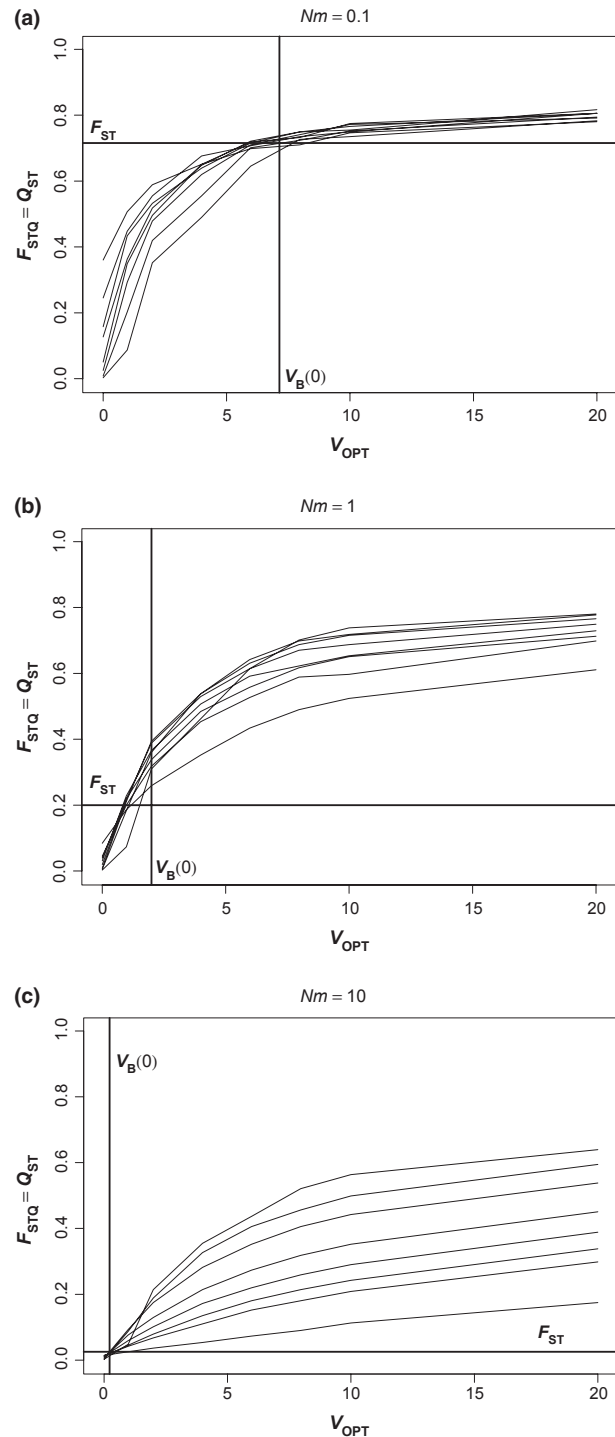
**Table 1** Parameters used in the models

Parameter	Value	Description
<b>Population model</b>		
$N_P$	25	Number of subpopulations
$N$	100	Number of individuals per subpopulation
$m$	0.001; 0.01; 0.1	Migration rate
<b>Selection model</b>		
$\omega^2$	1; 5; 10; 20; 30; 40; 50; 100	Selection intensity
$V_{ZOPT}$	0; 1; 2; 5; 10; 20	Variance of local phenotypic optima
<b>Genetic model</b>		
$V_T(0)$	5	Genetic variance for a randomly mating population with the same allele frequencies as the subdivided population considered
$V_E$	1	Environmental variance
$F_{ST}(0)$	$\frac{1}{2N[1-(1-1/2N)(1-m)^2]}$	Neutral genetic differentiation
$V_W(0)$	$[1-F_{ST}(0)] V_T(0)$	Initial genetic variance within population
$V_B(0)$	$2F_{ST}(0) V_T(0)$	Initial genetic variance among populations

### The single-locus, two-allele model

Under this model, response to selection is entirely via allele frequency changes (Box 1). Therefore, and as predicted by eqn (1), the genetic differentiation at the QTL,  $F_{STQ}$ , is equal to  $Q_{ST}$ . The single-locus model can be viewed as a boundary case corresponding to the largest possible difference between  $F_{STQ}$  and the neutral differentiation  $F_{ST}$ , that is, the most favourable case for a genome scan-based detection of QTL. We used this model to examine how the  $F_{STQ}-F_{ST}$  difference varies as a function of the interacting effects of migration and selection. Results were derived using recursion equations as in Santure & Wang (2009). The selected locus was considered unlinked to neutral markers, and we assumed that the neutral differentiation  $F_{ST}$  was constant. The additive effects of alleles at the selected locus were symmetrical and their values set to match the initial total genetic variance  $V_T(0) = 5$  (Table 1). Change in allele frequencies at the selected locus in response to selection were iterated based on relative fitness values for the different genotypes.

As shown in Fig. 1,  $F_{STQ}-F_{ST}$  is positive only when the variance of phenotypic optima  $V_{OPT}$  is higher than the initial (neutral) genetic variance among populations



**Fig. 1**  $F_{STQ}$  values as a function of the variance of local phenotypic optima ( $V_{OPT}$ ) in the single-locus, two-allele model. Results are presented for three levels of gene flow among local populations, from low ( $Nm = 0.1$ , a) to high ( $Nm = 10$ , c). The different curves on each graph correspond to different levels of selection intensity, from weak ( $\omega^2 = 100$ , lower curve) to strong ( $\omega^2 = 1$ , upper curve). The vertical line shows  $V_{OPT} = V_B(0)$ .

$V_B(0)$ . Under limited gene flow (Fig. 1a), moderately divergent local selection [ $0 < V_{OPT} < V_B(0)$ ] can result in a genetic differentiation at QTL lower than the neutral expectation. When selection is truly divergent [i.e. when  $V_{OPT} > V_B(0)$ ] but gene flow limited,  $F_{STQ}$  is only little higher than  $F_{ST}$ , because adaptive genetic differentiation can be attained with only a limited change in allele frequencies. Thus, under limited gene flow, there should be little power to detect QTL involved in response to local divergent selection because the background, neutral level of differentiation is high and can exceed the genetic differentiation driven by selection. Under moderate gene flow (Fig. 1b)  $F_{STQ}-F_{ST}$  depends mainly on the variance of optima, while under high gene flow (Fig. 1c), it also varies as a function of selection intensity. If selection intensity is weak (lower curve in Fig. 1c),  $F_{STQ}$  is only slightly higher than  $F_{ST}$ , because gene flow counteracts allelic differentiation in response to divergent selection. Under the single-locus model, the largest discrepancy between  $F_{STQ}$  and  $F_{ST}$  is expected under strong divergent local selection, when populations are connected by intermediate to high levels of gene flow.

#### *The Fisher–Bulmer infinitesimal model*

The Fisher–Bulmer infinitesimal model considers an infinite number of equivalent loci (Box 1). Response to divergent selection is then entirely attributable to the build-up of genetic covariance at the within-population ( $\theta_W$ ) and between-population ( $\theta_B$ ) levels, and allelic differentiation at QTL remains undistinguishable from that at neutral markers:  $F_{STQ} = F_{ST}$ . The infinitesimal model can be viewed as a boundary case corresponding to the maximum discrepancy between  $F_{STQ}$  and  $Q_{ST}$ . We used this model to examine how the  $Q_{ST}-F_{STQ}$  difference (caused by genetic covariance) varies as a function of the interacting effects of migration and selection.

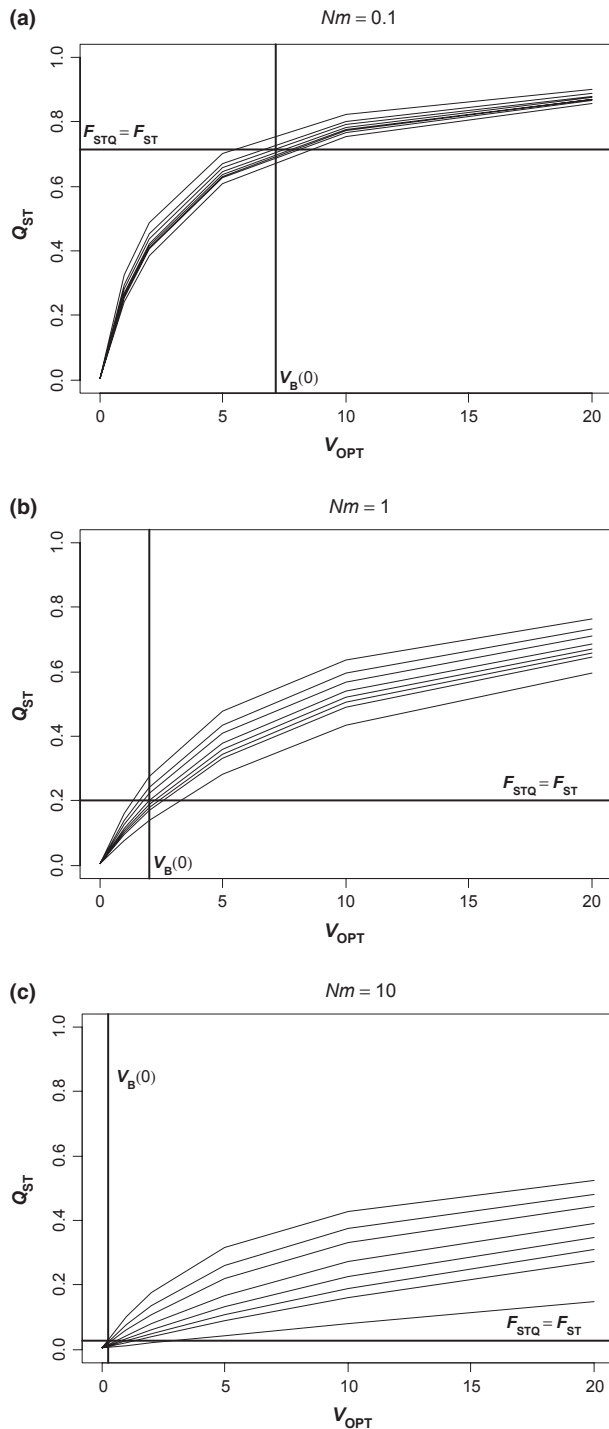
The infinitesimal model can be used to predict genetic variances over successive generations of selection based on normality assumptions. Here, we assumed that individual breeding values do not deviate from normality after migration (but see Yeaman & Guillaume 2009 and Tufto 2000 for a discussion of this), and we computed equilibrium values of genetic variances and covariances at the within- and between-population levels using recursion equations. A similar model was used by Lopez *et al.* (2008) to investigate the impact of seed and pollen migration on local adaptation in a set of plant populations. We used their recursion equations modified so that only diploid offspring migrate, and migration takes place after reproduction. Genetic differentiation at QTL ( $F_{STQ}$ ) was equal to the neutral differentiation (Table 1) and assumed to be constant across generations of selection.

Under all migration–selection schemes considered (Table 1), the amount of covariance among populations ( $\theta_B$ ) largely exceeded covariance at the within-population level ( $\theta_W$ ).  $\theta_B$  was positive when the variance of phenotypic optima  $V_{OPT}$  was higher than the initial (neutral) genetic variance among populations  $V_B(0)$ . As expected from eqn (1), and because  $\theta_W$  was much smaller than  $\theta_B$ ,  $\theta_B$  was the main driver of the difference between  $Q_{ST}$  and  $F_{STQ}$ . Thus, the sign of  $Q_{ST}-F_{STQ}$  is well predicted by the sign of  $V_{OPT}-V_B(0)$  (Fig. 2).

Under limited gene flow (Fig. 2a), moderately divergent local selection [ $0 < V_{OPT} < V_B(0)$ ] can result in a phenotypic differentiation lower than the allelic genetic differentiation. When gene flow is limited, local drift can drive allelic genetic differentiation to values higher than the genetic differentiation that would match the divergence of local optima. For a trait controlled by a very large number of loci, local adaptation is then attained through the antagonistic effect of negative genetic covariance among populations. When selection is truly divergent [i.e. when  $V_{OPT} > V_B(0)$ ],  $Q_{ST}$  is only little higher than  $F_{STQ}$ , as only a little amount of positive genetic covariance allows for local adaptation. Under moderate gene flow (Fig. 2b), the  $Q_{ST}-F_{STQ}$  difference increased with the variance of optima. Under high gene flow (Fig. 2c),  $Q_{ST}-F_{STQ}$  varies also as a function of selection intensity. If selection intensity is weak (lower curve in Fig. 2c), gene flow opposes divergent selection, driving phenotypic differentiation close to molecular differentiation. Under the infinitesimal model, the largest discrepancy between phenotypic differentiation and molecular differentiation is expected under strong divergent local selection, when populations are connected by intermediate to high levels of gene flow.

#### *Redefining expected signatures of selection*

Local selection has usually been classified as being uniform when all subpopulations are driven towards the same optimum phenotype and divergent otherwise. However, as exemplified by the two models considered above, this simple definition may be misleading. If population phenotypic means are more divergent than the optimum phenotypes because of limited migration and drift, divergent selection will result in a level of adaptive differentiation ( $Q_{ST}$ ) being lower than the neutral differentiation, which is usually interpreted as a signature of uniform selection. If the trait is encoded by a single locus, the differentiation at this selected locus will also drop below the neutral differentiation in order to meet optimal phenotypic values. The outcome will be different, however, for a multilocus trait: when populations are selected towards phenotypic optima less divergent than their neutral phenotypic means, selection



**Fig. 2**  $Q_{ST}$  values as a function of the variance of local phenotypic optima ( $V_{OPT}$ ) in the infinitesimal model. Results are presented for three levels of gene flow among local populations, from low ( $Nm = 0.1$ , a) to high ( $Nm = 10$ , c). The different curves on each graph correspond to different levels of selection intensity, from weak ( $\omega^2 = 100$ , lower curve) to strong ( $\omega^2 = 1$ , upper curve). The vertical line shows  $V_{OPT} = V_B(0)$ .

**Table 2** Comparative levels of genetic differentiation expected at neutral markers ( $F_{ST}$ ), quantitative trait loci ( $F_{STQ}$ ) and quantitative trait ( $Q_{ST}$ ) according to type of local selection and level of gene flow

Selection	Gene flow	
	Low	$\leftrightarrow$ High
Uniform to weakly divergent (variance of optima < neutral variance)	$F_{STQ} > F_{ST} \gg Q_{ST} \leftrightarrow F_{ST} \approx F_{STQ} \approx Q_{ST}$	
Divergent (variance of optima > neutral variance)	$F_{ST} < F_{STQ} < Q_{STQ} \leftrightarrow F_{ST} \ll F_{STQ} \ll Q_{ST}$	

induces a negative covariance of allelic frequencies that counterbalances the allelic differentiation at QTL (Latta 1998; Le Corre & Kremer 2003). Moreover, multilocus traits allow for genetic redundancy, meaning that similar phenotypic values can be produced by different combinations of alleles (Goldstein & Holsinger 1992). This accounts for the congruent observation by Latta (1998) and Le Corre & Kremer (2003) that uniform selection on a multilocus trait results in elevated differentiation at QTL in comparison with neutral markers.

The main expectations regarding the relative values of  $F_{ST}$ ,  $F_{STQ}$  and  $Q_{ST}$  for a multilocus trait can be roughly summarized as shown in Table 2. The limit put between 'weakly divergent' and 'divergent' selection, based on the comparison of the variance of phenotypic optima with the neutral genetic variance among populations, will be difficult to establish in practice as the neutral variance among populations depends on both the total available variance for the trait and the joint effects of migration and drift. It must be emphasized that it is a relative notion: under identical environmental constraints, genetically isolated populations are more likely to experience 'weakly divergent' selection.

In cases where contrasted values of  $F_{ST}$ ,  $F_{STQ}$  and  $Q_{ST}$  are expected, for example, when gene flow is high and selection divergent (Table 2), the success of outlier-based detection methods will depend on the number of loci involved in the trait: for a small number of loci,  $F_{STQ}$  should be closer to  $Q_{ST}$  than to  $F_{ST}$ , whereas for a high number of loci,  $F_{STQ}$  should be closer to  $F_{ST}$ . The effect of the number of loci as well as the effects of other attributes of the genetic architecture of a quantitative trait not taken into account by simplified quantitative genetic model will now be described.



## Impact of the genetic architecture

### Number of loci

Mathematically, for  $n$  quantitative trait loci, the genetic variance attributable to the variance of allelic frequencies can be as low as  $1/n$ , the rest of the variance being accounted for by covariance. As shown by Latta (1998), there is a much greater influence of covariance on genetic differentiation when the trait is controlled by numerous loci each of small effects than when it is controlled by a few loci with large effects. Figure 3 illustrates the effect of the number of loci on  $F_{STQ}$  for a set of 25 populations connected by intermediate level of gene flow ( $Nm = 1$ ) under divergent local selection. These results were obtained using individual-based simulations as in our previous papers (Le Corre & Kremer 2003; Kremer & Le Corre 2011). Effects of alleles were drawn in a Gaussian distribution with variance adjusted to match the fixed value of the initial genetic variance (Table 1) and were smaller as the number of loci increased. Simulation results show that mean  $F_{STQ}$  values approach phenotypic differentiation only when the number of loci is  $< 5$ . When the selected trait is controlled by 20 or more loci, genetic covariance is the main driver of genetic differentiation, and  $F_{STQ}$  became closer to the neutral differentiation  $F_{ST}$  than to  $Q_{ST}$ .

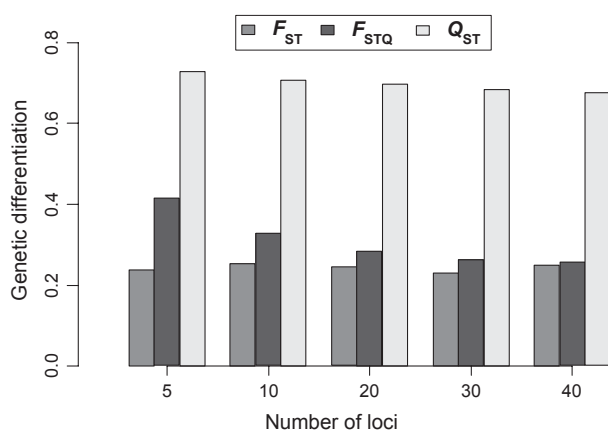
Throughout this paper, we considered QTL of identical effects with no physical linkage among them and evaluated their mean  $F_{STQ}$  value. However, as demonstrated by Griswold (2006) and Yeaman & Whitlock (2011), because gene flow can swamp divergence at weakly selected alleles, the interaction between migration and selection favours genetic architectures with some large alleles contributing most to adaptive

divergence. Under prolonged bouts of adaptation, a concentrated genetic architecture with a few large-effect loci emerges (Yeaman & Whitlock 2011). Under shorter time periods, the genetic architecture of adaptive variation would be highly dependent on the availability of mutations of large effects. Empirical data show disparate patterns among species and among traits, with some traits determined by large number of small-effect loci [e.g. many traits in mice, drosophila and human (Flint & Mackay 2009), flowering time in maize (Buckler *et al.* 2009)], while others are controlled by a few large-effect QTL [e.g. armour plating in sticklebacks (Albert *et al.* 2008), flowering time in *Arabidopsis thaliana* (Salomé *et al.* 2011)]. When few large-effect genes coexist with numerous small-effect genes, the former are much more likely to be detected by outlier approaches.

### Inbreeding, dominance and epistasis

The mating system of many plants and several animal species allows for selfing. Selfing has two distinct effects on genetic differentiation: first, selfing reduces gene flow among local populations by reducing male gametes dispersal; second, selfing reduces local effective sizes, thus enhancing local genetic drift (Duminil *et al.* 2009). Both effects result in increased genetic differentiation. In their simulation study, Le Corre & Kremer (2003) compared two contrasted selfing rates (0% vs. 90%) while maintaining identical migration rates. As expected, local inbreeding resulted in higher values of all three parameters of genetic differentiation:  $F_{ST}$ ,  $F_{STQ}$  and  $Q_{ST}$ . Inbreeding is accompanied by reduced recombination, so that higher amounts of genetic covariance can be maintained at the within-population level. Simulation results showed that under local divergent selection, inbreeding also resulted in smaller positive genetic covariance at the between-population level, and that there was no noticeable net effect on the difference between  $F_{STQ}$  and  $Q_{ST}$ .

The expectation that  $F_{ST} = F_{STQ} = Q_{ST}$  under neutrality only applies to QTL showing no dominance of allelic effects at any given locus and no epistasis, that is, no interaction between allelic effects at different loci. At neutral equilibrium, dominance can result in either  $Q_{ST} > F_{ST}$  or, more likely,  $Q_{ST} < F_{ST}$ , depending on the population structure and allele frequencies (Lopez-Fanjul *et al.* 2003, 2007; Goudet & Buchi 2006; Goudet & Martin 2007). The effect of dominance under divergent selection has been studied for a trait encoded by a single locus (Santure & Wang 2009). It was shown that dominance increased the value of  $Q_{ST}$  relative to  $F_{STQ}$  for an additive trait. The  $Q_{ST} - F_{STQ}$  contrast was found to be maximal when the dominance effect was large and the selfing rate was small. Thus, it seems that



**Fig. 3** Effect of the number of loci on  $F_{STQ}$  relative to  $F_{ST}$  and  $Q_{ST}$ . Results are based on computer simulation using parameter values as listed in Table 1, with  $Nm = 1$ ,  $\omega^2 = 20$  and  $V_{OPT} = 5$ .

dominance accentuates the decoupling between phenotypic differentiation and genetic differentiation at the selected loci. However, the parameters of divergent selection had a much greater impact on  $Q_{ST}$ – $F_{STQ}$  than dominance and inbreeding (Santure & Wang 2009). Epistasis is expected to decrease the neutral value of  $Q_{ST}$  in comparison with the neutral value of  $F_{ST}$  (Whitlock 1999; Lopez-Fanjul *et al.* 2003, 2007). Under stabilizing selection, epistasis increases the level of genetic variance maintained in a single population (Gimelfarb 1989). The consequences of epistasis in a subdivided population with divergent local selection have not been investigated. As conjectured by Le Corre & Kremer (2003), large epistatic effects should reinforce the contribution of genetic covariance to the genetic divergence among populations, hence accentuating the decoupling between  $Q_{ST}$  and  $F_{STQ}$ .

### Effects of gene flow on adaptive divergence

The difference between  $Q_{ST}$  and the molecular differentiation ( $F_{ST}$  and  $F_{STQ}$ ) increases as gene flow increases from low to moderate (Fig. 2a, b). This is coherent with earlier individual-based simulations, showing that the amount of genetic covariance among populations  $\theta_B$  increased as gene flow increased (Le Corre & Kremer 2003). As  $\theta_B$  is the main driver of  $Q_{ST}$  for multilocus traits, this solves the apparent contradiction of coexistence of large adaptive differentiation and large gene flow, as has been widely reported in forest trees (Savolainen *et al.* 2007) and in case studies of fishes (Saint-Laurent *et al.* 2003; Hemmer-Hansen *et al.* 2007) and in birds (Postma & van Noordwijk 2005). In this respect, gene flow has the same effect on  $\theta_B$  than the increase in the number of loci contributing to the trait. It is a combinatorial effect owing to the number of elements (loci or alleles) involved in the trait. Our biological understanding of the positive impact of gene flow on the  $\theta_B$  is that importation of new alleles in populations will generate more opportunities to have these alleles involved in beneficial allelic associations contributing to adaptive divergence.

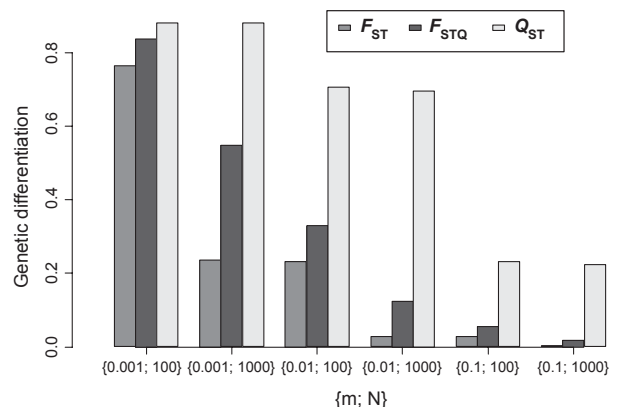
However, a further increase in gene flow may also prevent adaptive differentiation, especially when selection intensity is weak. Indeed, the equilibrium value of  $Q_{ST}$  is lower under high gene flow (Fig. 2b, c). A rapid decline in adaptive divergence with increasing gene flow was also predicted by Hendry *et al.* (2001) based on an infinitesimal model. Conclusions from the infinitesimal model should be taken with care, as this simplified model overestimates the constraints set by migration (Yeaman & Guillaume 2009). This is because the infinitesimal model assumes that genetic values are normally distributed within a population and ignores the skew towards

immigrants' mean phenotypic values generated by gene flow. Because of this skew, immigrant alleles are more efficiently eliminated by stabilizing selection, particularly when selection is strongly divergent and adaptive differentiation is attributed to a few large-effect alleles (Yeaman & Guillaume 2009). In any case, these results underline the ambivalent role of gene flow in local adaptation. On the one hand, it enhances adaptive population divergence through allelic covariance; on the other hand, it may induce a migration load, preventing populations from reaching their optimal phenotypic value (Hendry *et al.* 2001; Björklund *et al.* 2009).

While adaptive divergence is determined by the rate of migration ( $m$ ) relative to the strength of selection, neutral divergence is a function of both the migration rate and the effective population size. Figure 4 illustrates the response of  $F_{ST}$ ,  $F_{STQ}$  and  $Q_{ST}$  to different combinations of migration rate and population size. When  $m$  is held constant,  $Q_{ST}$  is unaffected by population size, whereas  $F_{ST}$  can be quite a bit lower with larger population sizes. The effect of population size is slighter on  $F_{STQ}$  than on  $F_{ST}$ . The difference between  $F_{ST}$  and  $F_{STQ}$ , as well as the difference between  $Q_{ST}$  and  $F_{STQ}$ , is much larger for a given migration rate when population size is high. This suggests that outlier methods would perform better for large-sized populations, although the level of differentiation at QTL would still understate adaptive divergence.

### Non-equilibrium situations

The temporal dynamics of  $F_{STQ}$  has to be considered to assess the power of genome scan methods to detect loci involved in the response to recent or ongoing selection. Unless there are genes with large effects, gene frequen-



**Fig. 4** Effect of population size ( $N$ ) and migration rate ( $m$ ) on  $F_{ST}$ ,  $F_{STQ}$  and  $Q_{ST}$ . Results are based on computer simulation using parameter values as listed in Table 1, with 10 quantitative trait loci,  $\omega^2 = 20$  and  $V_{OPT} = 5$ .

cies change slowly under selection. In contrast, the build-up of genetic covariance, or Bulmer effect, operates immediately (Falconer & Mackay 1996). Simulations (Kremer & Le Corre 2011) show that phenotypic differentiation as measured by  $Q_{ST}$  settles very quickly in response to local selection, in a few tens of generations during which genetic variance within populations decreases and genetic variance among populations increases. On the contrary, the dynamics of  $F_{STQ}$  values is very slow: hundreds of generations are needed before allele frequencies at QTL reach a steady state. This is because response to selection seems to operate as a two-step process. In a first step, genetic variance within populations and genetic variance among populations change mainly through the build-up of genetic covariance. In a second step, the gradual decrease in genetic covariance via recombination is compensated for by allele frequency changes at the QTL. This two-step dynamics implies that populations may become phenotypically differentiated long before detectable changes in allele frequencies occur at QTL, so that recent or ongoing divergent selection should be more difficult to detect using genome scans.

### Comparative estimates of $F_{ST}$ , $F_{STQ}$ and $Q_{ST}$ in experimental surveys

We conducted a review of case studies that estimated levels of differentiation at genes underlying adaptive traits. Although genes involved in adaptive traits are largely unknown, cataloguing of candidate genes has been completed for a large number of species, thanks to rapid technological development in genomics (Neale & Kremer 2011). Qualification of 'candidate' genes is based on three criteria (Pflieger *et al.* 2001): (i) their position on the genetic map matching with a QTL of a trait of interest ('positional candidates'), (ii) their differential expression between environments generating different responses of the trait ('expressional candidates'), (iii) their documented contribution to the expression of the trait in model species ('functional candidates'). Our review of published data on  $F_{STQ}$  refers to candidate genes and not to the true genes underlying the trait. This approximation may of course bias the estimates of  $F_{STQ}$ , as the set of candidate genes might be an incomplete set of the true genes. This approximation is well recognized in most cited papers in Table 3. This is because candidate genes were usually loosely identified regardless of the criteria used. Positional candidates matching a given QTL can comprise numerous false positives, given the large genomic portion covered by a QTL and the high linkage disequilibrium in mapping pedigrees. Qualification of functional candidate genes based on annotations in model organisms can be vague and may concern, for

example, gene families that have diverged between species, and is rarely supported by checking orthology in the recipient species. As a result, differentiation of these 'loosely defined' candidate genes is likely to underestimate differentiation of the true genes. We therefore subsequently screened the 'loosely defined' candidate genes by assessing their responses to natural selection. Indeed, most cited studies evaluated the imprint of natural selection on the allelic profiles of genes by using at least one of three methods: (i) departure of  $F_{STQ}$  values from neutral expectation (outlier method), (ii) clinal variation of allelic frequencies along environmental gradient (cline method), (iii) statistical association with phenotypic traits that have adaptive significance (association method). While each of this method may have its own shortcomings, as the signature of natural selection may be blurred by confounding effects (demography, population structure, etc.), they reduced the number of candidate genes by more than 50% (Table 3). It is likely that these methods miss true genes. Indeed, as suggested by earlier theoretical predictions (Le Corre & Kremer 2003; Kremer & Le Corre 2011), differentiation of most true genes will not deviate from neutral expectations and hence will not be detected by the outlier detection method. As a result, differentiation of the 'selected' candidate genes is likely to overestimate differentiation of the true genes. Based on this reasoning, differentiation of true genes should be located in between differentiation of the 'loosely defined' candidate genes and the 'selected' candidate genes, and Table 3 should be interpreted along this reasoning.

We focused on the comparison of observed  $F_{STQ}$  values—for 'loosely defined' and 'selected' candidate genes—with observed values of differentiation at neutral markers ( $F_{ST}$ ), and, when available,  $Q_{ST}$  values for the trait considered. Results for the four differentiation measures were sometimes presented in separate publications. We attempted however to assemble estimates from the same sets of populations when separate analyses were conducted. We were able to compile data for  $F_{ST}/F_{STQ}/Q_{ST}$  comparisons from 13 case studies, and from 15 additional studies for the  $F_{ST}/F_{STQ}$  comparison (Table 3). Most of the reported studies comparing  $F_{ST}$ ,  $F_{STQ}$  and  $Q_{ST}$  concern forest trees, mainly conifers and a few surveys in oak and poplar. These species share similar demographic and ecological features: large population sizes and extensive gene flow mainly because of pollen flow (Savolainen *et al.* 2007). The target traits that were investigated are in most cases related to tolerance to biotic or abiotic stresses, growth and phenology (timing of bud burst or growth cessation), for example, traits that are strongly correlated to fitness in trees. As reported in earlier reviews, they exhibit very large differentiation in common garden conditions, which

**Table 3** Summary of comparative studies of genetic differentiation at neutral markers, selected loci and quantitative traits

T	Species name	N <sub>P</sub>	Neutral markers		Loosely defined candidate genes			Selected candidate genes responding to natural selection			Traits		References
			Marker	N <sub>L</sub>	F <sub>ST</sub>	Selection	Ng	F <sub>STQ</sub>	Detection	Ng	F <sub>STQ</sub>	Type of traits	
P	<i>Arabidopsis thaliana</i>	12	SSR	12	0.595	F,P	1	0.607	NA	NA	NA	Flowering time	1
P	<i>Picea abies</i>	7	Control genes	11	0.157	F	11	0.143	NA	NA	NA	Seasonal growth cessation	2
P	<i>Picea glauca</i>	6	Allozymes	14+11	0.014	E	345	0.006	O	19	0.04 to 0.13	Wood traits, growth, phenology	3
P	<i>Picea sitchensis</i>	17	Control genes + ESTP (a)	98 SNPs	0.007	E	153	0.06	A	28	0.10	Growth, phenology, frost damage	4
P	<i>Picea mariana</i>	5	NA	NA	NA	F	313	0.005 to 0.006	O	25	0.041 to 0.078	Bud set, Height	24
P	<i>Pinus pinaster</i>	10	SSR	8	0.15	E, F	13	0.14	O	2	0.278	Drought resistance	5
P	<i>Pinus taeda</i>	8	SSR	22	0.009	E, F	19	0.023	O	4	0.109	Wood quality	14
P	<i>Pinus sylvestris</i>	2	Control genes	16	0.0095	F	14	0.003	O	1	0.13	Cold tolerance	15
P	<i>Populus tremula</i>	12	SSR	18	0.012	F	5	0.016	C	NA	0.015	Phenology	6
P	<i>P. tremula</i>	12	Control genes	21	0.016	F	23	0.018	C	4	0.01 to 0.03	Growth cessation	7
P	<i>Pseudotsuga Menziesii</i>	2	Allozymes	25	0.035	E,F,P	117	0.022	A	12	0.048	Cold damage	8
P	<i>Quercus petraea</i>	9	SSR	15	0.013	E,F,P	9	0.042	NA	NA	NA	Bud burst	9
P	<i>Quercus petraea</i>	12	SSR	16	0.023	E,F,P	73	0.023	O	15	0.058	Bud burst, Seed germination	10
P	<i>Triticum aestivum</i>	3	SSR	21	0.14	F, P	6	0.183	A	NA	NA	Flowering time	11
P	<i>Zea mays</i> ssp. <i>parviglumis</i>	6	Control genes	7	0.224	F	16	0.046 to 0.316	O	4	0.183 to 0.316	Plant defence against pathogens	17
V	<i>Coregonus clupeaformis</i>	8	SSR	6	0.24	P	27	0 to 0.865	O	8	0.199 to 0.865	Body weight	12
V	<i>Coregonus clupeaformis</i>	2	SSR	6	0.26	E	96	0.28	O	12	0.775	Body weight	25
V	<i>Coregonus clupeaformis</i>	2	SSR	6	0.14	E	96	0.11	O	7	0.488	Body weight	25
V	<i>Coregonus clupeaformis</i>	2	SSR	6	0.08	E	96	0.06	O	5	0.338	Body weight	25
V	<i>Coregonus clupeaformis</i>	2	SSR	6	0.06	E	96	0.02	O	3	0.193	Body weight	25
V	<i>Gasterosteus aculeatus</i>	8	SSR	9	0.056 to 0.157 (c)	P	7 (d)	0 to 0.46 (c)	NA	NA	NA	Morphological traits	13

Table 3 Continued

T	Species name	Neutral markers			Loosely defined candidate genes			Selected candidate genes responding to natural selection			Traits		References
		$N_p$	Marker	$N_L$	$F_{ST}$	Selection	$N_g$	$F_{STQ}$	Detection	$N_g$	$F_{STQ}$	Type of traits	
P	<i>Solanum peruvianum</i>	6	Control genes	8	0.008 to 0.218 (c)	NA	NA	NA	F	2	0.000 to 0.442 (c)	Drought tolerance	NA 16
I	<i>Solanum chilense</i>	18	Allozymes	17	0.029	NA	NA	NA	F	1	0.456	Thermal tolerance	NA 18
I	<i>Lycena tityrus</i>	3	Control genes	5	0.063	NA	NA	NA	F	1	0.285	Host specificity	NA 19
V	<i>Anas georgica</i>	116	Control genes	7	0.010	NA	NA	NA	F	1	0.37	Oxygen affinity of haemoglobin	NA 20
V	<i>Oncorhynchus tshawytscha</i>	2	SSR	9	0.005	NA	NA	NA	F	2	0.069	Spawning time	NA 21
V	<i>Platichthys flesus</i>	12	SSR	9	0.025	NA	NA	NA	E	1	0.14	Stress tolerance	NA 22
V	<i>Pomatoschistus minutus</i>	7	SSR	8	0.013	NA	NA	NA	F	1	0.455	Eye vision	NA 23

T, Taxa (P, plant; I, insect; V, vertebrate);  $N_p$ , number of populations;  $N_L$ , number of loci;  $N_g$ , number of genes; SSR, simple sequence repeats; SNPs, single nucleotide polymorphisms; NA, no data.

Method for identifying 'loosely defined' candidate genes: E, gene expression; F, gene function; P, gene located within a QTL.

Method for identifying 'selected' candidate genes responding to natural selection: O, outlier  $F_{ST}$  value; A, statistical association with an adaptive phenotypic trait; C, clinal variation along an environmental gradient.

a: Expression sequence tag polymorphism, b: Broad sense  $Q_{ST}$ , c: pairwise differentiation estimates, d: SSR located in QTL regions of traits investigated, e: a phenotypic analogue ( $P_{ST}$ ) to  $Q_{ST}$  was used as differentiation measure.

References: 1. Le Corre 2005; 2. Heuertz *et al.* 2006; 3. Jaramillo-Correa *et al.* 2001; Namroud *et al.* 2008; 4. Holliday *et al.* 2010; 5. Eveno *et al.* 2008; Sanchez-Gomez *et al.* 2010; 6. Hall *et al.* 2007; 7. Hall *et al.* 2007; Ma *et al.* 2010; 8. St Clair 2006; Eckert *et al.* 2009; 9. Ducousso *et al.* 2005; Derory *et al.* 2010; 10. Alberto *et al.* 2010, 2011 and unpublished data; 11. Rhoné *et al.* 2010; 12. Rogers & Bernatchez 2005; 13. Raeymaekers *et al.* 2007; 14. Gonzalez-Martinez *et al.* 2006; 15. Wachowiak *et al.* 2009; Pyhäjärvi *et al.* 2007; Aho 1994; 16. Xia *et al.* 2010; 17. Moeller & Tiffin 2008; 18. Karl *et al.* 2009a,b; 19. Heide-Fischer *et al.* 2010; 20. McCracken *et al.* 2009; 21. O'Malley *et al.* 2007; 22. Hemmer-Hansen *et al.* 2007; 23. Larmuseau *et al.* 2009; 24. Prunier *et al.* 2011; 25. Renaut *et al.* 2011.



contrast to the very low differentiation of neutral markers. In all reported cases, the average value of  $F_{STQ}$  is lower than  $Q_{ST}$  by an order of magnitude. While  $F_{STQ}$  of 'loosely defined' candidate genes is close to  $F_{ST}$ , there is however a substantial increase in  $F_{STQ}$  of 'selected' genes. It should be emphasized that microsatellites were used as neutral markers in 16 of 28 studies, and their differentiation values may consequently be lower than for biallelic markers as single nucleotide polymorphisms (SNPs) (Hedrick 2005). That  $F_{STQ}$  of 'selected' genes is lower than  $Q_{ST}$  but larger than  $F_{ST}$  (Fig. 2c) suggests that fitness traits in forest trees are under strongly diversifying selection pressures, but that local adaptation results from genetic covariance among large number of genes rather than from allele frequency shifts at a few selected loci. Because forest trees have long generation time, selection in relation to climate is comparatively more 'recent' than for more short-lived species, which would also favour adaptation via covariance of allelic effects as opposed to allele frequency changes. In lake whitefish (Renaut *et al.* 2011), elevated  $F_{STQ}$  values were also found at 'selected genes', and more such genes were identified in populations with stronger ecological diversification, suggesting that strong selection pressures induced detectable allele frequency changes. Among the studies that did not estimate quantitative differentiation ( $Q_{ST}$ ), many focused on traits with well-known, simple genetic determinism and compared genetic differentiation at one or two major-effect functional candidate genes with genetic differentiation at control genes or microsatellites. In these studies,  $F_{STQ}$  was substantially higher than  $F_{ST}$  (Table 3). This confirmed the predicted effect of the number of loci on  $F_{STQ}$  (Fig. 3).

### Conclusions and future directions

Historically, positive selection has been analysed under the paradigm of the 'selective sweep', which describes adaptation as resulting from the fixation of a single new beneficial mutation. The expected signature of positive selection under a selective sweep consists in shrinkage of allelic diversity and higher linkage disequilibrium at and around the selected gene (Kim & Stephan 2002). Recently, this approach has been challenged based on the following lines of reasoning: first, response to selection may arise from standing variation rather than from novel mutation, leading to 'soft sweeps' with a fainter molecular signature (Hermisson & Pennings 2005); second, selection may be too recent for fixation to have eventually occurred, so that both beneficial and neutral or deleterious alleles still segregate, which also leads to a weakened signature (Hohenlohe *et al.* 2010); finally and very importantly, it has

been recognized that most adaptive traits have a quantitative genetic basis, for example, they are controlled by large number of genes that may all together be the targets of selection (Pritchard & Di Rienzo 2010). Adaptation from standing variation at many loci, resulting in multiple but not necessarily substantial allele frequency changes, can thus be adopted as an alternative paradigm for studying positive selection. Quantitative genetics undoubtedly provides an adequate framework within this new paradigm (Kimura & Crow 1978) and already led to significant developments. As demonstrated by Chevin & Hospital (2008), the dynamics of a beneficial allele at a quantitative trait locus depends not only on its own effect, but also on the mean and variance for the trait contributed by other loci, which produces an attenuated selection signature.

Positive selection is often related to environmental factors that vary geographically, inducing spatially varying selective pressures and local adaptation (Savolainen *et al.* 2007; Barreiro *et al.* 2008; Hancock *et al.* 2010a, 2011; Fraser *et al.* 2011). Considering the degree of genetic differentiation among populations therefore appears as a preferred approach as it opens ways to establish local adaptation and decipher its genetic basis: evidence for local adaptation can be achieved from a comparison of quantitative genetic differentiation at the trait ( $Q_{ST}$ ) with neutral genetic differentiation ( $F_{ST}$ ) (Spitze 1993; Whitlock & Guillaume 2009), and selected loci can be identified from a comparison of their differentiation ( $F_{STQ}$ ) with the neutral differentiation ( $F_{ST}$ ) using a genome scan or candidate gene approach (Luikart *et al.* 2003). In this review, we have adopted a quantitative genetics approach to describe how population structure, the parameters of local selection and the genetic architecture for the trait modify the values of  $Q_{ST}$  and  $F_{STQ}$  in comparison with  $F_{ST}$ , thus influencing the ability to infer local adaptation and identify its genetic determinants. We have identified two main cases under which  $F_{ST}$ -based methods may fail to identify the genetic causes of adaptation.

The first case is when populations are connected by low levels of gene flow, which drives neutral differentiation to high values that can be similar or even higher than the genetic differentiation resulting from local divergent selection. The same confounding effects of population structure and local selection will arise if strong isolation by distance mimics local selection patterns, for example, along environmental gradients.

The second case is when phenotypic divergence is primarily because of positive covariance of allelic effects at the between-population level rather than allelic differentiation, so that  $F_{STQ}$  remains similar to  $F_{ST}$ . The  $Q_{ST}$ – $F_{STQ}$  difference is expected to be the largest in the following circumstances: when the trait is controlled by a large

number of QTL, when populations are linked by medium to high levels of gene flow, when selection is highly divergent among populations and selection intensity is strong within each population, and when selection is recent or ongoing. The effects of inbreeding, dominance and epistasis are much less than those of the number of loci and the selection parameters. These particular circumstances are likely to occur. Experimental data suggest they should be particularly prevalent for climate-related traits in forest trees, where loci detected as outliers seem to harbour low differentiation levels close to those measured at neutral markers, although the large values observed for the  $Q_{ST}$ – $F_{ST}$  difference provides evidence for local adaptation. Under these conditions, each individual QTL involved in local adaptation has an  $F_{STQ}$  value only slightly higher than the neutral  $F_{ST}$ . Alternatively, if the trait is defined by QTL with non-uniform effects, only a few large QTL can be easily detected while many small covarying alleles having significant effects on adaptation may be missed.

Outlier methods may have low power because, first, the markers assessed may be in incomplete linkage with the true QTL, so that their  $F_{ST}$  values are even lower than the  $F_{ST}$  values of the true QTL, and, second, because each marker is tested individually. In that respect,  $F_{ST}$ -based outlier detection studies share similarity with genome-wide association studies and exhibit the same methodological weakness. In genome-wide association studies, SNP variants that have a low effect on the trait's variance remain undetected when each SNP is tested individually for an association with the trait, because a very stringent  $P$ -value has to be used in multiple tests. As a consequence, a large part of the trait's variance is 'missing', that is, not explained by the validated SNPs. A prominent example of this comes from genome-wide association studies for human height, where associated variants were found to account for <5% of phenotypic variance, despite heritability for the trait is very high (Yang *et al.* 2010). Using a simple model that accounts for additive effects of SNPs on the trait's variance, Yang *et al.* (2010) were able to show that a much greater amount of variance (45%) can be explained by considering all SNPs simultaneously, not individually. Similarly, a large part of the adaptive differentiation among populations may be missed by studies that test markers individually for an outlier  $F_{ST}$  value.

These considerations prompt for considering not only individual allele frequency changes, but also covariance among the tested markers. Covariance has two components: the covariance of allele frequencies and the covariance of allelic effects. Across multiple loci, alleles having similar effects (decreasing or increasing the trait's value) are driven to similar high or low frequencies within populations selected for contrasted pheno-

typic values. Covariance of allele frequencies among populations can be estimated using a decomposition of linkage disequilibrium into within- and between-population components (Ohta 1982; Storz & Kelly 2008). Covariance of allelic effects can be estimated using additive allelic effects estimated from a separate association analysis (Ma *et al.* 2010). In their pioneering study of the differentiation of genes from the photoperiodic pathway in the tree *Populus tremula*, Ma *et al.* (2010) found strong evidence for selection from clinal patterns at several genes, but detected none of these genes from a genome scan for  $F_{ST}$  outliers. They found that locus pairs showing significant allele frequency clines also showed high pairwise disequilibrium among populations. Further, they demonstrated that 20–25% of the observed phenotypic variation could not be attributed directly to individual SNPs, but was rather explained by covariance among individual effects.

Another limitation of outlier  $F_{ST}$  methods is related to the fact they do not explicitly take into account information about local selection, as revealed by the pattern of variation of environmental factors or by the pattern of phenotypic divergence among populations. Usually, these factors are taken into account only implicitly via a sampling scheme that maximizes either the variance of environmental conditions at sampling locations or the phenotypic variance among populations, when known from a separate study (e.g. progeny tests in forest trees). Taking external information into account more explicitly would certainly improve the power of  $F_{ST}$ -based approaches. In a recent study that aimed at deciphering human adaptation to diet, Hancock *et al.* (2010b) found that a blind  $F_{ST}$ -based approach was less likely to detect the subtle allele frequency changes involved than a Bayesian approach that explicitly considered the variation of environmental factors. Their approach consisted in comparing a model in which allele frequencies were dependent on a given environmental variable in addition to population structure, relative to a model in which allele frequencies were dependent on population structure only. Their approach is thus similar to the Bayesian  $F_{ST}$  outlier detection method proposed by Foll & Gaggiotti (2008), while taking into account environmental data. In a similar approach, Prunier *et al.* (2011) successfully identified genes involved in climate adaptation in black spruce by using populations that maximized the contrast in climatic factors and used regression analyses between the frequencies of the identified SNPs and climatic variables (Joost *et al.* 2007) to support their implication in adaptive processes. The demonstration that integrating various external information may lead to the successful identification of important selected genes has been provided by a recent study by Marliac *et al.* (2011). This study that aimed at identifying the genetic basis of

adaptation of pearl millet to climatic variation used the  $F_{ST}$  outlier detection method of Foll & Gaggiotti (2008), in combination with several additional approaches: the study focused on a large set of candidate genes, populations were sampled among a well-described environmental gradient, correlation of allele frequencies with environmental factors were quantified and, finally, an independent association study was conducted to verify the link between allelic variation at the genes identified and phenotypic variation. Another example of a successful integrated approach is provided by the study from Renaut *et al.* (2011), where the integration of phenotypic, transcriptomic and functional genomic information resulted in the identification of genes involved in ecological speciation in lake whitefish.

The success of the several recent studies that pave the way to more integrated approaches for detecting selected loci suggest that much progress can still be expected in the deciphering of the genetic basis of adaptive quantitative traits. The preferred roads will consist in shifting from testing individual markers to multilocus approaches and integrating knowledge about candidate genes, phenotypic data as well as ecological and environmental data.

## Acknowledgements

We would like to thank Jean Bousquet, Sam Yeaman and one anonymous reviewer for their useful suggestions. We are grateful to Andrew Eckert, Santiago Gonzalez-Martinez, Jason Holliday and Pär Ingvarsson for sharing some of their data for the review of the experimental surveys.

## References

- Aho ML (1994) Autumn frost hardening of one-year old *Pinus sylvestris* (L) seedlings. Effect of origin and parent trees. *Scandinavian Journal of Forest Research*, **9**, 17–24.
- Albert AYK, Sawaya S, Vibes TH *et al.* (2008) The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution*, **62**, 76–85.
- Alberto F, Niort J, Derory J *et al.* (2010) Population differentiation of sessile oak at the altitudinal front of migration in the French Pyrenees. *Molecular Ecology*, **19**, 2626–2639.
- Alberto F, Bouffier L, Louvet JM, Delzon S, Kremer A (2011) Adaptive responses for seed and leaf phenology in natural populations of sessile oak along altitudinal gradients. *Journal of Evolutionary Biology*, **24**, 1442–1454.
- Atwell S, Huang YS, Vilhjálmsson BJ *et al.* (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature*, **465**, 627–631.
- Barreiro LB, Laval G, Quach H, Patin E, Quintana-Murci L (2008) Natural selection has driven population differentiation in modern humans. *Nature Genetics*, **40**, 340–345.
- Beaumont MA (2005) Adaptation and speciation: what can FST tell us? *Trends in Ecology and Evolution*, **20**, 435–440.
- Beaumont MA, Balding JB (2004) Identifying adaptive genetic divergence among 517 populations from genome scans. *Molecular Ecology*, **13**, 969–980.
- Björklund M, Ranta E, Kaitala V, Bach LA, Lundberg P, Stenseth NC (2009) Quantitative trait evolution and environmental change. *PLoS ONE*, **4**, e4521.
- Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution*, **23**, 773–783.
- Buckler ES, Holland JB, Bradbury PJ *et al.* (2009) The genetic architecture of maize flowering time. *Science*, **325**, 714–718.
- Bulmer MG (1980) *The Mathematical Theory of Quantitative Genetics*. Clarendon Press, Oxford, New York: Oxford University Press.
- Bulmer MG (1989) Maintenance of genetic variability by mutation-selection balance: a child's guide through the jungle. *Genome*, **31**, 761–767.
- Chevin LM, Hospital F (2008) Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics*, **180**, 1645–1660.
- Derory J, Scotti-Saintagne C, Bertocchi E *et al.* (2010) Contrasting correlations between diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. *Heredity*, **104**, 438–448.
- Ducousso A, Louvet JM, Jarret P, Kremer A (2005) Geographic variations of sessile oaks in French provenance tests. In: *Proceedings of the Joint Meeting of IUFRO Working Groups Genetic of oaks and Improvement and Silviculture of oaks* (eds Rogers R, Ducousso A, Kanazashi A), pp. 128–138. FFPRI (Forestry and Forest Products Research Institute) Scientific Meeting Report 3, Tsukuba, Japan.
- Dumilil J, Hardy OJ, Petit RJ (2009) Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evolutionary Biology*, **9**, 177.
- Eckert AJ, Bower AD, Wegrzyn JL *et al.* (2009) Association genetics of coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics*, **182**, 1289–1302.
- Eveno E, Collada C, Guevara MA *et al.* (2008) Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. *Molecular Biology and Evolution*, **25**, 417–437.
- Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.
- Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*. Addison Wesley Longman, New York City, New York.
- Fisher RA (1918) The correlation between relatives under the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, **52**, 399–433.
- Flint J, Mackay TFC (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research*, **19**, 723–733.
- Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics*, **180**, 977–993.
- Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extent and scale of local adaptation in

- salmonid fishes: review and meta-analysis. *Heredity*, **106**, 404–420.
- Gavrilets S, Hastings A (1995) Dynamics of polygenic variability under stabilizing selection, recombination, and drift. *Genetical Research*, **65**, 63–74.
- Gimelfarb A (1989) Genotypic variation for a quantitative character maintained under stabilizing selection without mutations: epistasis. *Genetics*, **123**, 217–227.
- Goldstein DB, Holsinger KE (1992) Maintenance of polygenic variation in spatially structured populations – roles for local mating and genetic redundancy. *Evolution*, **46**, 412–429.
- Gonzalez-Martinez SC, Krutovsky KV, Neale DB (2006) Forest tree population genomics and adaptive evolution. *New Phytologist*, **170**, 227–238.
- Goudet J, Buchi L (2006) The effects of dominance, regular inbreeding and sampling design on  $Q_{ST}$ , an estimator of population differentiation for quantitative traits. *Genetics*, **172**, 1337–1347.
- Goudet J, Martin G (2007) Under neutrality,  $Q_{ST} \neq F_{STQ}$  when there is dominance in an island model. *Genetics*, **176**, 1371–1374.
- Griswold CK (2006) Gene flow's effect on the genetic architecture of a local adaptation and its consequences for QTL analyses. *Heredity*, **96**, 445–453.
- Haldane JBS (1954) The measurement of natural selection. In: *Proceedings of the 9th International Congress of Genetics (Caryologia, Supplement to Volume 6) 1*: 480–487.
- Hall D, Luquez V, Garcia VM, St Onge KR, Jansson S, Ingvarsson PK (2007) Adaptive population differentiation in phenology across a latitudinal gradient in European Aspen (*Populus tremula*, L.): a comparison of neutral markers, candidate genes, and phenotypic traits. *Evolution*, **61**, 2849–2860.
- Hancock AM, Witonsky DB, Ehlerer E *et al.* (2010a) Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. *Proceedings of the National Academy of Sciences USA*, **107**, 8924–8930.
- Hancock AM, Alkorta-Aranburu G, Witonsky DB, Di Rienzo A (2010b) Adaptations to new environments in humans: the role of subtle allele frequency shifts. *Philosophical Transactions of the Royal Society, B*, **365**, 2459–2468.
- Hancock AM, Witonsky DB, Alkorta-Aranburu G *et al.* (2011) Adaptations to climate-mediated selective pressures in humans. *PLoS Genetics*, **7**, e1001375.
- Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.
- Heidel-Fischer HM, Vogel H, Heckel DG, Wheat CM (2010) Microevolutionary dynamics of a macroevolutionary key innovation in a Lepidopteran herbivore. *BMC Evolutionary Biology*, **10**, 60.
- Hemmer-Hansen J, Nielsen EE, Frydenberg J, Loeschcke V (2007) Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). *Heredity*, **99**, 592–600.
- Hendry AP, Day T, Taylor EB (2001) Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. *Evolution*, **55**, 459–466.
- Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, **173**, 579–588.
- Hermisson J, Pennings PS (2005) Soft sweeps – molecular population genetics of adaptation from standing genetic variation. *Genetics*, **169**, 2335–2352.
- Heuertz M, De Paoli E, Källmann T *et al.* (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium, and demographic history of Norway Spruce. *Genetics*, **174**, 2095–2105.
- Hoffmann AA, Willi Y (2008) Detecting genetic responses to environmental change. *Nature Reviews Genetics*, **9**, 421–432.
- Hohenlohe PA, Phillips PC, Cresko W (2010) Using population genomics to detect selection in natural populations: key concepts and methodological considerations. *International Journal of Plant Science*, **171**, 1059–1071.
- Holderegger R, Herrmann D, Poncet B *et al.* (2008) Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology and Diversity*, **1**, 273–283.
- Holliday JA, Ritland K, Aitken SN (2010) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytologist*, **188**, 501–514.
- Jaramillo-Correa JP, Beaulieu J, Bousquet J (2001) Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. *Molecular Ecology*, **10**, 2729–2740.
- Joost S, Bonin A, Bruford MW *et al.* (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, **16**, 3955–3969.
- Jumbo-Lucioni P, Ayroles JF, Moses Chambers M *et al.* (2010) Systems genetics analysis of body weight and energy metabolism traits in *Drosophila melanogaster*. *BMC Genomics*, **11**, 297.
- Karl I, Schmitt T, Fischer K (2009a) Genetic differentiation between alpine and lowland populations of a butterfly is related to PGI enzyme genotype. *Ecography*, **32**, 488–496.
- Karl I, Sorensen JG, Loeschcke V, Fischer K (2009b) HSP70 expression in the copper butterfly *Lycaena tityrus* across altitudes and temperatures. *Journal of Evolutionary Biology*, **22**, 172–178.
- Kavanagh KD, Haugen TO, Gregersen F, Jernvall J, Vøllestad LA (2010) Contemporary temperature-driven divergence in a Nordic freshwater fish under conditions commonly thought to hinder adaptation. *BMC Evolutionary Biology*, **10**, 350.
- Kim Y, Stephan W (2002) Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics*, **160**, 765–777.
- Kimura M, Crow JF (1978) Effect of overall phenotypic selection on genetic change at individual loci. *Proceedings of the National Academy of Sciences USA*, **75**, 6168–6171.
- Kremer A, Le Corre V (2011) Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity*, in press.
- Larmuseau MAD, Raeymaekers JAM, Ruddick KG (2009) To see in different seas: spatial variation in the rhodopsin gene of the sand goby (*Pomatoschistus minutus*). *Molecular Ecology*, **18**, 4227–4239.
- Latta RG (1998) Differentiation of allelic frequencies at quantitative trait loci affecting adaptive locally traits. *The American Naturalist*, **151**, 283–292.



- Latta RG (2004) Gene flow, adaptive population divergence and comparative population structure across loci. *New Phytologist*, **161**, 51–58.
- Le Corre V (2005) Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. *Molecular Ecology*, **14**, 4181–4192.
- Le Corre V, Kremer A (2003) Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. *Genetics*, **164**, 1205–1219.
- Lopez S, Rousset F, Shaw FH, Shaw RG, Ronce O (2008) Migration load in plants: role of pollen and seed dispersal in heterogeneous landscapes. *Journal of Evolutionary Biology*, **586**, 294–309.
- Lopez-Fanjul C, Fernandez A, Toro MA (2003) The effect of neutral nonadditive gene action on the quantitative index of population divergence. *Genetics*, **164**, 1627–1633.
- Lopez-Fanjul C, Fernandez A, Toro MA (2007) 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.
- Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.
- Ma XF, Hall D, Katherine R, St Onge KR, Jansson S, Ingvarsson PK (2010) Genetic differentiation, clinal variation and phenotypic associations with growth cessation across the *Populus tremula* photoperiodic pathway. *Genetics*, **186**, 1033–1044.
- Mackay TFC, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics*, **10**, 565–577.
- Marliac C, Jéhin L, Saïdou AA *et al.* (2011) Genetic basis of pearl millet adaptation along an environmental gradient investigated by a combination of genome scan and association mapping. *Molecular Ecology*, **20**, 80–91.
- McCracken KG, Bulgarella M, Johnson KP *et al.* (2009) Gene flow in the face of countervailing selection: adaptation to high-altitude hypoxia in the beta A hemoglobin subunit of yellow-billed pintails in the Andes. *Molecular Biology and Evolution*, **26**, 815–827.
- McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology and Evolution*, **17**, 285–291.
- Moeller DA, Tiffin P (2008) Geographic variation in adaptation at the molecular level: a case study of plant immunity genes. *Evolution*, **62**, 3069–3081.
- Namroud MC, Beaulieu J, Juge N *et al.* (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Molecular Ecology*, **17**, 3599–3613.
- Narum SR, Hess JE (2011) Comparison of  $F_{ST}$  outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.
- Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nature Reviews Genetics*, **12**, 111–122.
- Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic 602 divergence. *Molecular Ecology*, **18**, 375–402.
- Nunes VL, Beaumont MA, Butlin RK, Paulo OS (2011) Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (*Lacerta lepida*) along an environmental gradient. *Molecular Ecology*, **20**, 193–205.
- Ohta T (1982) Linkage disequilibrium due to random genetic drift in finite subdivided populations. *Proceedings of the National Academy of Sciences*, **79**, 1940–1944.
- O'Malley KG, Camara MD, Banks MA (2007) Candidate loci reveal genetic differentiation between temporally divergent migratory runs of Chinook salmon (*Oncorhynchus tshawytscha*). *Molecular Ecology*, **16**, 4930–4941.
- Pérez-Figueroa A, Garcia-Pereira MJ, Saura M, Rolan-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, **23**, 2267–2276.
- Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. *Molecular Breeding*, **7**, 275–291.
- Postma E, van Noordwijk AJ (2005) Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*, **433**, 65–68.
- Pritchard JK, Di Rienzo A (2010) Adaptation – not by sweeps alone. *Nature Reviews Genetics*, **11**, 665–667.
- Prunier M, Laroche J, Beaulieu J, Bousquet J (2011) Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. *Molecular Ecology*, **20**, 1702–1716.
- Pyhäjärvi T, García-Gil MR, Knürr T, Mikkonen M, Wachowiak W, Savolainen O (2007) Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics*, **177**, 1713–1724.
- Raeymaekers JA, Van Houdt J, Larmuseau MHD, Geldof S, Wolckaert FAM (2007) Divergent selection as revealed by PST and QTL-based FST in three-spined stickleback (*Gasterosteus aculeatus*) populations along a coastal-inland gradient. *Molecular Ecology*, **16**, 891–905.
- Rasanen K, Hendry AP (2008) Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, **11**, 624–636.
- Renaut S, Nolte AW, Rogers SM, Derome N, Bernatchez L (2011) SNP signatures of selection on standing variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). *Molecular Ecology*, **20**, 545–559.
- Reusch TBH, Wood TE (2007) Molecular ecology of global change. *Molecular Ecology*, **16**, 3973–3992.
- Rhoné B, Vitalis R, Goldringer I, Bonin I (2010) Evolution of flowering time in experimental wheat populations: a comprehensive approach to detect genetic signatures of natural selection. *Evolution*, **64**, 2110–2125.
- Rogers SM, Bernatchez L (2005) Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). *Molecular Ecology*, **14**, 351–361.
- Saint-Laurent R, Legault M, Bernatchez L (2003) Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchell). *Molecular Ecology*, **12**, 315–330.
- Salomé PA, Bomblies K, Laitinen RAE, Yant L, Mott R, Weigel D (2011) Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. *Genetics*, **188**, 421–433.
- Sanchez-Gomez D, Majada J, Alia R, Feito I, Aranda I (2010) Intraspecific variation in growth and allocation patterns in



- seedlings of *Pinus pinaster* Ait. submitted to contrasting watering regimes: can water availability explain regional variation? *Annals of Forest Science*, **67**, 505.
- Santure AW, Wang J (2009) The joint effects of selection and dominance on the  $Q_{ST} - F_{ST}$  contrast. *Genetics*, **181**, 259–276.
- Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 595–619.
- Spitze K (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.
- St Clair JB (2006) Genetic variation in fall cold hardiness in coastal Douglas-fir in western Oregon and Washington. *Canadian Journal of Botany*, **84**, 1110–1181.
- Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, **100**, 158–170.
- Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.
- Storz JF, Kelly JK (2008) Effects of spatially varying selection on nucleotide diversity and linkage disequilibrium: insights from deer mouse globin genes. *Genetics*, **180**, 367–379.
- Stranger BE, Stahl EA, Towfique R (2011) Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, **187**, 367–383.
- Tufto J (2000) Quantitative genetic models for the balance between migration and stabilizing selection. *Genetical Research*, **76**, 285–293.
- Turelli M, Barton N (1990) Dynamics of polygenic characters under selection. *Theoretical Population Biology*, **38**, 1–57.
- Wachowiak W, Balk PA, Savolainen O (2009) Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). *Tree Genetics & Genomes*, **5**, 117–132.
- 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, Guillaume M (2009) Testing for spatially divergent selection: comparing  $Q_{ST}$  to  $F_{ST}$ . *Genetics*, **183**, 1055–1063.
- Xia H, Camus-Kulandaivelu L, Stephan W, Tellier A, Zhang ZW (2010) Nucleotide diversity patterns of local adaptation at drought-related candidate genes in wild tomatoes. *Molecular Ecology*, **19**, 4144–4154.
- Yang J, Benyamin B, McEvoy BP *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, **42**, 565–569.
- Yeaman S, Guillaume F (2009) Predicting adaptation under migration load: the role of genetic skew. *Evolution*, **63**, 2926–2938.
- Yeaman S, Whitlock M (2011) The genetic architecture of adaptation under migration-selection balance. *Evolution*, **65**, 1897–1911.

---

V.L.C. is broadly interested in the ecology and population genetics of plants, with an emphasis on processes of local adaptation in metapopulations. Her experimental work focuses on the adaptation of weedy plant species to anthropogenic disturbances and agricultural environments. A.K. has long-standing interests in the evolution of forest trees with particular emphasis on population differentiation at various levels where diversity is expressed (from genes to phenotypes).

---