ELSEVIER

Adaptation and speciation: what can F_{st} tell us?

Mark A. Beaumont

School of Animal and Microbial Sciences, The University of Reading, Whiteknights, PO Box 228, Reading, UK, RG6 6AJ

A useful way of summarizing genetic variability among different populations is through estimates of the inbreeding coefficient, F_{st}. Several recent studies have tried to use the distribution of estimates of F_{st} from individual genetic loci to detect the effects of natural selection. However, the promise of this approach has yet to be fully realized owing to the pervasive dogma that this distribution is highly dependent on demographic history. Here, I review recent theoretical results that indicate that the distribution of estimates of F_{st} is generally expected to be robust to the vagaries of demographic history. I suggest that analyses based on it provide a useful first step for identifying candidate genes that might be under selection, and explore the ways in which this information can be used in ecological and evolutionary studies.

Introduction

It is becoming increasingly cheaper and faster to survey samples genetically from model and non-model organisms for a large number of loci across their genomes, an advance that derives largely from the activities of the biomedical community [1]. Here, I reappraise an old idea [2] for the analysis of such data, first proposed when multi-locus surveys were scarce and difficult to obtain; it was later discredited [3,4] and abandoned; and now, fuelled by the plentiful supply of these data, occasionally peeps apologetically out of research articles, smothered in caveats [5].

The ready availability of different classes of gene frequency information has rekindled an interest in natural selection and the development of a variety of methods for use in trying to infer the presence and mode of selection. Three main approaches can be identified [6]: (i) detailed modelling of selection at individual loci or sequences; (ii) multilocus comparisons, of which the Lewontin–Krakauer Method (see Glossary [2]), discussed here, is the oldest; and (iii) comparison of patterns of substitution among synonymous and non-synonymous sites. Analyses of the advantages and drawbacks of these different approaches are detailed by Nielsen [6]. Much of the research has been driven by the biomedical community with an aim to identify and characterize biochemical function and the phenotypic effect of natural variation throughout the human genome, often based on comparative analysis [7]. Judgements about the efficacies

of different methods implicitly tend to have these, ultimately medical, goals in mind.

In evolutionary biology there are, by contrast, several interesting hypotheses that can be tested by characterizing the number, position and fitness effects of genomic regions that show apparent adaptive divergence in allele frequency, without the need to delve into the physiological details. I argue that, for most organisms, the easiest way to achieve this is by using the Lewontin–Krakauer method, which appears, at least in recent versions, to be generally robust to the vagaries of demographic history. I invoke recent theoretical results that suggest why this is not so surprising and discuss possible sampling strategies that might maximize the power of the approach. I also outline areas of application, particularly the study of adaptive divergence in the face of gene flow and modes of speciation.

Inbreeding coefficients and the identification of loci subject to selection

The study of selection, particularly local adaptation, at the genetic level has a long history (usefully reviewed in [8]).

<div style="border:1px solid;background:#fffde0;padding:4px">

Glossary

AFLP: amplified fragment length polymorphism. A way of assaying nucleotide sites for polymorphisms, typically resulting in a dominant marker system. A relatively inexpensive way of obtaining many markers.

Ascertainment bias: bias in demographic inferences owing to the use of (typically) low mutation rate markers, such as SNPs, that have been previously identified in earlier smaller scale studies. The SNPs so identified will form a biased subset, with alleles at intermediate frequencies (otherwise they would not have been found in the first place) in the populations from which they were first surveyed.

Balancing selection: selection where gene frequencies tend to some equilibrium that maintains polymorphism.

Inbreeding coefficient: the probability that two genes at a locus are identical by descent. Classically refers to two genes within an individual, but currently used more widely, as here.

Lewontin–Krakauer method: a method that uses the variance in estimates of F_{st} to determine whether non-neutral loci are present in sample of loci whose allele frequencies are surveyed in several different populations.

Multinomial-Dirichlet formula: the multinomial-Dirichlet distribution gives the distribution of frequencies in a multinomial sample from a Dirichlet distribution. The Dirichlet distribution is the multivariate generalization of the beta distribution, which is useful for modelling proportions.

RFLP: restriction fragment length polymorphism. A way of assaying nucleotides sites for polymorphisms; important in the days before the use of the PCR.

Separation-of-timescales: a way of simplifying the description of a complicated system by separating out processes that occur rapidly from those that occur slowly. At any time, the fast processes are assumed to be at their equilibrium values, and the slow processes can then be made functions of these equilibrium values.

SNP: single nucleotide polymorphism; DNA sites known to be polymorphic, typically from earlier, smaller scale, studies.

Transient model: a model in which the quantities of interest change through time, as opposed to an equilibrium model.

</div>

Corresponding author: Beaumont, M.A. (M.A.Beaumont@reading.ac.uk).

Examples include the study of local crypsis in response to bird predation in the snail *Cepaea nemoralis* [9] and the peppered moth *Biston betularia* [10]. The genetics of adaptation was largely eclipsed during the 1980s and 1990s by an interest in the possibility of recovering the historical demography of populations through an analysis of genetic variation, in particular, of mitochondrial DNA sequence variation [11], a project that demands the absence of selection.

In response to the increasingly heated neutralist–selectionist controversy, Lewontin and Krakauer, stimulated by earlier suggestions of Cavalli-Sforza [12], proposed a test for identifying loci that are subject to selection that depended on estimating the variation in the apparent degree of inbreeding among different loci [2] (Box 1). Inbreeding is measured by the demographic parameter F_{st}, defined here to be the probability that two genes within a deme share a common ancestor within that deme. It can be estimated in several ways, most commonly now by the method of Weir and Cockerham [13]. This typically assumes that each deme has the same F_{st}, but can also be generalized to allow separate F_{st}s for each deme [14,15]. Importantly, inbreeding here refers to the departure of gene frequencies within a subpopulation from some metapopulation average as a result of random genetic drift within the subpopulation, rather than inbreeding of individuals owing to the ancestry of their parents within the subpopulation.

The intuition behind the Lewontin–Krakauer test is simple: loci in which different alleles are selectively favoured in different populations should exhibit larger allele frequency differences between populations than do loci with purely neutral alleles. Similarly, loci that are subject to strong balancing selection should have a lower level of genetic differentiation. Unfortunately, inbreeding itself can produce large frequency differences among populations. Lewontin and Krakauer proposed to estimate the variance in F_{st} among loci and to perform a statistical test to compare it with neutral expectations: a significant result indicates the presence of selection. They applied their method (Box 1) to allele frequency data at nine loci in

a worldwide survey of human populations [12], and concluded that there was good evidence of heterogeneity in estimates of F_{st} among loci. The also looked at data from 20 loci surveyed from local populations of native American Yanomama [16], and concluded that there was homogeneity in the estimates.

Avoiding dependence on demography

The test was soon severely criticised [3,4,17–19] (Box 1), and fell out of use. As originally proposed, it relied on several assumptions associated with the need to obtain analytical solutions at a time when extensive numerical simulations were unfeasible. The test also assumed that F_{st} is the same in each deme; that, for each locus, the deviation in allele frequency from the metapopulation average is independently distributed among demes; that mutation rates do not affect estimates of F_{st}. As a result of a series of studies than began during the early 1990s, the Lewontin–Krakauer test has increasingly been the subject of reassessment [14,20,21]. The reliance on analytical approximations has been addressed through the use of computer simulations, which directly give the distribution of F_{st} among loci, and also enable the influence of different demographic scenarios to be explored. For example Bowcock and colleagues [20] used simulations to analyse the distribution of human RFLP data under what was considered at the time a more realistic demographic model, concluding that the observed distribution of F_{st} did not match neutral expectations.

The original conception of the Lewontin–Krakauer test, however, is that the distribution of estimates of F_{st} should be relatively robust to demographic effects, avoiding the need to model the demographics explicitly. This robustness to the effects of demography is supported by a study [21] that directly tested the influence of different models of population structure on the distribution of estimates of F_{st}.

A theoretical foundation for why we might expect the distribution of estimates of F_{st} to be robust to demographic effects comes from studies of gene genealogies in structured populations [22–26]. In many cases, the

---

Box 1. The Lewontin–Krakauer test

Lewontin and Krakauer [2] argued that the expected amount of inbreeding experienced at different loci should be the same because of the shared demographic history experienced at those loci. Selection, by contrast, can alter the apparent degree of inbreeding at individual loci.

Behind the Lewontin–Krakauer test, and indeed, any measurement of differentiation, is the idea that there is some baseline metapopulation allele frequency, denoted for a particular allele at the *l*th locus by p_l. Inbreeding then allows the deme allele frequencies, a_{lj}, to be shifted from this baseline; a result that can be justified under several different models [15] is that $\text{Var}[a_{lj}] = p_l(1-p_l)F_{stj}$, where F_{stj} is the inbreeding coefficient in the *j*th deme.

The Lewontin–Krakauer test, as originally proposed, involves an approximation that can be laid out as follows (see also [15]). Suppose that all the F_{stj} are identical and denoted by F_{st}, and are estimated across all loci with sufficient precision that we can equate this estimate with F_{st}. Similarly, we assume that p_l is estimated with sufficient precision that we can equate the estimate with p_l. F_{st} is then parametrically $\frac{\text{Var}[a_{lj}]}{p_l(1-pl)}$. Each estimate of F_{st} for a particular locus, $\hat{F}_{st}$, is then a sample variance divided by $p_l(1-p_l)$. Making the

assumption that the distribution of a_{lj} is approximately normal, $(n-1)\hat{F}_{st}/F_{st}$ is expected to be chi-square with $n-1$ degrees of freedom, where $n=d$, the number of subpopulations. The variance of $\hat{F}_{st}$ among loci is then $2\hat{F}_{st}^2/(n-1)$. The same approximation can be extended to multiple alleles, in which case n is given by $(d-1)(K-1)$, where K is the number of alleles.

The test was initially criticised in 1975 by Nei and Maruyama, and Robertson [3,4,17–19]. The main criticisms were: (i) if there is a high rate of migration among a subset of demes, then their gene frequencies will be correlated contrary to the assumptions of the test; and (ii) at the molecular level, most mutations are unique and the expected $\hat{F}_{st}$ for particular alleles will vary. They pointed out that both these effects could substantially inflate the variance in estimates of F_{st}.

Another criticism is that the chi-square approximation, although a useful rule-of-thumb guideline, is unlikely ever to be particularly accurate. For example, it requires large sample sizes, identical numbers of alleles at each locus, equal F_{st} among demes, and intermediate baseline frequencies. However, given the ease of computer simulation methods, there is no need for such an additional level of approximation.

genealogy naturally tends towards a simple structure and can be described mathematically through a separation-of-timescales approximation in which the genealogy is broken up into two phases (Figure 1): (i) the recent genealogy of each deme, called the 'scattering phase' by John Wakeley [23]; and (ii) the ancestral genealogy of the metapopulation as a whole (the 'collecting phase'), which typically behaves as if from a single population. In this approximation, there are no mutations in the scattering phase, and the mean and variance in gene frequency in each deme depends only on F_{st} and the gene frequencies in the collecting phase. Given samples from independent demes, we can estimate the collecting-phase gene frequency without the need to model its demographic history, and thereby estimate F_{st} from the deme gene frequency. If the separation-of-timescales approximation holds, most of the information that is necessary to model the gene frequencies under any neutral structured population model, and thereby detect the signatures of selection in the form of outliers, is provided by a multinomial-Dirichlet formula [27–30]. Based on this, a Bayesian approach has been developed that can detect loci subject to selection [31].

Alternative methods

The more recent studies have not invalidated the original criticisms of the Lewontin–Krakauer test [3,4,17–19], but, instead, suggest that they are often not applicable to the real world. The criticisms can be re-expressed as violations of the separation-of-timescales approximation. Potentially problematic are high mutation rate loci, such as microsatellites, in which mutations occur in the scattering phase [32]. Another problem arises when the gene frequency in the collecting phase is not the same for each deme [4,17]. One way of overcoming this is to restrict the analysis to pairwise comparisons [14,33]. Two issues that need to be explored further in this regard are the

potential loss of power in reducing the number of subpopulations to just two, and how best to protect against false positives in the large number of pairwise comparisons that may be needed. Simulations from a large two-dimensional stepping-stone model with differing configurations of neighbouring samples, and measurable isolation-by-distance [21] did not lead to strong effects on the distribution of F_{st}, an observation that is supported by the recent analysis of Wilkins [26].

If the separation-of-timescales approximation is believed not to hold for a particular system, it might be necessary to abandon the simplicity of the Lewontin–Krakauer approach and to model the effects of demographic history in detail [34–36]; however, it might be found that detailed modelling gives similar results to standard methods [35]. The detection of outliers can be regarded as part of some model-fitting exercise when inferring demographic parameters [36,37]. Indeed, another motivation for detecting outliers is to make robust inferences about demographic history, unaffected by aberrant loci [38–40]. Ideally, rather than detect outliers from a neutral model, it would be statistically more efficient to model selection at each locus, as well as the demographic history, but this must be considered a long-term goal.

Alternative multilocus simulation-based tests that use summary statistics other than F_{st} have been proposed by Schlötterer and colleagues [41–44], and involve the idea of a 'selective sweep', a useful conceptual metaphor that arises from positive (directional) selection [42,45]. Underlying it is a transient model, in which, for example, a population recently invades a new environment. This model can be contrasted with the underlying assumption behind the method of Beaumont and Balding [31] of a mutation–drift–selection equilibrium, in which there is gene flow between demes, particular alleles have different selective advantages in different demes, and the expected degree of differentiation (measured, for example, by estimates of F_{st}) remains the same over time.

Although there are uncertainties associated with the robustness of the Lewontin–Krakauer approach, there are also uncertainties associated with these other approaches: Can we be confident that we have chosen the correct demographic and mutational model? How robust are alternative estimators to variability in demographic history? With tests based on that of Lewontin and Krakauer, there is a good theoretical foundation and, in my opinion, for most data sets it is a worthwhile first approach for identifying loci subject to selection.

However, there is now no longer a single Lewontin–Krakauer test, but several approaches based on the original idea. These can give different answers based on the same data. For example Vitalis *et al.* [14] obtained different answers using their method on one data set from those obtained using the test of Beaumont and Nichols [21]. Similarly Beaumont and Balding [31] observed that their method, and that of Beaumont and Nichols [21], gave somewhat different results based on the same data. Given the differences between the methods, these results are unsurprising, but also point to the need for more detailed
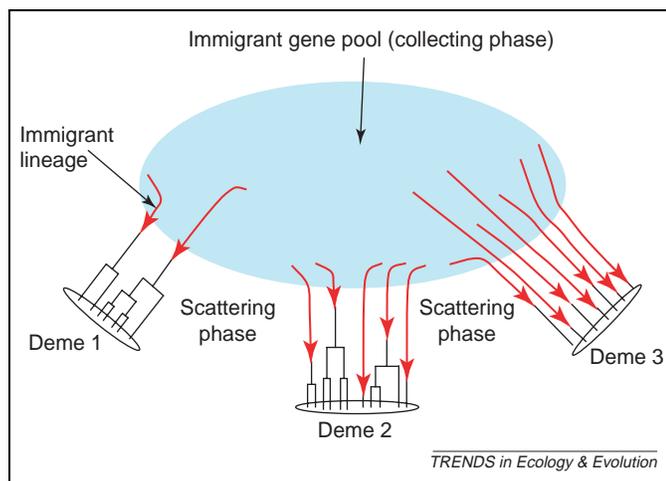
Figure 1. The separation-of-timescales model. The genealogies of samples typed at a locus are shown for three demes, which correspond to the scattering phase. The ancestor of each cluster of lineages is an immigrant drawn from a gene pool (shown as red lines). The gene frequencies in the immigrant gene pool are determined by the collecting phase genealogy (not shown), and the mutation model. If F_{st} is high, the sample in a deme will have few immigrant ancestors (e.g. Deme 1), and the sample frequencies will tend to be different from those in the immigrant gene pool. If F_{st} is low (e.g. Deme 3), the sample frequencies will tend to be similar to those in the immigrant gene pool.

power-testing to highlight the advantages and drawbacks of different approaches.

Design of surveys
How should surveys be designed to maximize the chances of picking up loci that are subject to selection? The approximation given by Lewontin and Krakauer (Box 1) can be used to obtain some idea of the expected variability in estimates of F_{st} among loci. This shows that there is substantial variability and skew in estimates of F_{st} when biallelic markers are surveyed in only a pair of populations, and hence there is potentially little power to detect outlier loci unless selection is strong relative to the immigration rate. Low mutation rate markers, such as SNPs and AFLPs, might also suffer from problems of ascertainment bias if the loci differ in how they were ascertained. Ideally, one should use markers with a large number of alleles surveyed in a large number of demes. However, the more alleles at a locus, the more likely it is that there are mutations in the scattering phase, which will affect the distribution of estimates of F_{st} [32,35]; and the more demes surveyed, the less likely it is that their observed deviations in gene frequency from the meta-population mean are independent of each other. It might be most efficient to survey short sequences, widely spaced around the genome, typically with three or four SNPS, in a large number of samples, covering a wide range of environments. Potential problems of correlations of gene frequencies among different demes can then be explored by studying subsets of samples, chosen to minimize correlations.

Example applications
There have been an increasing number of studies that aim to identify loci subject to selection [46,47], often using the distribution of F_{st} among loci. An illustrative case stems from the work of Pogson and colleagues [48]. They studied a mixture of RFLPs and allozymes in populations of Atlantic cod *Gadus morhau*. On the basis of tests closely related to those of Lewontin and Krakauer [2], they showed that the mean F_{st} in allozymes was lower than that of RFLPs, and suggested that the allozymes were under balancing selection and the RFLPs were neutral. As a result, they therefore suggested that cod populations were much more structured than hitherto believed. A reanalysis of the cod data [48] by Beaumont and Nichols [21] noted that most RFLPs behaved similarly to the allozymes, and all the results could be explained by two 'rogue' RFLPs. One is in the pantophysin locus, now much studied [49,50], for which there is additional strong evidence of adaptive selection from detailed sequence analysis, although the mechanism of selection is unknown. This has led to a reassessment of the degree of substructure in the cod [50].

Another example where the distribution of F_{st} has been used to identify candidate loci, and where later independent observations support previous conclusions, comes from studies by Storz and colleagues on the rodent genus *Peromyscus* (which includes the deermouse, *Peromyscus maniculatus*) [51,52]. A re-examination of allozyme frequencies among different populations in each of four species of *Peromyscus*, using the distribution of F_{st} in simulations from island models, has detected good evidence of selection on some loci, particularly that encoding albumin [51]. The different allelic variants at the albumin locus appear to be correlated with altitude. A subsequent study on new samples from different populations of the deermouse along an altitudinal gradient upholds the evidence of earlier surveys that albumin appears to be adaptively differentiated [52], possibly in response to different levels of anoxia.

Testing models of adaptation and speciation
Once interesting genomic regions have been identified through an analysis of F_{st}, or related method, how might the information be used? In conjunction with demographic information, such as immigration rate, it might be possible to quantify the distribution of fitness effects that are necessary to lead to the observed distribution of estimates of F_{st}, and thereby test evolutionary hypotheses [47]. Although predictions of phenotypic evolution are often still based on the infinitesimal model [53], it is likely that genes of large effect are prevalent [53], possibly in the tails of an exponential distribution of effects [54,55], and analyses of F_{st} might help quantify this.

F_{st}-based surveys might also help to quantify levels of adaptive differentiation among populations and recently derived species. For example, recent molecular genetic analysis has uncovered widespread evidence of gene flow between species previously thought to be reproductively isolated, resulting in the reappraisal of the importance of strict allopatry in speciation [56]. This has led to the suggestion that speciation is better viewed from a genic perspective [57,58]: genomic regions that are subject to strong disruptive selection or assortative-mating are effectively in two different species, whereas other regions are in a single structured population. An F_{st}-based survey of genetically mapped loci might highlight these different regions, and also provide a means of comparing the degree of adaptive differentiation among different species. The genus *Geospiza*, one of the genera comprising Darwins finches on the Galapagos, might be a strong candidate group to study in this regard. In spite of strong phenotypic differentiation, inferred mtDNA gene-trees and nuclear ITS (internal transcribed spacer) gene trees appear to be uncorrelated either with each other, or with traditional taxonomic groups and their geographical location [59]. Thus, it might be that, what are called 'species' consist of several regions in the genome that, by virtue of their phenotypic effects, are under strong disruptive selection and/or subject to assortative mating, and thus have high F_{st}; whereas the remainder of the genome is exchanged relatively freely among populations and has lower F_{st}. Seen from a metapopulation perspective, these cases are no different from the complex polymorphisms observed in, for example, *Cepaea*.

A recent study highlights the usefulness of F_{st}-based surveys to pick out regions of the genome that appear to be subject to strong disruptive selection and/or assortative mating [60]. In this case, two morphologically differentiated parapatric populations of periwinkle *Littorina saxatilis*, surveyed for almost 300 AFLP markers, have

been shown to exhibit strong differentiation at 15 markers. In this example, the ability to identify outliers depends on the presumed 'neutral' F_{st} being low (~ 0.01–0.02), whereas the F_{st} of the outliers is typically > 0.4, and this pattern holds for the same 15 markers in different paired samples collected from shores around the UK. This suggests that strong selection is operating somewhere in the genomic region in which these markers reside.

As pointed out by Campbell and Bernatchez [61], an important enhancement to F_{st}-based surveys is the genetic mapping of marker loci [62], because we can then see, for example, whether highly differentiated markers all tend to map in the same location, possibly in a region with reduced recombination such as an inversion. This information will then help us test recent ideas about the importance of reduced recombination in promoting adaptation in the face of gene flow [63–67]. Strong disruptive selection can impede gene flow [68] in the vicinity of the selected locus. In a chromosomal segment with reduced recombination, perhaps brought about by some sort of rearrangement such as an inversion, the effect might extend quite broadly over a chromosome through linkage disequilibrium and make it easier for the fixation of other alleles that are only favourable in one of the populations, thereby increasing the strength of disruptive selection. In many ways, these ideas are similar to the concept of the 'supergene' [8,69] pervasive in classic ecological genetics [8], in which multilocus polymorphisms are held together in a nonrecombining block so that only synergistic sets of alleles segregate, and, antagonistic sets, which would otherwise be generated by independent assortment, are suppressed.

Conclusions

The abandonment of Lewontin and Krakauer's idea could be regarded as a major advance in the wider acceptance of the usefulness of the neutral theory, and the importance of demographic events in shaping gene frequency data. This can then be seen as leading directly to the significant research programme of the 1980s and 1990s, that of trying to recover population history from genetic information. Ironically, however, the reinstatement of Lewontin and Krakauer's ideas depends largely on a theoretical result in which the older demographic history is unrecoverable in principle. It is likely that only relatively recent demographic events can be usefully inferred, and genotypic methods such as [70] point to one possible way forward in their analysis.

There is no doubt that an increasing number of genetic surveys will be carried out on a wide diversity of organisms of intrinsic evolutionary interest. In my opinion, the Lewontin–Krakauer test, in its recent incarnations, largely based on the infinite-island model, provides a useful and fairly robust tool for interpreting the results of such studies. Furthermore, the endpoint of analysis need not be the biochemical characterization of individual genes, but the method can be used to help answer more evolutionarily interesting questions about modes of adaptation and speciation.

References

1 Gibbs, R.A. *et al*. (2003) The International HapMap Project. *Nature* 426, 789–796
2 Lewontin, R.C. and Krakauer, J. (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74, 175–195
3 Nei, M. and Maruyama, T. (1975) Lewontin–Krakauer test for neutral genes. *Genetics* 80, 395
4 Robertson, A. (1975) Remarks on the Lewontin–Krakauer test. *Genetics* 80, 396
5 Akey, J.M. *et al*. (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Res.* 12, 1805–1814
6 Nielsen, R. (2001) Statistical tests of selective neutrality in the age of genomics. *Heredity* 86, 641–647
7 Clark, A.G. *et al*. (2003) Inferring nonneutral evolution from human–chimp–mouse orthologous gene trios. *Science* 302, 1960–1963
8 Ford, E.B. (1975) *Ecological Genetics*, Chapman & Hall
9 Cain, A.J. and Sheppard, P.M. (1950) Selection in the polymorphic land snail *Cepaea nemoralis*. *Heredity* 4, 275–294
10 Kettlewell, H.B.D. (1958) A survey of the frequencies of *Biston Betularia* (L) (Lep.) and its melanic forms in Great Britain. *Heredity* 12, 51–72
11 Avise, J.C. (1994) *Molecular Markers, Natural History, and Evolution*, Chapman & Hall
12 Cavalli-Sforza, L.L. (1966) Population structure and human evolution. *Proc. R. Soc. Lond. B Biol. Sci.* 164, 362–379
13 Weir, B.S. and Cockerham, C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370
14 Vitalis, R. *et al*. (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics* 158, 1811–1823
15 Weir, B.S. and Hill, W.G. (2002) Estimating F-statistics. *Annu. Rev. Genet.* 36, 721–750
16 Arends, T. *et al*. (1966) Intra-tribal genetic differentiation among the Yanomama Indians of Southern Venezuela. *Proc. Natl. Acad. Sci. U. S. A.* 57, 1252–1259
17 Robertson, A. (1975) Gene frequency distribution as a test of selective neutrality. *Genetics* 81, 775–785
18 Nei, M. and Chakravarti, A. (1977) Drift variance of F_{st} and G_{st} statistics obtained from a finite number of isolated populations. *Theor. Popul. Biol.* 11, 307–325
19 Nei, M. *et al*. (1977) Mean and variance of F_{st} in a finite number of incompletely isolated populations. *Theor. Popul. Biol.* 11, 291–306
20 Bowcock, A.M. *et al*. (1991) Drift, admixture, and selection in human-evolution – a study with DNA polymorphisms. *Proc. Natl. Acad. Sci. U. S. A.* 88, 839–843
21 Beaumont, M.A. and Nichols, R.A. (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B Biol. Sci.* 263, 1619–1626
22 Nordborg, M. (1997) Structured coalescent processes on different time scales. *Genetics* 146, 1501–1514
23 Wakeley, J. (1999) Nonequilibrium migration in human history. *Genetics* 153, 1863–1871
24 Wakeley, J. and Aliacar, N. (2001) Gene genealogies in a metapopulation. *Genetics* 159, 893–905
25 Wakeley, J. (2001) The coalescent in an island model of population subdivision with variation among demes. *Theor. Popul. Biol.* 59, 133–144
26 Wilkins, J.F. (2005) A separation-of-timescales approach to the coalescent in a continuous population. *Genetics* 168, 2227–2244
27 Rannala, B. and Hartigan, J.A. (1996) Estimating gene flow in island populations. *Genet. Res.* 67, 147–158
28 Balding, D.J. and Nichols, R.A. (1994) DNA profile match probability calculations: how to allow for population stratification, relatedness, database selection and single bands. *For. Sci. Int.* 64, 125–140
29 Balding, D.J. (2003) Likelihood-based inference for genetic correlation coefficients. *Theor. Popul. Biol.* 63, 221–230

30 Porter, A.H. (2003) A test for deviation from island-model population structure. *Mol. Ecol.* 12, 903–915

31 Beaumont, M.A. and Balding, D.J. (2004) Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* 13, 969–980

32 Flint, J. *et al.* (1999) Minisatellite mutational processes reduce F_{st} estimates. *Hum. Genet.* 105, 567–576

33 Tsakas, S. and Krimbas, C.B. (1976) Testing the heterogeneity of *F* values: a suggestion and a correction. *Genetics* 84, 399–401

34 Kauer, M.O. *et al.* (2003) A microsatellite variability screen for positive selection associated with the 'Out of Africa' habitat expansion of *Drosophila melanogaster*. *Genetics* 165, 1137–1148

35 Storz, J.F. *et al.* (2004) Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. *Mol. Biol. Evol.* 21, 1800–1811

36 Akey, J.M. *et al.* (2004) Population history and natural selection shape patterns of genetic variation at 132 genes. *PLoS Biol.* 2, 1591–1599

37 Galtier, N. *et al.* (2000) Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics* 155, 981–987

38 Luikart, G. *et al.* (2003) The power and promise of population genomics: from genotyping to genome typing. *Nat. Rev. Genet.* 4, 981–994

39 Storz, J.F. *et al.* (2002) Genetic evidence for long-term population decline in a savannah-dwelling primate: inferences from a hierarchical Bayesian model. *Mol. Biol. Evol.* 19, 1981–1990

40 Storz, J.F. and Beaumont, M.A. (2002) Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution* 56, 154–166

41 Schlotterer, C. (2002) Towards a molecular characterization of adaptation in local populations. *Curr. Opin. Genet. Dev.* 12, 683–687

42 Schlotterer, C. (2002) A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics* 160, 753–763

43 Schlotterer, C. (2003) Hitchhiking mapping – functional genomics from the population genetics perspective. *Trends Genet.* 19, 32–38

44 Kayser, M. *et al.* (2003) A genome scan to detect candidate regions influenced by local natural selection in human populations. *Mol. Biol. Evol.* 20, 893–900

45 Przeworski, M. (2002) The signature of positive selection at randomly chosen loci. *Genetics* 160, 1179–1189

46 Guinand, B. *et al.* (2004) How to detect polymorphisms undergoing selection in marine fishes? A review of methods and case studies, including flatfishes. *J. Sea Res.* 51, 167–182

47 Storz, J.F. (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.* 14, 671–688

48 Pogson, G.H. *et al.* (1995) Genetic population-structure and gene flow in the Atlantic cod *Gadus morhua* – a comparison of allozyme and nuclear RFLP loci. *Genetics* 139, 375–385

49 Karlsson, S. and Mork, J. (2003) Selection-induced variation at the pantophysin locus (PanI) in a Norwegian fjord population of cod (*Gadus morhua* L.). *Mol. Ecol.* 12, 3265–3274

50 Pogson, G.H. and Fevolden, S.E. (2003) Natural selection and the genetic differentiation of coastal and Arctic populations of the Atlantic cod in northern Norway: a test involving nucleotide sequence variation at the pantophysin (PanI) locus. *Mol. Ecol.* 12, 63–74

51 Storz, J.F. and Nachman, M.W. (2003) Natural selection on protein polymorphism in the rodent genus *Peromyscus*: evidence from interlocus contrasts. *Evolution* 57, 2628–2635

52 Storz, J.F. and Dubach, J.M. (2004) Natural selection drives altitudinal divergence at the albumin locus in deer mice, *Peromyscus maniculatus*. *Evolution* 58, 1342–1352

53 Barton, N.H. and Keightley, P.D. (2003) Understanding quantitative genetic variation. *Nat. Rev. Genet.* 3, 11–21

54 Otto, S.P. and Jones, C.D. (2000) Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics* 156, 2093–2107

55 Farrall, M. (2004) Quantitative genetic variation: a post-modern view. *Hum. Mol. Genet.* 13, R1–R7

56 Machado, C.A. and Hey, J. (2003) The causes of phylogenetic conflict in a classic Drosophila species group. *Proc. R. Soc. Lond. B Biol. Sci.* 270, 1193–1202

57 Wu, C.I. (2001) The genic view of the process of speciation. *J. Evol. Biol.* 14, 851–865

58 Wu, C.I. and Ting, C.T. (2004) Genes and speciation. *Nat. Rev. Genet.* 5, 114–122

59 Freeland, J.R. and Boag, P.T. (1999) The mitochondrial and nuclear genetic homogeneity of the phenotypically diverse Darwin's ground finches. *Evolution* 53, 1553–1563

60 Wilding, C.S. *et al.* (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *J. Evol. Biol.* 14, 611–619

61 Campbell, D. and Bernatchez, L. (2004) Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Mol. Biol. Evol.* 21, 945–956

62 Emelianov, I. *et al.* (2004) Genomic evidence for divergence with gene flow in host races of the larch budmoth. *Proc. R. Soc. Lond. B Biol. Sci.* 271, 97–106

63 Noor, M.A.F. *et al.* (2001) Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. U. S. A.* 98, 12084–12088

64 Rieseberg, L.H. (2001) Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* 16, 351–358

65 Navarro, A. and Barton, N.H. (2003) Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* 57, 447–459

66 Hey, J. (2003) Speciation and inversions: chimps and humans. *Bioessays* 25, 825–828

67 Zhang, J.Z. *et al.* (2004) Testing the chromosomal speciation hypothesis for humans and chimpanzees. *Genome Res.* 14, 845–851

68 Barton, N.H. and Bengtsson, B.O. (1986) The barrier to genetic exchange between hybridising populations. *Heredity* 56, 357–376

69 Darlington, C.D. and Mather, K. (1949) *The Elements of Genetics*, Allen and Unwin

70 Falush, D. *et al.* (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587