Heredity

Heredity (2013) 110, 409–419
© 2013 Macmillan Publishers Limited All rights reserved 0018-067X/13
www.nature.com/hdy

REVIEW

Inferring population size changes with sequence and SNP data: lessons from human bottlenecks

LM Gattepaille[1], M Jakobsson[1,2] and MGB Blum[3]

Reconstructing historical variation of population size from sequence and single-nucleotide polymorphism (SNP) data is valuable for understanding the evolutionary history of species. Changes in the population size of humans have been thoroughly investigated, and we review different methodologies of demographic reconstruction, specifically focusing on human bottlenecks. In addition to the classical approaches based on the site-frequency spectrum (SFS) or based on linkage disequilibrium, we also review more recent approaches that utilize atypical shared genomic fragments, such as identical by descent or homozygous segments between or within individuals. Compared with methods based on the SFS, these methods are well suited for detecting recent bottlenecks. In general, all these various methods suffer from bias and dependencies on confounding factors such as population structure or poor specification of the mutational and recombination processes, which can affect the demographic reconstruction. With the exception of SFS-based methods, the effects of confounding factors on the inference methods remain poorly investigated. We conclude that an important step when investigating population size changes rests on validating the demographic model by investigating to what extent the fitted demographic model can reproduce the main features of the polymorphism data.

Heredity (2013) **110,** 409–419; doi:10.1038/hdy.2012.120; published online 20 February 2013

Keywords: demographic inference; bottlenecks; out-of-Africa; site-frequency spectrum; SNPs; linkage disequilibrium

INTRODUCTION

Making inferences about demographic history is valuable for understanding the evolutionary trajectories of species. Recovering the temporal fluctuations of the population size and understanding their causes convey information about how species reacted to past events such as glacial cycles, climate change or human-driven perturbations (for example, Ruzzante *et al.*, 2008; Mondol *et al.*, 2009). When deciphering the potential causes of speciation, bottlenecks are also of primary interest since speciation can result from enhanced genetic drift that has been caused by bottlenecks and founder effects (Barton and Charlesworth, 1984). In conservation biology, detecting ongoing bottlenecks is crucial because reductions in genetic diversity may reduce the evolutionary potential of species (England *et al.*, 2003). Indeed, as predicted by population genetic theory, bottlenecks and founder effects can lead to an increased presence of deleterious alleles, which has been shown, for example to occur in humans that went through the out-of-Africa bottleneck (Lohmueller *et al.*, 2008). Furthermore, the construction of an historical model of demography provides a statistical null model that can be used to reveal outlier loci. These outlier loci may display patterns of genetic variation that are at odds with the null model, and could instead be targets of natural selection (Nielsen, 2005; Roux *et al.*, 2013).

Population genetic data have the potential to uncover historical fluctuations of demographic sizes and several approaches to infer population size changes from genetic information have been developed. A particularly well-studied example is the population bottleneck associated with human migration out-of-Africa some 70 000 years ago, when anatomically modern humans colonized the world outside Africa (for example, Voight *et al.*, 2005; Gutenkunst *et al.*, 2009). The human out-of-Africa bottleneck shaped many properties of human variation in non-Africans and, since this event has been extensively investigated, it provides a nice example illustrating the different methodologies of demographic reconstruction. The fruit fly, *Drosophila melanogaster*, also provides additional insights because the relationship between population size changes and polymorphism data has been thoroughly investigated in this species (Haddrill *et al.*, 2005; Thornton and Andolfatto, 2006). Additionally, the fruit fly's demographic history shares a resemblance with human demographic history because of its African origin and a bottleneck that has left signatures in the genomes of non-African *Drosophila melanogaster* populations (Li and Stephan, 2006). Demographic reconstruction based on molecular data has also been undertaken for numerous other species, including *Arabidopsis thaliana* (François *et al.*, 2008), maize and rice (Wright *et al.*, 2005; Caicedo *et al.*, 2007) among many others. Because of the wealth of available data from model organisms, approaches for demographic reconstruction have often been pioneered for these species. In this review, we will use human and fruit fly demographic reconstruction as case studies, including the well-studied bottleneck associated with the out-of-Africa migration. Due to the recent advances in sequencing technologies, these reviewed approaches are quickly becoming applicable for an increasing number of species.

Demographic events can leave traces in polymorphism data by modifying the shape of the genealogical tree of a sample of

[1]Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden; [2]Science for Life Laboratory, Uppsala University, Uppsala, Sweden and [3]Laboratoire TIMC-IMAG UMR 5525, Centre National de la Recherche Scientifique, Université Joseph Fourier, Grenoble, France
Correspondence: Dr MGB Blum, Laboratoire TIMC-IMAG UMR 5525, Centre National de la Recherche Scientifique, Université Joseph Fourier, Grenoble 38706, France.
E-mail: michael.blum@imag.fr

Received 4 July 2012; revised 14 December 2012; accepted 18 December 2012; published online 20 February 2013

chromosomes. A bottleneck can drastically increase the rate of coalescence of lineages and cause severe deviations from the expectations of the standard neutral coalescent with constant population size. If the bottleneck is strong enough, then all the lineages might coalesce during the bottleneck, creating a star-shaped genealogy leading to many rare variants and singletons, usually interpreted as a sign of demographic expansion (Figure 1b). Such a genealogy has short internal branch lengths close to the root of the genealogy, and lead to a lack of common variants when compared with the standard neutral expectation. However, if the strength of the bottleneck is moderate, some lineages might escape from the bottleneck leading to an unexpectedly large proportion of basal lineages, which produces an excess of intermediate frequency variants (Figure 1a). In such cases, the proportion of the time to the most recent common ancestor that is represented by the last few coalescent events can be much greater than the proportion expected under a constant size model. In particular, these genealogies can mimic the shapes of gene genealogies expected in a structured population, although data from multiple loci should provide enough information for distinguishing population bottlenecks from population structure (Peter *et al.*, 2010).

Various types of molecular markers have been employed to infer population sizes and we review methods that use sequence data from multiple loci (or for the entire genome) and single-nucleotide polymorphisms (SNPs), but note that several approaches have also been developed for inferring population size changes based on microsatellite data (Williamson-Natesan, 2005) and based on mitochondrial markers (Ho and Shapiro, 2011). This review is structured into three sections that correspond to the different types of genetic information utilized to recover past population sizes. The first section reviews how the site-frequency spectrum (SFS), or statistical summaries of the SFS, conveys information about past population sizes. The second section deals with haplotypic and linkage disequilibrium (LD)-based methods and the third section presents recently developed methods based on full sequence data. We conclude with a discussion on the caveats that should be kept in mind when reconstructing past population sizes from polymorphism data.

SFS SUMMARY STATISTICS

The empirical SFS is the distribution of allele frequencies at a large number of SNPs. If the ancestral allele is known, then it is possible to reconstruct the SFS that provides the probabilities that an allele/variant is carried by k individuals $p_k = p(k|n)$, $k = 1\ldots(n-1)$, in a sample of n sequences (Marth *et al.*, 2004). Numerous statistical summaries, such as Tajima's *D*, have been proposed to provide one-dimensional summaries of the SFS (Tajima, 1989; Equations (23–26) in Tajima, 1997). The footprints of salient demographic events such as bottlenecks may be detected through these statistical summaries of the SFS, sometimes called class I summary statistics (Ramos-Onsins and Rozas, 2002), and the typical signatures caused by a bottleneck are outlined below.

Tajima's *D*

A well-known consequence of a recent bottleneck involves the reduction of low-frequency polymorphism (Nei *et al.*, 1975), which will generate positive Tajima's *D* values. During a bottleneck, genetic drift is enhanced because of the reduced (effective) population size and low-frequency variants are the most prone to extinction. To show this pattern, it is convenient to consider the scaled SFS, which corresponds to the ratio between the observed SFS and the SFS expected under a panmictic constant-size model (Achaz, 2009). For a recent moderate bottleneck (reduction in effective population size

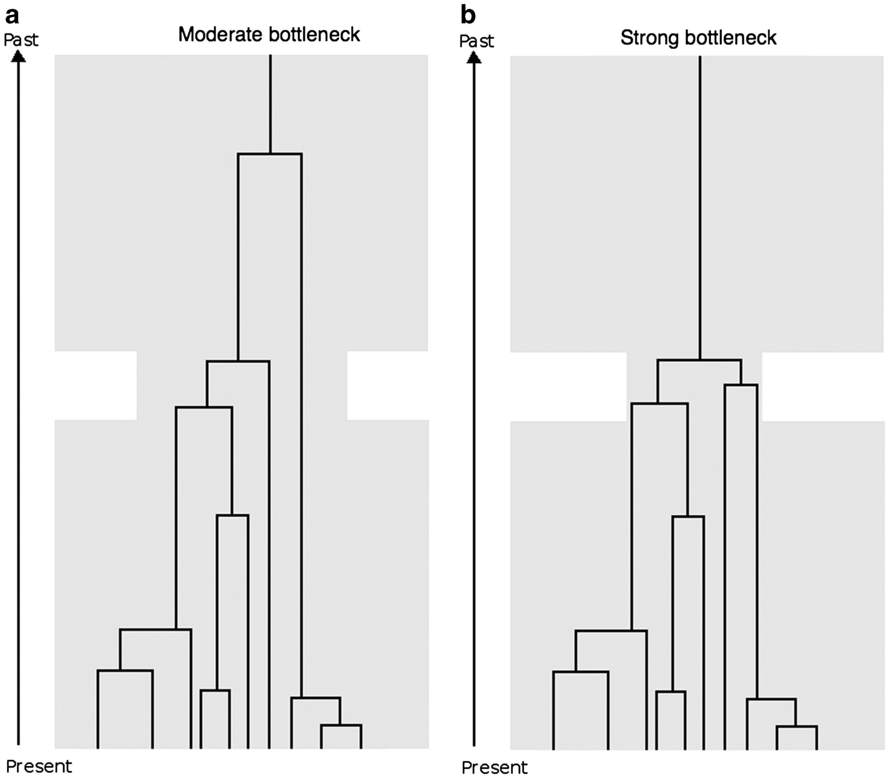

Figure 1 Genealogies of a sample under a bottleneck model. (**a**) Two lineages escape a moderate bottleneck and coalesce further back in the past. (**b**) All the remaining lineages coalesce during a severe bottleneck. A full color version of this figure is available at the *Heredity* journal online.

by 80% for 1000 generations in a population of 10 000 diploid individuals), the scaled SFS has a 'tick mark' pattern (Figure 2a) with a slight deficit of singletons, a more pronounced deficit of other low-frequency variants and an excess of high frequency variants. The deficit of singletons disappears for sufficiently ancient bottlenecks but the deficit of other low-frequency variants remains long after the end of the bottleneck (Figure 2a). For stronger bottlenecks (reduction in effective population by 95% for 1000 generations), the deficit of singletons is only visible for the most recent bottlenecks (Figure 2b) and an excess of singletons quickly appears as the bottleneck becomes more ancient (negative Tajima's *D*, see Figure 3).

In humans, a paradox emerged when investigating the pattern of polymorphism of non-African populations: the average Tajima's *D* of autosomal sequences was found to be positive whereas the mitochondrial Tajimas's *D* values were negative and even smaller than −1 (Garrigan and Hammer, 2006). The paradox disappears if we consider the findings of Fay and Wu (1999) who showed that the out-of-Africa bottleneck affected the mitochondrial and the autosomal values of the Tajima's *D* differently and the reason is as follows: for a strong enough bottleneck, no lineage will escape the bottleneck so that the resulting star-like gene tree (Figure 1b) creates an excess of rare variants and negative values of Tajima's *D* (Figure 3b). In a more moderate bottleneck, some lineages will escape the bottleneck and the basal coalescent lineages will constitute an unexpectedly large part of the coalescent tree (Figure 1a), and generate positive Tajima's *D* values (Figure 3a). The discrepancy can be explained by the fact that effective population size of autosomes is four times that of the mitochondrion. The out-of-Africa bottleneck can therefore be considered strong for the mitochondrion and moderate for the autosomes. Additionally, the out of-Africa bottleneck is more ancient if scaled in units of the

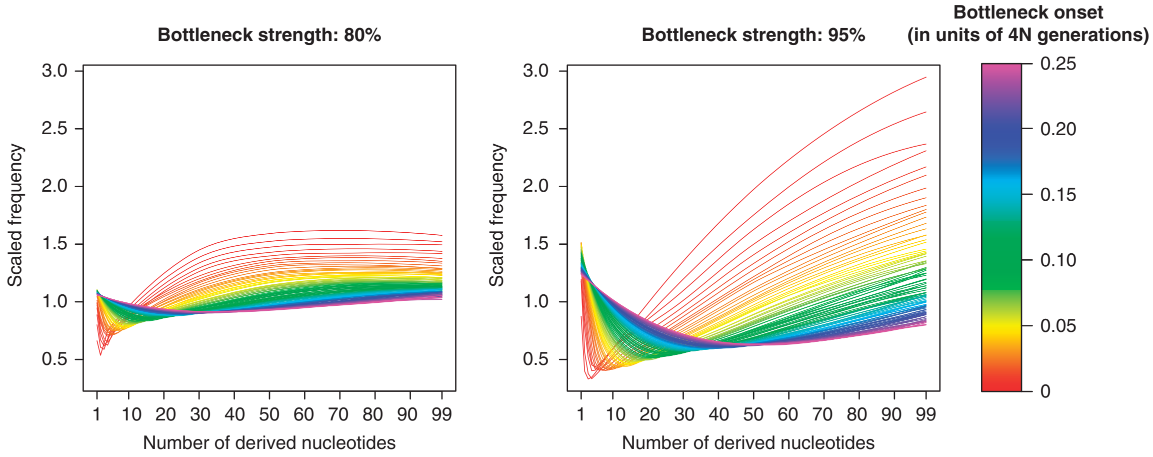

Figure 2 Scaled SFS under a bottleneck model. The scaled SFS is the ratio of the SFS and the expected SFS for a model of constant population size. A sample of 100 chromosome of 50 000 bp long was simulated for a population of 10 000 present-day diploid individuals. (**a**) Recombination rate of 1.5×10^{-8} per site and per generation as well as a mutation rate of 1.2×10^{-8} per site and per generation was used (Scally and Durbin, 2012). Each curve corresponds to the ratio between the observed SFS under a bottleneck model and the expected SFS under a standard neutral model with constant population size. The color of the curves indicates the onset of the bottleneck, in units of four times the present-day population size. Each bottleneck lasted 1000 generations and the strength of the bottleneck measures the reduction in population size during the bottleneck: 80% reduction (left panel) or 95% reduction (right panel). A total of 20 000 simulations per model were performed with the software ms (Hudson, 2002).

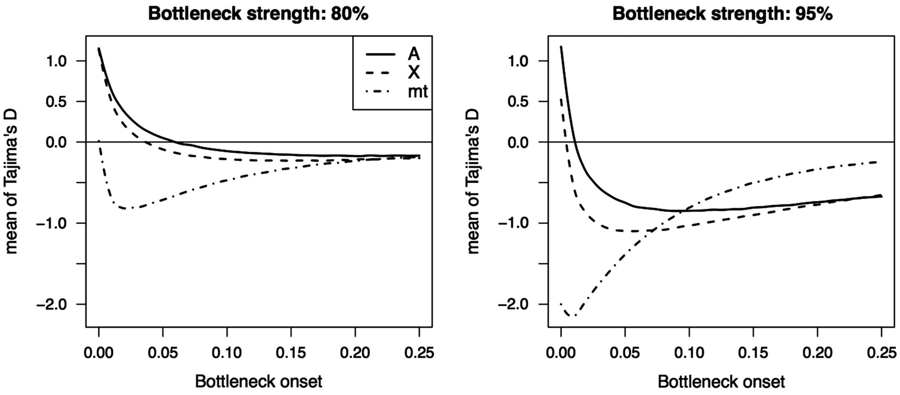

Figure 3 Mean of Tajima's *D* as a function of the time since the onset of the bottleneck. Average values of Tajima's *D* over 20 000 replicate simulations of bottleneck models, for an 80% reduction in population size (left panel) and a 95% reduction in population size (right panel). Different chromosomes are considered: autosome (A, solid line), X chromosome (X, dashed line) and mitochondrion (mt, dash-dotted line). The bottleneck lasted 1000 generations for all models, and the onset of the bottleneck varies, as indicated on the x axis (given in units of four times the present-day population size). The simulated samples are composed of 100 segments of 50 000 bp (the recombination rate is 1.5×10^{-8} per site and per generation), with a mutation rate of 1.2×10^{-8} per site and per generation, and they are randomly sampled from a present-day population of 10 000 diploid individuals. mtDNA segments are also 50 000 bp long, but they are non-recombining and the mutation rate is 2.5×10^{-6}. Time is measured in units of 4 N generations.

mitochondrial effective population size compared with scaling time in units of the autosomal effective population size. Tajima's D is only positive for the short period of time that follows the bottleneck, and it becomes negative with the appearance of new mutations (Figure 3). Note also that bottlenecks impact autosomes and the X-chromosome differently, which is caused by the different effective population sizes (see Figure 3).

With sequence data from multiple loci, additional information arises from the variance of the per-locus Tajima's D values. Technically speaking, the variance of Tajima's D is not an SFS summary statistic because it measures the among-loci variation of the SFS and conveys information that is not contained in the SFS. This being said, moderate bottlenecks lead to large variances in the genealogical history of samples from different loci along a chromosome (McVean, 2002). Unexpected large values of the variance of the Tajima's D have been used as evidence for the human out-of-Africa bottleneck (Voight *et al*., 2005) as well as supporting a bottleneck affecting a non-African *Drosophila* population (Haddrill *et al*., 2005). By contrast, the observation of small among-locus variation of Tajima's D for African populations has shown evidence against a human 'speciation bottleneck' that would have occurred during the penultimate ice age 190–130 kya (Sjödin *et al*., 2012). Contrasting the variance of Tajima's D for different genetic compartments also conveys relevant information with regard to past population sizes. The values of Tajima's D might be difficult to compare between different loci because of potential differences in the sample size and in the number of segregating sites, and some authors prefer to scale the Tajima's D with its minimum possible value $D_{\min}$ (Schaeffer, 2002). In an African population of *Drosophila melanogaster*, the variance of $D/D_{\min}$, where $D_{\min}$ is the theoretical minimum value of Tajima's D, was found to be reduced for the X chromosome (compared with the autosomes and accounting for the difference in population size) and this finding indicated a population expansion for an African *Drosophila* population (Hutter *et al*., 2007).

Combining class I summary statistics: Fisher's method and Approximate Bayesian Computation

While Tajima's D measures departures from neutrality that are reflected in the difference between low and intermediate-frequency alleles, there are many other SFS-based summary statistics commonly used to convey information about past demography (Ramos-Onsins and Rozas, 2002). A typical difficulty arises when considering several different summary statistics that should be combined into a single coherent statistical framework.

Fisher's method combines a number (s) of summary statistics and computes the one-dimensional summary statistic

$$C = -2\sum_{i=1}^{s}\log(P_{\text{val}}(i)) \tag{1}$$

where $P_{\text{val}}(i)$ is the estimated P-value of the i-th summary statistic (Voight *et al*., 2005). Using C as a test statistic, a P-value for each demographic model is then reported using simulations. A second approach consists of minimizing over the range of demographic parameters a goodness-of-fit criterion that measures the discrepancy between simulated and observed summary statistics (Schaffner *et al*., 2005). A common drawback of these approaches is that they might favor more complex models because they do not always account for the different number of parameters in each demographic model. Because a parametric bottleneck model usually encompasses a standard constant-size model, it is expected that it would provide a better fit to the data and the bottleneck model should be penalized for its larger complexity. Approximate Bayesian Computation provides a coherent framework to compare alternative demographic models (Pritchard *et al*., 1999). It is based on an acceptance-rejection algorithm where the simulations of each model that provide the best fit to the data are accepted. Then, the relative proportions of the accepted simulations under each model provide the posterior model probabilities. More advanced statistical approaches to evaluate the posterior model probabilities are now frequently used to compare alternative demographic models (Fagundes *et al*., 2007; Sjödin *et al*., 2012). In the following, we will see that comparing the statistical support of alternative demographic models is also possible when considering the full SFS without summarizing it (Gutenkunst *et al*., 2009).

Considering the entire SFS

We denote the SFS by $m=(m_1,\ldots,m_{n-1})$ where m_k denotes the number of variants/alleles that are carried by k individuals. Conditional on the number of segregating sites and assuming independence between sites, the vector m follows a multinomial distribution that provides the likelihood function of the vector of demographic parameters D

$$P(m;D) = \prod\nolimits_{k-1}^{n-1} p_k^{m_k}(D) \tag{2}$$

where p_k is the probability that a variant is carried by k individuals conditional on being segregating in the sample. Because we assume that mutations occurred according to a Poisson process, the p_k's are given by the expected ratio between the branches leading to k individuals and the total length of the coalescent tree, and this expected ratio can be evaluated with simulations (Adams and Hudson, 2004). This likelihood framework is very flexible because it can account any demographic model for which a coalescent approximation is available. In some sufficiently simple situations, closed-form formulae exist for p_k (Polanski and Kimmel, 2003) making the model-fitting step using a multidimensional grid much more rapid (Marth *et al*., 2004). Model checking can be conducted easily in this setting because chi-square statistics are naturally constructed for multinomial models (Adams and Hudson, 2004). Similarly to the posterior probabilities obtained with Approximate Bayesian Computation, the statistical support of alternative models can be compared with likelihood-ratio tests keeping in mind that standard chi-square asymptotic results might fail because the likelihood formula is only a *composite likelihood* in the likely presence of LD between SNPs (Gutenkunst *et al*., 2009). Instead of estimating the probabilities p_k with coalescent simulations, the p_k's can also be estimated with diffusion equations using

$$p_k = \int\limits_{x=0}^{1} \binom{n}{k} x^k(1-x)^{n-k} f(x)dx \tag{3}$$

where $f(x)$ is the solution of a diffusion equation providing the probability that the derived variant is at frequency x in the population. Likelihood formulae similar to Equation (2) are known as the Poisson random field model (Sawyer and Hartl, 1992), which has been thoroughly used to calibrate models of demography for humans (Keinan *et al*., 2007; Gutenkunst *et al*., 2009) and other species for which SNP data are available (Caicedo *et al*., 2007). In addition, it can be generalized to models of divergence and isolation between populations by considering the joint or multiple (across populations) SFS (Gutenkunst *et al*., 2009). It is also possible to account for genotyping error in the probabilistic model that provides the theoretical SFS (Gravel *et al.*, 2011), which can be useful for new sequence data that may contain many genotyping errors. An additional advantage of using the full SFS is purely graphical because we can readily observe why a demographic model is inadequate: because of the out-of-Africa bottleneck, the non-African SFS is clearly lacking low frequency polymorphisms for instance (Keinan *et al.*, 2007). Despite its success, the likelihood framework is nonetheless not a compulsory way for fitting an SFS. Using simulations to minimize a measure of discrepancy between observed and theoretical SFS is also a fruitful approach (François *et al.*, 2008).

LD AND HAPLOTYPIC SUMMARY STATISTICS

Class II summary statistics

Haplotypic patterns and LD are affected by demographic history (Pritchard and Przeworski, 2001). Many haplotypic summary statistics—sometimes called class II statistics by opposition to the class I SFS statistics—can be used to reconstruct population history (Ramirez-Soriano *et al.*, 2008) and we provide an admittedly partial overview of five different approaches here. The first approach is based on the idea that LD decays more slowly after the occurrence of a bottleneck. For instance, the LD of European populations typically extends to 60 kb for common alleles whereas the LD of African populations extends markedly less far and was taken as evidence for a bottleneck experienced by the European population (Reich *et al.*, 2001). The second approach considers that simple haplotypic quantities such as the number of haplotypes, the frequency of the most common frequent haplotypes or the haplotype heterozygosity are informative about past demography (Depaulis *et al.*, 2003). Lohmueller *et al.* (2009) used the bivariate histogram of the number of haplotypes and the frequency of the most common frequent haplotypes for windows of fixed genetic length and they fitted a demographic model with an approximate-likelihood method. For a given number of haplotypes, the signature of a bottleneck would be an excess of windows where the most common haplotype is at unusually high frequency (Lohmueller *et al.*, 2009). This signal of a bottleneck was found for a population of European descent but not for an admixed population of African origin (Lohmueller *et al.*, 2009).

Identical by descent (IBD) segments and runs of homozygosity

The third approach capitalizes on that homologous long segments that are IBD from a common ancestor provide clues about past population sizes. Compared with identical by state segments, IBD segments are at the same time more informative about remote relatedness (Gusev *et al.*, 2012) but also more difficult to detect so that dedicated algorithms have been devised for this purpose (Gusev *et al.*, 2009). Between pairs of individuals, the length of the IBD segments decay exponentially with the size of the segments (in centimorgans) and the rate of decay depends on the number of meioses d on the path relating the two individuals (Huff *et al.*, 2011). The histogram of the sizes of the long IBD segments is typically used to estimate the degree of relatedness between individuals (Huff *et al.*, 2011), but it is additionally informative about recent founder events (or bottlenecks). Using the fact that the distribution of the size L (in centimorgans) of IBD segments between a pair of individuals separated by d meioses is an exponential distribution of rate $d/100$ (Huff *et al.*, 2011), we can integrate over the number of meioses and show that in a standard coalescent model with a diploid constant population size N, the probability distribution function of the length of an IBD segment is given by

$$p(L) = \frac{2}{100\times2\times N}\frac{1}{(2L/100+1/(2N))^2} \tag{4}$$

If a bottleneck occurred G generations ago followed by a strong expansion that prevented lineages to coalesce during the expansion phase, then the number of meioses can be approximated by twice the sum of G and an exponential random variable of rate $1/(2N)$ where N is the number of diploid individuals before the bottleneck. The distribution of the tract length is then given by

$$p_{\mathrm{bott}}(L) = \frac{2}{10}e^{-2GL/100}\left(\frac{1}{2N(2L/100+1/(2N))^2}+\frac{G}{4NL/100+1}\right) \tag{5}$$

(see Chapman and Thompson 2003) for more accurate formulae for varying population sizes). Figure 4 shows that when a bottleneck occurred, there is a linear relationship between the frequency of IBD tracts (on a log scale) and the length of IBD tracts, and this linear relationship fails for a constant-size population (Gusev *et al.*, 2012). This approach provided evidence for a bottleneck in the Ashkenazi Jewish population that occurred some 23 generations ago and was followed by a rapid expansion (Atzmon *et al.*, 2010, Gusev *et al.*, 2012). Although valid for sufficiently recent bottlenecks, investigating the distribution of IBD tract lengths is not adequate for older bottlenecks such as the out-of-Africa event because of the difficulty, at least with the 500k SNP chips, of finding IBD segments shorter than 3 cM (Gusev *et al.*, 2012).

The fourth method interrogates two homologous chromosomes within the same individual and computes the runs of homozygosity (ROHs) for different length classes (Kirin *et al.*, 2010; Howrigan *et al.*, 2011). Long ROHs provide information on recent ancestry such as recent effective population size but also on individual inbreeding occurring because of marriage between relatives (Kirin *et al.*, 2010). For instance, the Oceanian populations have a very large numbers of shorter ROHs (1–2 Mb), but few long ROHs (>16 Mb), and it was

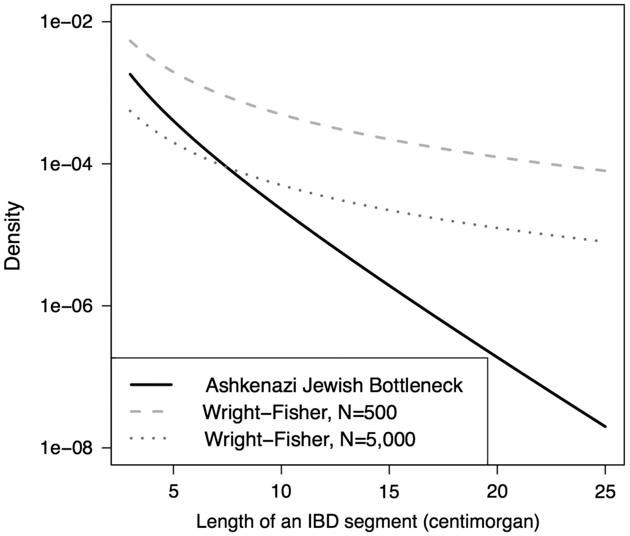

Figure 4 Probability density function of the length of IBD segments. A model of a bottleneck that occurred 20 generations ago with $N=1000$ founding diploid individuals, mimicking the Ashkenazy Jewish bottleneck (Gusev *et al.*, 2012). The model assumes a constant size with $N=1000$ diploid individuals followed by a strong expansion 20 generations ago that prevented coalescence events. The two other models correspond to constant-size populations with 500 or 1000 individuals.

interpreted as the result of little inbreeding in recent times (large recent effective population size) but reduced effective population size in the past (Kirin *et al.*, 2010). Unlike class II haplotypic summary statistics, ROH is individual-based measures and provides a record of the demographic history of an individual's ancestors (Kirin *et al.*, 2010). The outlier individuals who show significantly higher levels of autozygosity are likely offspring of consanguineous unions (Auton *et al.*, 2009). However, a large number of genomic regions that are homozygous across a high proportion of individuals, belonging to a given population, rather provide an indication of a recent founding bottleneck, which was suggested for East Asian and Mexican populations (Auton *et al.*, 2009). Demographic history additionally affects the fraction of the genome that is in ROHs (fROHs) (Henn *et al.*, 2011). The large fROH of 8% found in the Hadza population of Tanzania, which contrasts with the standard 1–3% found for the populations of the Human Genome Diversity Panel, was suggestive of a recent and severe bottleneck in that population (Henn *et al.*, 2011).

LD at different genetic distances

The fifth and last method based on LD uses analytic relationships between measures of LD and the effective population size to reconstruct ancient demographic history. For biallelic loci, denoted as locus A and locus B, a common statistic used to compute LD is the correlation between alleles at pairs of loci that can be expressed as

$$r = \frac{D_{AB}}{\sqrt{p_A(1-p_A)p_B(1=p_B)}} \tag{6}$$

where p_A and p_B are the allele frequencies of one (say the most frequent) of the alleles at the first and second locus and $D_{AB}=p_{AB}-p_Ap_B$ where p_{AB} is the proportion of chromosomes carrying the most frequent alleles at both loci. Providing an analytic formula for the expectation of r or r^2 has proven difficult (McVean, 2002); however, Ohta and Kimura (1971) showed that $\sigma_d{}^2$, the 'standard linkage deviation', which is the ratio of the expectation of $D^2{}_{AB}$ and the expectation of the product of allele frequencies (the denominator of Equation (6)), is a function of the scaled recombination rate $\rho=4Nc$ and c denotes the recombination distance in morgans (McVean, 2002):

$$\sigma_d^2 = \frac{10+\rho}{22+13\rho+\rho^2} \tag{7}$$

Approximating the right-hand side of Equation (7) and considering that the expectation of r^2 can be approximated by $\sigma_d{}^2$, the ratio of expected values, McEvoy *et al.* (2011) estimated N using the approximation $E[r^2]=1/(2+\rho)$. However, the approximation of the expectation of r^2 by the ratio $\sigma_d{}^2$ of the expectations (Equation (7)) is rather poor for small genetic distances between markers (Figure 5; Maruyama, 1982). Moreover, under a scenario of recurrent bottlenecks, analytic approximations for $\sigma_d{}^2$ showed that this measure failed to capture the long-range LD that was generated by the successive bottlenecks (Schaper *et al.*, 2012).

Hayes *et al.* (2003) argued that LD over different genetic distances could be used to reconstruct the change in effective population size over time if it has evolved at a constant linear rate. Their main idea, first introduced by Hill (1981), is that LD over long genetic distances should reflect the recent effective population size, whereas LD over short genetic distances should reflect effective population size further back in time. In particular, Hayes *et al.* (2003) and McEvoy *et al.* (2011) argued that r^2 (or a related measure) between markers at a genetic distance of c can be used to infer the effective population size $1/(2c)$ generations ago. To infer the effective population size using this approximation, an accurate genetic map is needed. Additionally, the method rests on the approximation $E[r^2]=1/(2+\rho)$, which can be rather crude (see Figure 5). Nevertheless, this LD-based approach has been used to re-confirm a substantial reduction of the effective population size accompanying the out-of-Africa event as well as a dramatic re-expansion of the non-African populations (McEvoy *et al.*, 2011).

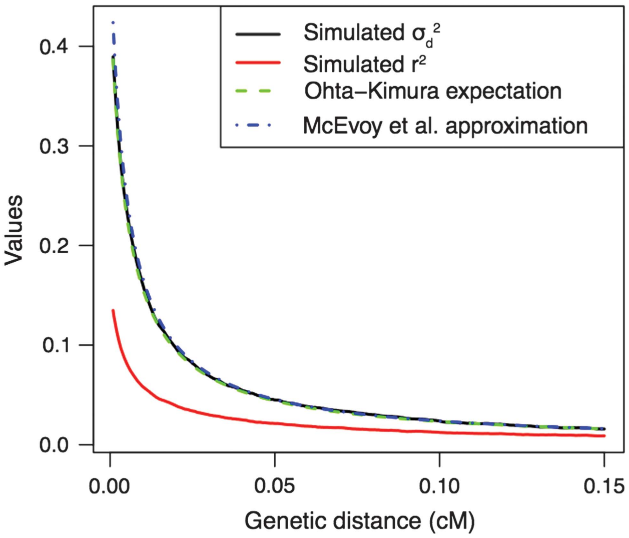

Figure 5 LD measures and analytic approximations. A sample of 100 chromosomes (of length 500 kb) was simulated for a constant population size model of 10 000 diploid individuals assuming a recombination rate of 1.5×10^{-8} per site and per generation, and a mutation rate of 1.2×10^{-8} per site and per generation. Only pairs of segregating sites with genetic distance <0.25 cM were considered. At least 10 million simulated pairs of segregating sites were grouped into 500 equally spaced bins according to their genetic distance between 0 and 0.25 cM. The product of allele frequencies, D^2 values and r^2 values were averaged across all pairs of segregating sites in each bin. The Ohta and Kimura (1971) expectation was computed following Equation (7). The McEvoy *et al.* (2011) approximation corresponds to $1/(2+\rho)$.

FULL SEQUENCE METHODS

Skyline plots

For all methods discussed so far, the demographic history has to be specified with a parametric model. A simple bottleneck model can for instance involve four parameters: the ancestral effective population size equal to the present-day population size, the time at the onset of the bottleneck, the duration of the bottleneck and the effective population size during the bottleneck. Because population history is much more complex than those described by parametric models, it motivated the development of semi- and non-parametric methods of demographic reconstruction (see Ho and Shapiro, 2011 for a review). Pybus *et al.* (2000) introduced the 'skyline plot' framework that provides a very flexible estimation framework of population history. Assuming that the genealogy of the sequence data is known, the coalescence time t_i during which there are i lineages in the coalescent tree is exponentially distributed with a mean of $(2\times2N_i)/i(i-1)$ generations where N_i is the diploid effective population size during this time period. The skyline plot is therefore given by the trajectory of the estimated effective population $\tilde{N}_i=i(i-1)t_i/4$ during the time when there are i lineages in the coalescent tree. Because the obtained trajectories were rugged, more sophisticated methods have been implemented in the BEAST software and in the ape R package

(Ho and Shapiro, 2011). These improvements include the integration over different possible gene genealogies to account for multi-loci data, and the smoothing of the trajectory of the effective population size (Ho and Shapiro, 2011).

Scaling to whole-genome sequences

With the development of new sequencing technologies, future methods of demographic inference would have to scale with the dimension of full genomes. Skyline-plot methodologies typically require the reconstruction of the genealogy of each non-recombining segment (Heled and Drummond, 2008), and it might become computationally prohibitive as the number of segments grows. Recent progress in reconstructing demographic history with large-scale whole-genome sequence data was achieved by Li and Durbin (2011). With a single diploid genome sequence, the principle is to reconstruct the TMRCAs (Times since the Most Recent Common Ancestor) along the genome. Using a hidden Markov model where the TMRCA values are the hidden states, the TMRCAs are reconstructed based on the local density of heterozygotes. Because the transition probability of the hidden Markov chain (conditional distribution of TMRCAs) depends on the temporal variation of the effective population size, estimates of the change of the effective population size over time are provided when fitting the hidden Markov model. Although this approach confirmed the signature of a bottleneck in non-African populations, it also revealed an unforeseen bottleneck in African populations ~50 kya that requires additional confirmation (Li and Durbin, 2011).

CAVEATS

The sensitivity of the SFS-summary statistics to demographic and genetic assumptions has been thoroughly investigated (Thornton, 2005; Ramírez-Soriano *et al.*, 2008; Städler *et al.*, 2009). One of the most prominent confounding effect is population structure that is known to have a profound effect on the SFS (De and Durrett, 2007; Städler *et al.*, 2009). Most other recent statistical methodologies for inferring past population sizes—ROH, IBD tracts, Skyline plots, LD at different genetic distances, and whole-genome sequence methods—are also likely to be affected by many confounding factors including population structure, but their impact on these methods have not yet been thoroughly investigated.

Confounding and biasing factors: genetic factors

Confounding factors can lead to false positive results: a bottleneck can be incorrectly inferred because population structure was not accounted for. Biasing factors are slightly different because they bias the estimates of the demographic parameters, such as the time at which a bottleneck starts. Confounding and biasing factors consist of at least two types: genetic factors that arise from a poor specification of the molecular properties of genetic change over time, and demographic factors, which involve particular events (for example, migration or stratification) that are not accounted for. A straightforward biasing genetic factor concerns the mutation rate. For example, for humans where high-quality genetic data are abundant, estimated mutation rates differ by a factor of 2–3: the rate estimated using mother-father-child trios (or quartet) (The 1000 Genomes Project Consortium, 2010) is 2–3 times lower (10^{-8}/bp/generation) compared with estimates from human-chimpanzee comparisons (2.5×10^{-8}/bp/generation; Nachman and Crowell, 2000) but a consensus is emerging for a mutation rate around 1.2×10^{-8}/bp/generation (Scally and Durbin, 2012). For most other organisms, the uncertainty about mutation rates is much greater. Another important genetic effect arises from a poor specification of the recombination process. For instance, the mean of Tajima's *D* depends on the value of the recombination rate and incorrect assumptions about the recombination rate can lead to substantial biases in estimates of bottleneck parameters (Thornton, 2005). Class II statistics are even more sensitive to the recombination process (Ramírez-Soriano *et al.*, 2008). Lohmueller *et al.* (2009) acknowledged that one of the major disadvantages of their haplotypic summary statistic method to infer bottleneck parameters is its dependence on accurately modeling the distribution of recombination rates across the genome. They suggested that their method should only be used in combination with an accurate genetic map. Ramírez-Soriano *et al.* (2008) also warned that class II summary statistics should not be used if recombination levels are unknown because tests of demographic expansion based on class II summary statistics were too liberal for poorly specified recombination processes. This sensitivity is unfortunate because class II summary statistics based on haplotypic summaries are generally more powerful than class I SFS summary statistics to detect demographic expansions (Ramírez-Soriano *et al.*, 2008) and bottlenecks (Depaulis *et al.*, 2003).

Although class I summary statistics are more robust with respect to the recombination process, reconstructing an unbiased SFS involves several difficulties. First, there is the recurring ascertainment bias with SNP data that is difficult to account for except if the ascertainment strategy is known precisely (Nielsen *et al.*, 2004). With the new sequencing technologies, the ascertainment bias will no longer be a problem but reconstructing the SFS is still challenging; for example, the conjugate effect of sequencing errors and random sampling of alleles at heterozygous sites affect the SFS, in particular for low- to medium-coverage data. To address these issues, Keightley and Halligan (2011) developed a maximum likelihood approach to reconstruct the SFS from error-prone medium-coverage sequence data.

Confounding and biasing factors: population structure and sampling

The second type of problematic factors is the demographic processes that are not accounted for when fitting simple demographic models to empirical data, such as panmictic models of population size changes. Population structure is a notable confounding factor for demographic expansion: pooling samples from different populations skews the SFS toward low-frequency polymorphism (for example, negative Tajima's *D*) mimicking the pattern generated by population expansion (Ptak and Przeworski, 2002). Figure 6 illustrates this 'pooling effect' (Städler *et al.*, 2009); under a scenario of recent range fragmentation, the mitochondrial Tajima's *D* decreases as a function of the number of sub-populations in the fragmented habitat. More ancient range fragmentation will additionally affect the autosomal values of Tajima's *D* (Sjödin *et al.*, 2012). Adding to the confusion, population structure can also mimic the SFS of a bottleneck (Nielsen and Beaumont, 2009; Chikhi *et al.*, 2010). Both a deficit of singletons and an excess of high-frequency polymorphisms (positive Tajima's *D*) can be generated under island and stepping-stone models (De and Durrett, 2007) when the sampling is performed locally (no pooling, see Figure 6). More generally, the sampling scheme affects the SFS: the SFS of pooled samples (multiple individuals from multiple demes) being intermediate between the SFS of local samples (multiple individuals from one deme) and the SFS of scattered samples (one individual per deme) (Städler *et al.*, 2009).

Besides the SFS, the confounding effects of population structure have also been investigated for other summaries of polymorphism

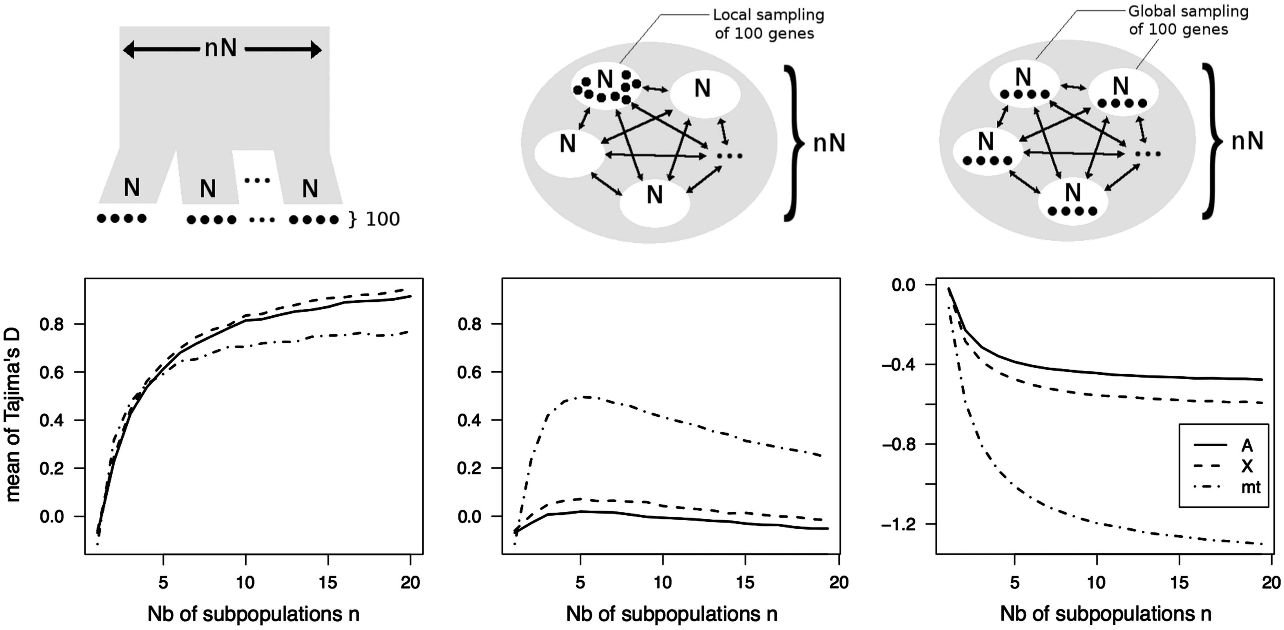

Figure 6 Effect of different models of population structure and different sampling strategies on Tajima's *D*. Average values of Tajima's *D* over 100 000 replicate simulations of divergence models with global sampling (left panel), island models with local sampling (central panel) and island models with global sampling (right panel), as function of the number of demes. The simulated samples are composed of 100 chromosomes (of length 50 000 bp), with a mutation rate of 1.2×10^{-8} per site and per generation and a recombination rate of 1.5×10^{-8} per site and per generation. In all models, the demes are of size 10 000 present-day diploid individuals, and *n* represents the number of demes. In the local sampling scheme, all chromosomes are sampled from one deme only. In the global sampling schemes, however, the chromosomes are sampled from all the demes, with an almost equal number of chromosomes within each deme (plus or minus one chromosome). No migration is allowed in the divergence model, whereas a migration rate of 1×10^{-3} migrants/generation is assumed in the island model.

data. Under the stepping-stone model and the island model, De and Durrett (2007) also showed that LD decays more slowly with genetic distance than expected for a panmictic population mimicking the LD decay of a bottleneck. Another potential biasing factor for haplotypic summary statistics is recent admixture. It biases the estimates of demographic parameters based on class II statistics whereas the SFS summary statistics are more robust (Lohmueller *et al.*, 2010). For example, Lohmueller *et al.* (2010) managed to recover the demographic history of West African populations with SFS summaries of SNPs from African-American individuals whereas the haplotypic statistics did not provide consistent estimates of past African population sizes.

Providing a demographic reconstruction based on the SFS or based on haplotypic measures has consequently numerous caveats that should be addressed. Fortunately, some of these caveats are known and their impact on the inference can be tested. For the ROH, length of IBD tracts, methods based on LD at different distances, Skyline plots and whole-genome sequence methods, we currently lack extensive investigations about the potential confounding and biasing factors, partly because some of these methods have recently been introduced.

Interpreting a signal for a bottleneck

Having controlled for all of these potential caveats, the genetic signature for a bottleneck may be found to be robust. The remaining question is then to provide an explanation or interpretation of the bottleneck. *Sensu stricto*, a bottleneck is the situation in which a population goes through a severe but temporary population restriction (Maruyama and Fuerst, 1984). A naïve interpretation of the bottleneck signature found in the human European population would then be that this population experienced a bottleneck of ~60 kya within Europe, although we know that this bottleneck should rather be interpreted in a spatial context under a scenario of serial founder effects connected with the out-of-Africa migration (DeGiorgio *et al.*, 2011). The bottleneck signal in non-Africans is caused by the relatively small number of individuals migrating out of Africa some 60 kya carrying a subsample of the much greater genetic diversity present among African populations. Quoting the psychologist William James 'we must be careful not to confuse data with the abstractions we use to analyze them'. This problem of interpretation can be particularly acute for non-model organisms for which we may lack knowledge about past demography.

Another challenge is to interpret the fluctuations of the effective population size. The fluctuation of census size is not the only factor affecting the effective population size. For instance, the very long ROHs found in the South and West Asiatic populations were interpreted as the result of inbreeding occurring because of consanguineous marriages, that is, non-random mating (Kirin *et al.*, 2010). Similar patterns were rather interpreted as evidence for founding bottlenecks in the East Asian and Mexican populations (Auton *et al.*, 2009). Distinguishing between the two potential explanations (which are not mutually exclusive) might be difficult because the two scenarios involve a reduction in the effective population size.

PERSPECTIVES

There is a large spectrum of methods to infer historical changes in population size. An emerging class of techniques, which was not addressed in this review, concerns the use of ancient DNA for demographic reconstruction. The ability to infer population histories

Table 1 Statistical measures used to detect bottlenecks

Data	*Species*	*Statistics*	*Simulations*	*Bottleneck*	*Ref.*
SNPs	*Homo sapiens*	LD decay	Coal.-1 pop	Non-African bott.	Reich *et al.* (2001)
SNPs	*Homo sapiens*	SFS-LD-Fst	Coal.-IM model	Out-of-Africa	Schaffner *et al.* (2005)
SNPs	*Homo sapiens*	SFS	Coal. 1 pop	Out-of-Africa	Adams and Hudson (2004), Marth *et al.* (2004) and Keinan *et al.* (2007)
50 10-kb sequences	*Homo sapiens*	Class I-LD decay	Coal. 1 pop	Out-of-Africa	Voight *et al.* (2005)
50 500-bp sequences	*Homo sapiens*	Class I-Fst	Coal.-IM model	Out-of-Africa + Toward America	Fagundes *et al.* (2007)
24 1-kb sequences	*Homo sapiens*	Class I-Fst-Stats based on Multipop-SFS	Coal.-IM model	Pygmies	Patin *et al.* (2009)
SNPs	*Homo sapiens*	Class II	Coal. 1 pop	European bott.	Lohmueller *et al.* (2009)
SNPs	*Homo sapiens*	Joint-SFS	Diffusion IM model	Out-of-Africa	Gutenkunst *et al.* (2009); Gravel *et al.* (2011)
SNPs	*Homo sapiens*	r^2 measure of LD	No simulations	Out-of-Africa	McEvoy *et al.* (2011)
SNPs	*Homo sapiens*	IBD segments	Coal. 1 pop	Ashkenazy Jewish	Atzmon *et al.* (2010) and Gusev *et al.* (2012)
SNPs	*Homo sapiens*	ROH	Coal. 1 pop	Hadza	Henn *et al.* (2011)
Whole genome	*Homo sapiens*	Local density of heterozygotes	Coal. 1 pop	Bott. in all continents	Li and Durbin (2011)
61 20-kb sequences	*Homo sapiens*	Class I	Coal. 1 pop	No speciation bott. during ice age	Sjödin *et al.* (2012)
X-linked and autosomal sequences	D*rosophila melanogaster*	Class I and II-LD	Coal. 1 pop	Non-African bott.	Haddrill *et al.* (2005) and Thornton and Andolfatto (2006)
X-linked and autosomal sequences	*Drosophila melanogaster*	SFS	Coal.-2pop IM model	European bott.	Hutter *et al.* (2007)
~300 500-kb sequences	*Drosophila melanogaster*	Class I and II	Coal.-IM model	Southeast Asia	Laurent *et al.* (2011)

Abbreviations: bott., bottleneck; Coal., coalescent; IBD, identical by descent; IM model, Isolation and Migration model; LD, linkage disequilibrium; ROHs, runs of homozygosity; SFS, site-frequency spectrum.

can be enhanced by including ancient DNA because it adds the temporal information to the DNA data (Mourier *et al.*, 2012). Among the different methods reviewed here, the ones based on haplotypes and on shared genomic segments (ROH, IBD tracts) have been shown to provide information about more recent demographic event (see Table 1; Atzmon *et al.*, 2010; Henn *et al.*, 2011; Gusev *et al.*, 2012). Aside from the specific nature of statistical summaries, we believe that contrasting statistical summaries for different genomic compartments (Fay and Wu, 1999, Blum and Jakobsson, 2011) and different sampling strategies (Städler *et al.*, 2009) can be very powerful because it can provide additional information about demography for the study species/population.

In addition to the classical methods that fit statistical summaries with numerical simulations of polymorphism data, more sophisticated methods (Heled and Drummond, 2008; Li and Durbin, 2011) have been developed that add to the current toolbox of population geneticists. These methods considerably facilitate the fitting of demographic processes because the user does not need to specify a parametric model of demographic changes. However, little is known about the effects of confounding and biasing factors on these methods and interpreting the result in terms of demography is often difficult since many factors can affect effective population sizes. In addition, a recommended practice would be to investigate the predictions obtained with the calibrated demographic models. Based on SFS, LD or haplotypic statistical measures for example, simulations can readily be used to check if the fitted demographic model can reproduce the main features of the polymorphism data (see for example, Thornton and Andolfatto, 2006; Sjödin *et al.*, 2012). In the end, proposing the model of population size changes that best fits the data will give us an idea of the demographic scenario that can have generated the patterns of variation and it will help us to advance our understanding of the evolutionary history of the population or species, at least by producing models that future inference methods can be tested against.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Michael Stocks for comments on the manuscript. This work was supported by a grant from the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) awarded to MJ and MB. A grant from the French national research agency provided support to MB (DATGEN project, ANR-2010-JCJC-1607-01) and a grant from the Swedish research council provided support to MJ.

Achaz G (2009). Frequency spectrum neutrality tests: one for all and all for one. *Genetics* **183**: 249–258.

Adams AM, Hudson RR (2004). Maximum-likelihood estimation of demographic parameters using the frequency spectrum of unlinked single-nucleotide polymorphisms. *Genetics* **168**: 1699–1712.

Atzmon G, Hao L, Pe'er I, Velez C, Pearlman A, Palamara PF *et al.* (2010). Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry. *Am J Hum Genet* **86**: 850.

Auton A, Bryc K, Boyko AR, Lohmueller KE, Novembre J, Reynolds A *et al.* (2009). Global distribution of genomic diversity underscores rich complex history of continental human populations. *Genome Res* **19**: 795–803.

Barton NH, Charlesworth B (1984). Genetic revolutions, founder effects, and speciation. *Annu Rev Ecol Evol Syst* **15**: 133–164.

Blum MGB, Jakobsson M (2011). Deep divergences of human gene trees and models of human origins. *Mol Biol Evol* **28**: 889–898.

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL *et al.* (2007). Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genet* **3**: e163.

Chapman NH, Thompson EA (2003). A model for the length of tracts of identity by descent in finite random mating populations. *Theor Popul Biol* **64**: 141–150.

Chikhi L, Sousa V, Luisi P, Goossens B, Beaumont MA (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics* **186**: 983–995.

De A, Durrett R (2007). Stepping-stone spatial structure causes slow decay of linkage disequilibrium and shifts the site frequency spectrum. *Genetics* **176**: 969–981.

DeGiorgio M, Degnan JH, Rosenberg NA (2011). Coalescence-time distributions in a serial founder model of human evolutionary history. *Genetics* **189**: 579–593.

Depaulis F, Mousset S, Veuille M (2003). Power of neutrality tests to detect bottlenecks and hitchhiking. *J Mol Evol* **57**: S190–S200.

England PR, Osler GH, Woodworth LM, Montgomery ME, Briscoe DA, Frankham R (2003). Effects of intense versus diffuse population bottlenecks on microsatellite genetic diversity and evolutionary potential. *Conserv Genet* **4**: 595–604.

Fagundes NJR, Ray N, Beaumont MA, Neuenschwander S, Salzano FM, Bonatto SL *et al.* (2007). Statistical evaluation of alternative models of human evolution. *Proc Natl Acad Sci USA* **104**: 17614–17619.

Fay JC, Wu C-I (1999). A human population bottleneck is not incompatible with the discordance between patterns of mitochondrial vs. nuclear DNA variation. *Mol Biol Evol* **16**: 1003–1006.

François O, Blum MGB, Jakobsson M, Rosenberg NA (2008). Demographic history of European populations of Arabidopsis thaliana. *PLoS Genet* **4**: e1000075.

Garrigan D, Hammer MF (2006). Reconstructing human origins in the genomic era. *Nat Rev Genet* **7**: 669–680.

Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG *et al.* (2011). Demographic history and rare allele sharing among human populations. *Proc Natl Acad Sci USA* **108**: 11983–11988.

Gusev A, Lowe JK, Stoffel M, Daly MJ, Altshuler D, Breslow JL *et al.* (2009). Whole population, genome-wide mapping of hidden relatedness. *Genome Res* **19**: 318–326.

Gusev A, Palamara PF, Aponte G, Zhuang Z, Darvasi A, Gregersen P *et al.* (2012). The architecture of long-range haplotypes shared within and across populations. *Mol Biol Evol* **29**: 473–486.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP data. *PLoS Genet* **5**: e1000695.

Haddrill P, Thornton KR, Charlesworth B, Andolfatto P (2005). Multilocus patterns of nucleotide variability and the demographic and selection history of Drosophila melanogaster populations. *Genome Res* **15**: 790–799.

Hayes BJ, Visscher PM, McPartlan HC, Goddard ME (2003). Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res* **13**: 635–643.

Heled J, Drummond AJ (2008). Bayesian inference of population size history from multiple loci. *BMC Evol Biol* **8**: 1–15.

Henn BM, Gignoux CR, Jobin M, Granka JM, Macpherson JM, Kidd JM *et al.* (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *Proc Natl Acad Sci USA* **108**: 5154–5162.

Hill WG (1981). Estimation of effective population size from data on linkage disequilibrium. *Genet Res* **38**: 209–216.

Ho SYW, Shapiro B (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. *Mol Ecol Resour* **11**: 423–434.

Howrigan DP, Simonson MA, Keller MC (2011). Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics* **12**: 460.

Hudson RR (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* **18**: 337–338.

Huff CD, Witherspoon DJ, Simonson TS, Xing J, Watkins WS, Zhang Y *et al.* (2011). Maximum-likelihood estimation of recent shared ancestry (ERSA). *Genome Res* **21**: 768–774.

Hutter S, Li H, Beisswanger S, De Lorenzo D, Stephan W (2007). Distinctly different sex ratios in African and European populations of Drosophila melanogaster inferred from chromosome-wide SNP data. *Genetics* **177**: 469–480.

Keightley PD, Halligan DL (2011). Inference of site frequency spectra from high-throughput sequence data: quantification of selection on nonsynonymous and synonymous sites in humans. *Genetics* **188**: 931–940.

Keinan A, Mullikin JC, Patterson N, Reich D (2007). Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nat Genet* **39**: 1251–1255.

Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF (2010). Genomic runs of homozygosity record population history and consanguinity. *PLoS ONE* **5**: e13996.

Laurent SJY, Werzner A, Excoffier L, Stephan W (2011). Approximate Bayesian Analysis of *Drosophila melanogaster* polymorphism data reveals a recent colonization of southeast Asia. *Mol Biol Evol* **28**: 2041–2051.

Li H, Durbin R (2011). Inference of human population history from individual whole-genome sequences. *Nature* **475**: 493–497.

Li HP, Stephan W (2006). Inferring the demographic history and rate of adaptive substitution in Drosophila. *PLoS Genet* **2**: 1580–1589.

Lohmueller KE, Bustamante CD, Clark AG (2009). Methods for human demographic inference using haplotype patterns from genome-wide SNP data. *Genetics* **182**: 217–231.

Lohmueller KE, Bustamante CD, Clark AG (2010). The effect of recent admixture on inference of ancient population history. *Genetics* **185**: 611–622.

Lohmueller KE, Indap AR, Schmidt S, Boyko AR, Hernandez RD *et al.* (2008). Proportionally more deleterious genetic variation in European than in African populations. *Nature* **451**: 994–997.

Marth GT, Czabarka E, Murvai J, Sherry ST (2004). The allele frequency spectrum in genome-wide human variation data reveals signals of differential demographic history in three large world populations. *Genetics* **166**: 351–372.

Maruyama T (1982). Stochastic integrals and their application to population genetics. In: Kimura M (ed) *Molecular Evolution, Protein Polymorphism and the Neutral Theory.* Springer-Verlag: Berlin, pp 151–166.

Maruyama T, Fuerst PA (1984). Population bottlenecks and nonequilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. *Genetics* **108**: 745–763.

McEvoy BP, Powell JE, Goddard ME, Visscher PM (2011). Human population dispersal "Out of Africa" estimated from linkage disequilibrium and allele frequencies of SNPs. *Genome Res* **21**: 821–829.

McVean GAT (2002). A genealogical interpretation of linkage disequilibrium. *Genetics* **162**: 987–991.

Mondol S, Karanth KU, Ramakrishnan U (2009). Why the Indian subcontinent holds the key to global tiger recovery. *PLoS Genet* **5**: e1000585.

Mourier T, Ho SYW, Gilbert MTP, Willerslev E, Orlando L (2012). Statistical guidelines for detecting past population shifts using ancient DNA. *Mol Biol Evol* **29**: 2241–2251.

Nachman MW, Crowell SL (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics* **156**: 297–304.

Nei M, Maruyama T, Chakraborty R (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.

Nielsen R (2005). Molecular signatures of natural selection. *Annu Rev Genet* **39**: 197–218.

Nielsen R, Beaumont MA (2009). Statistical inferences in phylogeography. *Mol Ecol* **18**: 1034–1047.

Nielsen R, Hubisz MJ, Clark AG (2004). Reconstituting the frequency spectrum of ascertained single-nucleotide polymorphism data. *Genetics* **168**: 2373–2382.

Ohta T, Kimura M (1971). Linkage disequilibrium between two segregating nucleotide sites under the steady flux of mutations in a finite population. *Genetics* **68**: 571–580.

Patin E, Laval G, Barreiro LB, Salas A, Semino O, Santachiara-Benerecetti S *et al.* (2009). Inferring the demographic history of African farmers and Pygmy hunter-gatherers using a multilocus resequencing data set. *PLoS Genet* **5**: e1000448.

Peter BM, Wegmann D, Excoffier L (2010). Distinguishing between population bottleneck and population subdivision by a Bayesian model choice procedure. *Mol Ecol* **19**: 4648–4660.

Polanski A, Kimmel M (2003). New explicit expressions for relative frequencies of single-nucleotide polymorphisms with application to statistical inference on population growth. *Genetics* **165**: 427–436.

Pritchard JK, Przeworski M (2001). Linkage disequilibrium in humans: models and data. *Am J Hum Genet* **69**: 1–14.

Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW (1999). Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol Biol Evol* **16**: 1791–1798.

Ptak S, Przeworski M (2002). Evidence for population growth in humans is confounded by population structure. *Trends Genet* **18**: 559–563.

Pybus OG, Rambaut A, Harvey PH (2000). An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* **155**: 1429–1437.

Ramirez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A (2008). Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. *Genetics* **179**: 555–567.

Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**: 2092–2100.

Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ *et al.* (2001). Linkage disequilibrium in the human genome. *Nature* **411**: 199–204.

Roux C, Pauwels M, Ruggiero M-V, Charlesworth D, Castric V, Vekemans X (2013). Recent and ancient signature of balancing selection around the S-locus in Arabidopsis halleri and A. lyrata. *Mol Biol Evol* **30**: 435–447.

Ruzzante DE, Walde SJ, Gosse JC, Cussac VE, Habit E, Zemlak TS *et al.* (2008). Climate control on ancestral population dynamics: insight from Patagonian fish phylogeography. *Mol Ecol* **17**: 2234–2244.

Sawyer SA, Hartl DL (1992). Population genetics of polymorphism and divergence. *Genetics* **132**: 1161–1176.

Scally A, Durbin R (2012). Revising the human mutation rate: implications for understanding human evolution. *Nat Rev Genet* **13**: 745–753.

Schaeffer SW (2002). Molecular population genetics of sequence length diversity in the Adh region of *Drosophila pseudoobscura*. *Genet Res* **80**: 163–175.

Schaffner S, Foo C, Gabriel S, Reich D, Daly MJ, Altshuler D (2005). Calibrating a coalescent simulation of human genome sequence variation. *Genome Res* **15**: 1576–1583.

Schaper E, Eriksson A, Rafajlovic M, Sagitov S, Mehlig B (2012). Linkage disequilibrium under recurrent bottlenecks. *Genetics* **190**: 217–229.

Sjödin P, Sjöstrand AE, Jakobsson M, Blum MGB (2012). Resequencing data provide no evidence for a human bottleneck in Africa during the penultimate glacial period. *Mol Biol Evol* **29**: 1851–1860.

Städler T, Haubold B, Merino C, Stephan W, Pfaffelhuber P (2009). The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics* **182**: 205–216.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Tajima F (1997). Estimation of the amount of DNA polymorphism and statistical tests of the neutral mutation hypothesis based on DNA polymorphism. In: Tavaré S, Donnelly P (eds) *'Progress in Population Genetics and Human Evolution'*. Springer: Berlin, pp 149–164.

The 1000 Genomes Project Consortium (2010). A map of human genome variation from population-scale sequencing. *Nature* **467**: 1061–1073.

Thornton K (2005). Recombination and the properties of Tajima's D in the context of approximate likelihood calculation. *Genetics* **171**: 2143–2148.

Thornton KR, Andolfatto P (2006). Approximate Bayesian Inference reveals evidence for a recent, severe, bottleneck in a Netherlands population of Drosophila melanogaster. *Genetics* **172**: 1607–1619.

Voight BF, Adams AM, Frisse LA, Qian YD, Hudson RR, Di Rienzo A (2005). Interrogating multiple aspects of variation in a full resequencing data set to infer human population size changes. *Proc Natl Acad Sci USA* **102**: 18508–18513.

Williamson-Natesan EG (2005). Comparison of methods for detecting bottlenecks from microsatellite loci. *Conserv Genet* **6**: 551–562.

Wright SI, Bi IV, Schroeder SG, Yamasaki M, Doebley JF, McMullen MD *et al*. (2005). The effects of artificial selection on the maize genome. *Science* **308**: 1310–1314.