Human Biology

Volume 81
Issue 2 *Special Issue on Demography and Cultural Macroevolution*

Article 3

2009

Inferring Past Demography Using Spatially Explicit Population Genetic Models

Nicolas Ray
University of Bern, Nicolas.Ray@unige.ch

Laurent Excoffier
University of Bern

Inferring Past Demography Using Spatially Explicit Population Genetic Models

Abstract
Inferring the past demography of human populations has been classically approached through data from the archaeological record but more recently by the use of genetic data from contemporary samples. Building realistic demographic models at the continental scale is a necessary step toward the improvement of current genomic methods aimed at fi nding genes under selection. In light of recent advances in Bayesian statistical inference, we discuss here the importance of considering spatially explicit approaches for modeling population expansion and dispersal. Neutral processes, such as the surfi ng phenomenon, that occur at the front of a range expansion may indeed mimic selection, and they may have played a signifi cant role in spreading particular alleles over large geographic areas. Finally, we discuss a few important issues that require further investigation, notably the use of archaeological data to inform population genetic models, the simulation of range contraction and reexpansion, and the importance of long-distance dispersal.

Keywords
range expansion, human evolution, introgression, surfing phenomena, Bayesian inference, paleodemography, genetic diversity, population genetic models, SPLATCHE software.

Cover Page Footnote
This work was supported by the Swiss National Science Foundation through grant 3100A0-112072 awarded to L. Excoffier. We thank two anonymous reviewers for their comments on an earlier draft.

Inferring Past Demography Using Spatially Explicit Population Genetic Models

NICOLAS RAY[1,2] AND LAURENT EXCOFFIER[1]

Abstract Inferring the past demography of human populations has been classically approached through data from the archaeological record but more recently by the use of genetic data from contemporary samples. Building realistic demographic models at the continental scale is a necessary step toward the improvement of current genomic methods aimed at finding genes under selection. In light of recent advances in Bayesian statistical inference, we discuss here the importance of considering spatially explicit approaches for modeling population expansion and dispersal. Neutral processes, such as the surfing phenomenon, that occur at the front of a range expansion may indeed mimic selection, and they may have played a significant role in spreading particular alleles over large geographic areas. Finally, we discuss a few important issues that require further investigation, notably the use of archaeological data to inform population genetic models, the simulation of range contraction and reexpansion, and the importance of long-distance dispersal.

Importance of Knowing the Past Demography of Human Populations: Most Relevant and Significant Previous Work

The past demographic history of humans is a fascinating topic, but it is also important for detecting genetic loci under positive (adaptation) or negative (diseases) selection. Inferring the paleodemography of populations is therefore becoming a major step in ecological and medical genetics. Data from the archaeological record provide an important source of information for paleodemography [see Chamberlain 2009 (this issue)], but genetic data could become increasingly important and useful for making demographic inferences in the future. Genetic data have indeed entered the field of paleodemography more recently, but they have rapidly become crucial to defining phylogeographic relationships among

[1]Computational and Molecular Population Genetics Lab, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland; and Swiss Institute of Bioinformatics, Switzerland.

[2]Present address: Climatic Change and Climate Impacts/enviroSPACE Group, University of Geneva, Geneva, Switzerland.

Human Biology, April–June 2009, v. 81, nos. 2–3, pp. 141–157.

KEY WORDS: RANGE EXPANSION, HUMAN EVOLUTION, INTROGRESSION, SURFING PHENOMENA, BAYESIAN INFERENCE, PALEODEMOGRAPHY, GENETIC DIVERSITY, POPULATION GENETIC MODELS, SPLATCHE SOFTWARE.

extant populations (see, e.g., Rosenberg et al. 2002). These relationships can inform us about past demography because events such as migrations, population bottlenecks, population expansion, or subdivision may drastically alter current patterns of molecular diversity. The molecular signatures left by these past events can therefore be used to make inferences about these demographic events.

Human curiosity is the main driver behind the quest for knowledge about the history of our species. However, as mentioned, developing a realistic demographic model of human prehistory is also important for biomedical applications that study genes that are or have been under selection. The correct identification of these genes should allow one to better understand the genetic bases of adaptation and speciation (Nielsen et al. 2007; Wu and Ting 2004) and could lead to the identification of genes involved in complex diseases. A common approach for the detection of genes under selection is based on the detection of genes that show extreme allele frequency differences or similarities between populations. An index of population differentiation, such as F_{ST}, can be computed for thousands or millions of SNPs (single nucleotide polymorphisms) (Akey et al. 2002; Shriver et al. 2004) assayed in samples from different regions, and its empirical distribution can be computed. The loci located at the tails of this distribution are then generally considered as being potentially under selection (Lewontin and Krakauer 1973). The rationale for this model-free approach is that gene frequencies at loci involved in adaptive events should show increased levels of differentiation among populations, whereas genes under balancing selection should present relatively uniform allele frequencies across populations and lead to low F_{ST} values (see Barreiro et al. 2008). The main problem with this approach is that the empirical distribution can include a mixture of selected and neutral genes. Moreover, the neutral distribution is actually unknown and its shape should be dependent on the past demography of the populations.

Another approach used to target genes under selection is to try to model past global demography to obtain the null distribution of allele frequency differences from which outliers can be recognized. The problem here is that this past demography has always been considered simple, for instance, by assuming the occurrence of a bottleneck and demographic recovery (Storz et al. 2004; Williamson et al. 2005) or by considering simple models of population structure such as an island model (Akey et al. 2004). More realistic models that take into account past range expansions, the spatial locations of the populations, or their shared common history could be more powerful and decrease the rate of false-positive signals.

Models of Past Human Demography. Earlier human population genetic studies were restricted to simple demographic models, generally involving isolated population(s) with constant population size (e.g., Reich and Goldstein 1998). However, recent studies on patterns of human genetic diversity suggest that human demography has been much more complex, with extensive levels of subdivisions in Africa (Campbell and Tishkoff 2008; Gonder et al. 2007; Harding and McVean 2004), migrations back to Africa (Templeton 2002), corridors of migration (e.g.,

Wang et al. 2007), and potential interbreeding with local populations [Eswaran 2002; Eswaran et al. 2005; but also see Currat and Excoffier (2005), Currat et al. (2008), and our later discussion for a different opinion].

The recent availability of large-scale genetic studies of many populations [e.g., the Centre d'Etude du Polymorphisme Humain/Human Genome Diversity Project (CEPH-HGDP) panel] has helped to gain a global picture of human diversity (Rosenberg et al. 2002). Based on these data, simple descriptive approaches have shown that the distance from East Africa is a good predictor of the genetic diversity in human populations worldwide (Handley et al. 2007; Prugnolle et al. 2005; Ramachandran et al. 2005), making geography a better determinant of human genetic differentiation than ethnicity (Manica et al. 2005). These studies have strengthened the case for the recent African origin (RAO) model of human evolution, as opposed to alternative models, such as the multiregional evolution model (Wolpoff et al. 2000) or some forms of assimilation models in which modern humans migrating outside Africa hybridized with local *Homo erectus* and incorporated old lineages (Eswaran et al. 2005; Evans et al. 2006). Using a Bayesian framework, we have been able to statistically assess the relative probabilities of different flavors of all these models using DNA data from 50 nuclear loci sequenced in African, Asian, and native American samples (Fagundes et al. 2007). The analyses led to the largest posterior probability for an RAO model (80%) and thus possibly to the invalidation of other models. This approach showed the importance of using realistic models of past demography and of considering the interplay between different regions of the world.

Another aspect of past demography that has received much attention recently is the consequence of spatial expansions on genetic diversity (Excoffier 2002). These expansions are typically accompanied by a demographic expansion (i.e., a large increase in the total population size), and they have most certainly occurred in many places and at different times of human history. Prominent examples of such expansions are the exit of modern humans out of Africa, the colonization of Europe by hunter-gatherers, and the peopling of the Americas from Asia. To be able to model these events, our team has developed the software SPLATCHE (Currat et al. 2004), which enables us to model complex demography of populations arranged on a two-dimensional lattice and to subsequently simulate the genealogy of a sample of genes based on the simulated demography. Each subpopulation can have its own independent population size, and these subpopulations can be linked to environmental information, such as the type of land cover, topography, or the river network (Ray et al. 2008). This modeling framework has proved useful to understanding the molecular signatures left by a range expansion under various migration rates and population sizes (Ray et al. 2003) and to investigating the effect of spatially and temporally varying resource abundances on patterns of genetic diversity (Wegmann et al. 2006).

The importance of spatially explicit models of genetic diversity can be best exemplified by the surfing phenomenon (Edmonds et al. 2004; Klopfstein et al. 2006) (see Figure 1). Surfing describes the geographic spread of an allele that

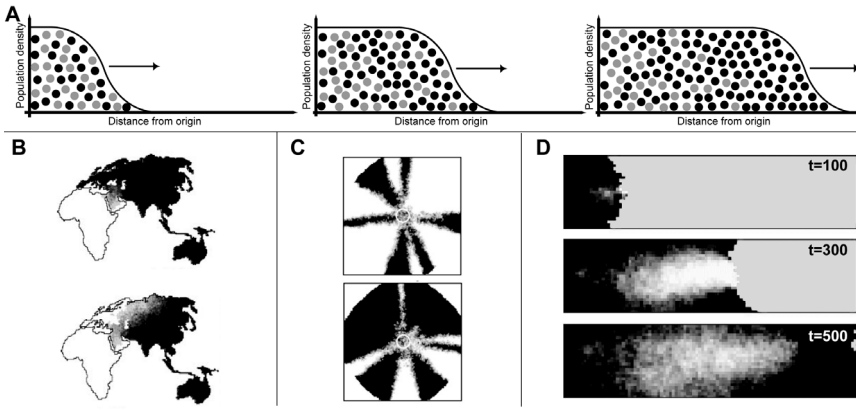

Figure 1. The surfing phenomenon: Examples from simulations. (A) Surfing phenomena showing the consequences of drift occurring at the wave front of a spatially expanding population. Initial conditions show an equal proportion of two alleles (gray and black dots). As the expansion progresses from left to right, the black allele found by chance at the tip of the wave increases in frequency and eventually becomes fixed in downstream populations. (B) Two outputs of the spatial frequency distribution of an allele after the simulation of an expansion out of Africa [see Currat et al. (2006) for simulation details]. Darker shades of gray indicate higher allele frequencies. The rare allele is initially present in Africa at a frequency of 5%. Because of the spatial bottleneck during the exit from Africa, this allele can sometimes increase in frequency at the wave front and surf. It can even become fixed in certain regions of the Old World or show frequency gradients over large areas. (C) Two random simulation outputs (after 300 generations) of the spatial segregation of two alleles during a radiating range expansion [see Excoffier and Ray (2008) for simulation details]. The initial state is an equal proportion of both alleles (black and white) in the central white circle. As the expansion wave progresses outward, one of the two alleles becomes rapidly fixed, which generates typical radiating zones with a unique allele (Hallatschek et al. 2007). (D) Example of the possible dynamics of the spatial distribution of a new neutral mutation (in white) occurring at the wave front of an expansion, at times $t = 100$, $t = 300$, and $t = 500$ generations after the start of an expansion in a rectangular world (of 25 by 100 demes) (Matthieu Foll, personal communication, 2008). In this example the mutation occurs at a relatively high frequency over a large but limited area.

rides on the wave of advance of a spatial expansion. This can more particularly happen when the spatial expansion occurs after a demographic bottleneck, when the population density is small, which causes increased rates of genetic drift at the wave front. This situation makes it possible for a rare allele to reach high frequencies and to be propagated to newly colonized territories. Surfing is favored when populations at the wave front grow rapidly and exchange few genes with their neighbors (Excoffier and Ray 2008; Hallatschek and Nelson 2008; Hallatschek et al. 2007). When an allele surfs successfully, it can increase in frequency over a large area in the newly colonized region (Klopfstein et al. 2006). In principle, any allele, neutral or not, can surf, and deleterious alleles can also be propagated by

surfing over long distances, even though the probability of such an event is small (Travis et al. 2007). However, if an allele is found at high frequencies over a large area, it is difficult to decide whether it was due to surfing, positive selection, or both. Therefore specific adaptations to new environments and to new pathogens that lead to local increases in allele frequencies (Storz et al. 2004) should be difficult to distinguish from surfing on a wave of advance. This may affect many genomewide tests of selection that identify outlier loci as candidates for further investigation (Nielsen et al. 2007), because we recently found that surfing may have affected a larger proportion of loci than originally thought (Hofer et al. 2009). Until now it was commonly admitted that the distinction between the effect of past demography and selection was possible because demographic events should affect all neutral loci across the genome, whereas selection would act on only a single or a few loci at a time (Stajich and Hahn 2005). However, this paradigm does not hold anymore, because surfing can affect only a few loci and can mimic all aspects of the molecular signatures left by selective events.

The fact that patterns of allele frequencies after surfing mimic selective processes may have led to erroneous conclusions in many studies that used a strong geographic structure of particular alleles as evidence of selective effects. A possible example of such a problem are two genes involved in the control of brain size, *MCPH1* and *ASPM* (Evans et al. 2006; Mekel-Bobrov et al. 2005), which show an increased frequency of a derived haplotype outside Africa, interpreted to be the consequence of local positive selection for increased cognitive abilities in non-African populations. However, the link between the presence of the derived haplotype and a phenotypic trait could not be found (Dobson-Stone et al. 2007; Mekel-Bobrov et al. 2007; Timpson et al. 2007; Woods et al. 2006; Yu et al. 2007). Moreover, we were able to show by performing spatially explicit simulations that similar geographic distributions of allele frequencies could be generated by allelic surfing of a neutral gene during the range expansion outside Africa (Currat et al. 2006).

Comparing Alternative Models of Human Evolution in a Spatially Explicit Context: Our Most Recent and Current Work

Applying SPLATCHE on a continental scale allowed us to test for alternative geographic origins of the global human expansion (Ray et al. 2005). In this study, we first showed that the origin of an expansion leaves a signature in the genetic data and that it is therefore possible to recover this origin provided that a large number of independent markers are used and precise information on past demography and potential places of origin is available. Using the original CEPH-HGDP panel with 377 STR loci, we simulated spatial expansions from 25 putative sources of dispersal evenly spread over the Old World. We also contrasted these unique origin models with nine multiregional evolution models. In all these multiregional evolution models, we considered three ancestral populations located in Africa, Europe, and central Asia. These three panmictic populations could exchange migrants

at different rates during a large number of generations. Each of these populations then underwent a spatial expansion to colonize the Old World. Our results showed that it was possible to unambiguously distinguish between a unique origin and a multiregional evolution model, and the application to the CEPH-HGDP panel pointed to a unique but surprising North African origin of modern human evolution (Ray et al. 2005). Because this result could be due to ascertainment bias in favor of markers selected to be polymorphic in Europeans, we explicitly considered this bias in our simulations. The new estimation that models this bias revealed that East Africa is the most likely place of origin for modern humans.

The study by Ray et al. (2005) was one of the first to try to quantitatively contrast alternative models of human evolution based on multilocus data. However, the study was based on differences in allele frequencies (through F_{ST}) and therefore did not make use of the full molecular information contained in the genetic data. We thus needed to develop a statistical framework that would allow us to use more information, in the form of many summary statistics computed on the observed data.

Estimation of Demographic Parameters of Complex Scenarios from Genetic Data. Since the advent of coalescent theory (Hudson 1983; Kingman 1982a, 1982b; Tajima 1983), tremendous progress has been made in the estimation of mutation and demographic parameters from genetic data, with the generalization of maximum-likelihood techniques that allow one to use more information present in the molecular data (Marjoram and Tavare 2006). Several methods relying on Markov chain Monte Carlo (MCMC) or importance sampling have been introduced to estimate population size (e.g., Kuhner et al. 1998; Wilson and Balding 1998), time to the most recent common ancestor (e.g., Tavare et al. 1997), population growth or decline (e.g., Beaumont 1999), and divergence times and/or migration rates between populations (e.g., Nielsen and Wakeley 2001) [see Kuhner (2009) for a review of available coalescent genealogy samplers]. However, full-likelihood methods are limited to relatively simple mutation models. Moreover, they are difficult to extend to the analysis of more than a few populations and still require a fairly long processing time.

Approximate Bayesian Computation

To tackle models for which the likelihood could not be computed, some investigators have proposed basing demographic estimations on summary statistics instead of the full data and estimating parameters by comparing summary statistics computed on observed data sets with those obtained from simulations, using an acceptance-rejection algorithm (Pritchard et al. 1999; Tavare et al. 1997). The idea consists in performing a large number of simulations and retaining only a small fraction of simulations (e.g., 1%) for which summary statistics are close to the observed ones. The range of parameter values found in this subset of simulations is then used to derive posterior distributions for each parameter of interest.

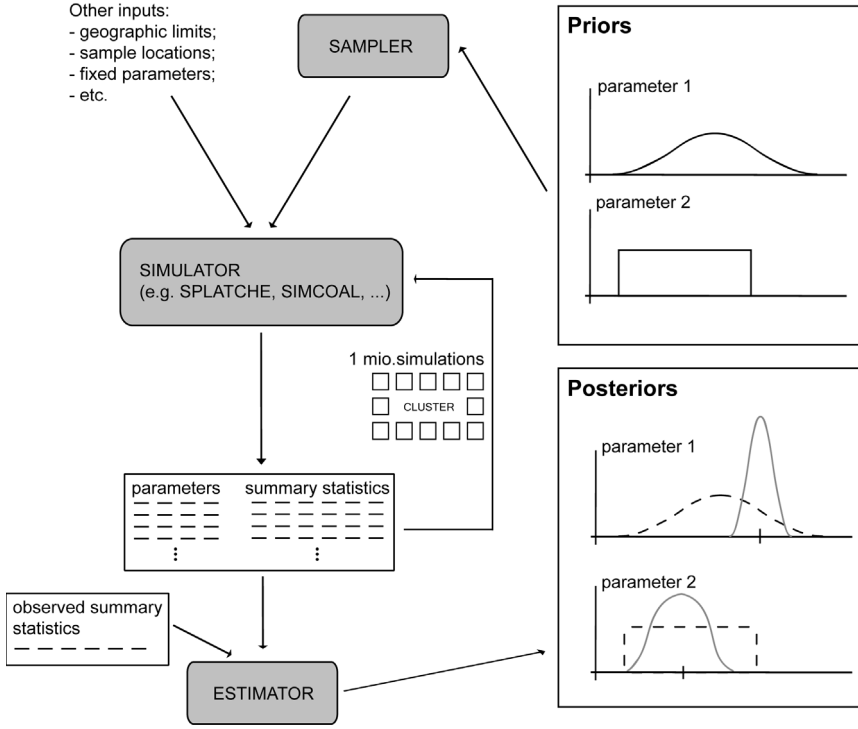

Figure 2. Schematic view of the SPLATCHE ABC framework for parameter inference under complex demographic models. For each demographic and/or genetic parameter of interest (only two in this example), prior probability distributions (priors) must first be defined. Priors can be uniformly distributed (uninformative), but they can also have a different shape (as for parameter 1) depending on prior knowledge. Based on these priors, sampled values are drawn at random (by a sampler) and taken as input for a simulator (e.g., SPLATCHE) that carries out one simulation and outputs the corresponding summary statistics. The sampling and simulation steps are repeated a large number of times (e.g., 1 million times), typically using a computer cluster. Once the required number of simulations is obtained, an estimator computes the Euclidean distances between observed and simulated summary statistics, and the posterior distribution of each parameter (gray lines in the Posteriors panel) are derived using the regression method of Beaumont et al. (2002). Point estimates and associated credible intervals can then be derived from the posterior distributions.

Beaumont et al. (2002) extended and formalized this procedure, which they called approximate Bayesian computation (ABC) (see Figure 2). They proposed estimating parameters from a set of accepted values by performing a local linear regression of statistics on parameters in which simulations are weighted as a function of the Euclidean distance computed between observed and simulated summary statistics. In that way, simulated parameter values that showed more similarity to the observation would be given more weight. This procedure im-

proved on previous methods, making it possible to obtain similar estimates with much fewer simulations. This ABC technique has been applied essentially to relatively complex scenarios of colonization (Pascual et al. 2007), spatial expansion (Estoup et al. 2004; François et al. 2008; Hamilton et al. 2005a; Neuenschwander et al. 2008), pest introduction (Miller et al. 2005), pathogen transmission (Tanaka et al. 2006), population bottleneck (Chan et al. 2006), complex phylogeography (Hickerson et al. 2006), and admixture (Excoffier et al. 2005).

Application of ABC to Human Evolution. Until recently, the application of ABC to questions related to human evolutionary biology was scarce. This was probably a result of a lack of appropriate data sets, lack of computational power, and lack of tools to automate ABC inference. Our team has applied these techniques in the context of human evolution and ethology (Fagundes et al. 2007; Hamilton et al. 2005b). Concerning the ethology application (Hamilton et al. 2005b), we have analyzed patterns of mitochondrial and Y-chromosome diversity in three matrilocal and three patrilocal groups from northern Thailand to infer by means of ABC the number of males and females arriving each generation in these populations and to estimate the age of their range expansion. Our ABC estimates indicate many fewer male immigrants in patrilocal than in matrilocal populations and a reverse situation for female immigrants, as one would have predicted. However, the interesting result was that matrilocal populations were much more porous to male immigration than patrilocal populations were to female immigrants, showing a clear asymmetry between these two societal types.

We also used the ABC framework to statistically evaluate the relative probabilities of alternative models of human evolution (Fagundes et al. 2007). This simulation study was partly spatial, because the three main regions of interest (Africa, Asia, and the Americas) were considered unsubdivided. Using DNA data from 50 nuclear loci sequenced in samples from these three regions, we showed that a simple African replacement model with exponential growth had a higher probability (78%) than alternative multiregional evolution or assimilation scenarios. We also inferred the time of colonization of the Americas to be about 10,000 years BP (95% highest posterior density credible interval: 7,647–15,945 years BP). To more accurately study the details of the early history of Amerindians, many more populations and markers are necessary. We are currently using ABC to estimate various parameters linked to the colonization of the Americas by using a data set of 29 Amerindian populations (Wang et al. 2007) typed at 680 STR markers.

Improvements of ABC Techniques. In principle, the ABC method could allow one to estimate parameters of any scenario for which genetic data can be simulated. However, the applicability of this method is limited because it might require millions of simulations to give accurate results for complex scenarios; this might be prohibitive when simulations are computer intensive (Marjoram et al. 2003). Moreover, the choice of informative summary statistics can sometimes be problematic. To improve the efficiency of ABC techniques and to perform fewer simu-

lations for the same final accuracy, Marjoram et al. (2003) proposed performing a Markov chain Monte Carlo simulation without computing likelihoods. This ABC-MCMC technique has been recently refined (Wegmann et al. 2009) and leads to estimations comparable to conventional ABC with about 50 times less simulations. Another way to gain efficiency is to perform sequential rounds of simulations, each time concentrating in the parameter space leading to summary statistics close to the observations (Beaumont et al. 2009; Sisson et al. 2007). A way to address the problem of choosing the right statistics has been addressed (Wegmann et al. 2009) by using partial least-squares (PLS) (see, e.g., Tenenhaus et al. 1995). PLS indeed transforms summary statistics into a reduced set of orthogonal components that also maximally explain the variance of the parameters. Note that Blum and François (2009) recently addressed the same issue by using nonlinear regression models and regularization (ridge regression).

These improvements should be helpful to study complex and realistic scenarios of human evolution, which are extremely costly to simulate. For example, spatial expansions on a continental scale coupled with long-distance dispersal can be extremely computer intensive, and any reduction in the required number of simulations to be performed would allow us to test many more alternative scenarios (e.g., various intensities of long-distance dispersal, different preferred routes of migrations), which would increase our confidence in the inferential process.

Range Expansion into an Occupied Territory: Interactions with Local Species. The expansion of a species into an occupied territory leads to the question of its interactions with local populations. If the two species can interact and interbreed or if a population invades the territory of another population of the same species, one would expect the invading species to be progressively introgressed by the genes of the local species and that a gradient of gene frequencies would develop along the expansion axis. This phenomenon is the basis for the theory of demic diffusion (Ammerman and Cavalli-Sforza 1971, 1973; Sokal et al. 1991), which was developed in the context of the colonization of Europe by Neolithic farmers to explain gradients of genetic diversity between the Near East and the British Isles. These gradients have been shown mainly using principal components analysis, which is a general method for summarizing high-dimensional data. For example, the classic work of Cavalli-Sforza et al. (1994) makes use of principal components analysis to concentrate hundreds of allele frequency maps into a few synthetic maps corresponding to the first few principal component axes. The clines or gradients visually inferred from these maps were thought to correspond to historical migration events. However, these claims have been controversial, and a recent paper [Novembre and Stephens (2008), but see also Reich et al. (2008)] clearly demonstrated that these clines do not necessarily reflect directed migrations but may simply be due to isolation by distance, whereby there is gene flow at equilibrium only among neighboring populations.

Understanding the genetic signature left by admixture events is important. We have been able to study this phenomenon using an extended version of our

SPLATCHE software applied to two major interaction events in human history. In our first study, we addressed the demic diffusion issue by modeling the spread of Neolithic populations in Europe some 10,000 years ago (Currat and Excoffier 2005) and their potential interaction with existing Paleolithic populations [see Bentley et al. 2009 (this issue)]. We found that if Neolithic individuals freely admixed with Paleolithic people, we should expect a massive Paleolithic contribution to the current European gene pool, whereas current estimates of the Paleolithic contribution are around only 50% at most. Currat and Excoffier (2005) also examined the effect of range expansions and of population interaction on the formation of allele frequency clines and found that a pure range expansion can lead to clines with or without contacts with a preexisting population. Therefore the observation of clines does not in itself support the idea of a progressive dilution of a Neolithic gene pool in the Paleolithic population.

In a second study on admixture we modeled the range expansion of modern humans into Europe and their potential interaction with preexisting Neanderthals some 40,000 years ago (Currat and Excoffier 2004). The main result of this study was to show that if modern humans had admixed with Neanderthals during the colonization process, we would expect to see a large Neanderthal contribution in the gene pool of modern humans. Because we know this is not the case [from the study of Neanderthal mitochondrial and nuclear DNA (e.g., Caramelli et al. 2003; Excoffier 2006; Green et al. 2006)], we estimated the maximum possible contribution of Neanderthal input into our gene pool. We found an extremely small contribution, arguing in favor of no genetic contacts between Neanderthals and modern humans. This finding was later backed up by simulated work showing that massive introgression of neutral genes is to be expected during the invasion of an occupied territory if interbreeding is not severely prevented between the invading and the local species (Currat et al. 2008).

Significant Unsolved Problems

Building Priors with Archaeological Data (Dates of First Arrival, Spatial Density and Distribution of Artifacts). The ABC framework and other Bayesian approaches rely on prior information drawn from previous studies. When implementing human evolutionary models, one can make use of archaeological or paleoanthropological information to build model prior probability distributions. Prior information on parameters such as times of the onset of a range expansion or of the entrance into a given area (e.g., into the Americas through the Bering Strait) would greatly benefit from dates of archaeological findings. What is typically done is to consider the date of the oldest artifact as a lower bound for the prior. When it comes to calibrating a wave of expansion without genetic information, first occupancy dates of many sites over a large area can be used (Martino et al. 2007; Pinhasi et al. 2005). However, the ability to accurately predict colonization rate can be severely constrained depending on the interplay between sample locations and the uncertainty on radiocarbon dates (Hazelwood and Steele 2004).

Other key parameters that archaeological data could provide are those linked to population size, notably the carrying capacity of different types of environment. Under some assumptions, artifact densities can be used as a surrogate for population density [Chamberlain 2009 (this issue); Steele et al. 1998], which can help calibrate carrying capacities. However, these studies do not incorporate the two main types of uncertainty linked to these dates: the uncertainty of the date itself and the latency period that exists between the real first occupancy date and the time of visibility in the archaeological record. For the uncertainty of the date itself, the probability distribution of radiocarbon dates could be used instead of the point estimate. Within an ABC framework, this would require a method that could correctly account for the probabilistic nature of an arrival date by retaining only those simulations that satisfy the constraints imposed by the empirical distribution. For the second type of uncertainty, it should be possible to model the latency period. If one assumes that archaeological visibility at a given site is partly related to the accumulated number of individuals who lived on this site, then estimated carrying capacities could be used to parameterize the latency model. Correctly accounting for the uncertainty surrounding first occupancy dates could help to decipher apparent discrepancies in the temporal succession of first occupancy dates, notably in the contexts of the Neolithic in Europe (Pinhasi et al. 2005) and the rapid diffusion of humans in South America (Hazelwood and Steele 2004; Martino et al. 2007).

Archaeological (or historical) and genetic data should ideally be considered jointly in the inferential process. Our team is currently involved in a project in which an ABC framework based on SPLATCHE simulations is used to analyze the range expansion of cane toads in Australia (Estoup et al. 2009), for which dates of first occupancy are known at several locations. Preliminary results show that coupling the historical data with genetic data is the most powerful way to infer demographic parameters such as migration rates and population densities. However, the temporal scale of this project (a few dozen generations) and the ecology of the species (high migration rates and population sizes) are quite different from our human evolutionary context. It is still not known how to best exploit different types of information for inferences about past human demography. One way to approach this would be to first retain simulations compatible with historical information (e.g., dates of first occupancy, accounting for uncertainties) (see Currat and Excoffier 2005) and then compute genetic summary statistics on the retained simulations to perform classical ABC parameter estimation. This approach would be helpful to better understand the settlement dynamics and past demography of North and South America, especially when long-distance dispersal is modeled.

Range Contractions and Reexpansions. Earth's climate has been extremely variable during the Pleistocene and the Holocene, with long glacial maxima occurring approximately every 100,000 years, interrupted by relatively short interglacial periods (Jouzel et al. 2007). The climatic changes that have occurred in the Pleistocene are thought to have thus greatly affected the distribution of most animal and plant species in paleoarctic regions. The effects have been a series

of range shifts, range contractions and expansions, allopatric isolation, local extinctions, demographic bottlenecks, and demographic expansions. The habitat of most species has thus considerably shrunk during cold periods and increased during warmer periods. In Europe the location of several last glacial maximum refugia have been inferred from the pattern of molecular diversity (e.g., Hewitt 1999, 2001, 2004; Melo-Ferreira et al. 2007; Taberlet et al. 1998) identifying the Iberian peninsula, Italy, and the Balkans as likely refuge areas during glaciations and therefore as potential sources of genetic diversity for species having presently a more northern range.

Climatic oscillations have certainly had a drastic impact on the demography of prehistoric human populations [see Riede 2009 (this issue)]. It is quite likely that the southern European refuge areas have been used by past hunter-gatherer populations and potentially also by Neanderthals. Range contractions and reexpansions probably had a lasting effect on the genetic makeup of Europeans. However, if the consequences of range expansion are beginning to be well understood, this is not the case for range contractions. It is still unclear how a range contraction differs from a bottleneck in unsubdivided populations at the molecular level. Moreover, we do not know how long a genetic signature of a contraction is visible in reexpanding populations or under what conditions a range contraction can erase genetic diversity. These questions could be addressed using spatially explicit simulations and would greatly benefit from studies in paleodemography. The key parameters of such simulations are indeed anticipated to be those linked to migration rates and population carrying capacities during the contraction and expansion phases.

Long-Distance Dispersal During Initial Human Colonization and More Recently as Recurrent Gene Flow. Range expansions are more easily simulated using a stepping-stone model, in which migrants are sent only to directly adjacent demes. This approach is best suited for organisms with low vagility and an effective dispersal strategy that reflects this simulated behavior. However, long-distance dispersal (LDD) events, defined as rare migration events that occur at much longer distances than the mean dispersal distance, may have played a major role in the colonization dynamics of many animal species. Long-range migrations speed up the colonization process and lead to a more patchy spatial distribution of allele frequencies (Ibrahim et al. 1996; Nichols and Hewitt 1994). Moreover, Bialozyt et al. (2006) recently showed, using forward simulations, that genetic diversity can be reduced for low levels of LDD but increased for higher levels of LDD. In simulations LDD parameters are linked to the dispersal kernel that characterizes the probability of occurrence and the distance of LDD events. For a given mean dispersal distance, the shape of the dispersal kernel (e.g., fat- or thin-tailed) does not seem to matter too much for the establishment of geographic patterns of allele frequencies in European human populations (Novembre et al. 2005), although this shape has been shown to play a major role in the mixing of plant propagules from discrete sources at long distance (Bohrer et al. 2005; Klein et al. 2006).

It is not clear whether LDD was important during the initial colonization of the world by past human populations. It is also not known whether LDD events were recurrent among distant populations once these populations were established (long after the initial colonization). Some inferences about past LDD from genetic data would therefore be valuable. To better investigate the effect of LDD in humans, we modified our SPLATCHE software (Currat et al. 2004) by incorporating LDD into our simulations. We are currently exploring the signature left by LDD events in a simple world with uniform resource distribution. Preliminary results show that it might be possible to infer the amount of LDD using a large number of markers in an ABC framework.

However, the extension of this work to a realistic landscape for particular episodes of past human dispersal (e.g., colonization of the Americas) seems difficult. The main challenge that needs to be addressed is the building of the priors linked to LDD parameters. The frequency of LDD, the mean LDD distance, and the environment-specific occurrence of LDD events (e.g., only along coastlines) are parameters that would greatly benefit from information from other disciplines. Information on these priors would help to reduce the parameter space to be explored in the ABC framework and would certainly provide much better estimates of heterogeneous dispersal.

Acknowledgments This work was supported by the Swiss National Science Foundation through grant 3100A0-112072 awarded to L. Excoffier. We thank two anonymous reviewers for their comments on an earlier draft.

Received 13 January 2009; revision accepted for publication 24 April 2009.

Literature Cited

Akey, J. M., M. A. Eberle, M. J. Rieder et al. 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biol*. 2(10):e286.

Akey, J. M., G. Zhang, K. Zhang et al. 2002. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res*. 12(12):1805–1814.

Ammerman, A. J., and L. L. Cavalli-Sforza. 1971. Measuring the rate of spread of early farming in Europe. *Man* 6(4):674–688.

Ammerman, A. J., and L. L. Cavalli-Sforza. 1973. A population model for the diffusion of early farming in Europe. In *The Explanation of Culture Change*, C. Renfrew, ed. London: Duckworth, 343–357.

Barreiro, L. B., G. Laval, H. Quach et al. 2008. Natural selection has driven population differentiation in modern humans. *Nat. Genet.* 40(3):340–345.

Beaumont, M. A. 1999. Detecting population expansion and decline using microsatellites. *Genetics* 153(4):2013–2029.

Beaumont, M. A., J.-M. Cornuet, J.-M. Marin et al. 2009. Adaptivity for ABC algorithms: The ABC-PMC scheme. *Biometrika* (in press).

Beaumont, M. A., W. Y. Zhang, and D. J. Balding. 2002. Approximate Bayesian computation in population genetics. *Genetics* 162(4):2025–2035.

Bentley, R. A., R. H. Layton, and J. Tehrani. 2009. Kinship, marriage, and the genetics of past human dispersals. *Hum. Biol.* 81(2–3):159–179 (this issue).

Bialozyt, R., B. Ziegenhagen, and R. J. Petit. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *J. Evol. Biol.* 19(1):12–20.

Blum, M. G. B., and O. François. 2009. Nonlinear regression models for approximate Bayesian computation. *Statist. Comput*. (in press).

Bohrer, G., R. Nathan, and S. Volis. 2005. Effects of long-distance dispersal for metapopulation survival and genetic structure at ecological time and spatial scales. *J. Ecol.* 93(5):1029–1040.

Campbell, M. C., and S. A. Tishkoff. 2008. African genetic diversity: Implications for human demographic history, modern human origins, and complex disease mapping. *Annu. Rev. Genom. Hum. Genet.* 9(1):403–433.

Caramelli, D., C. Lalueza-Fox, C. Vernesi et al. 2003. Evidence for a genetic discontinuity between Neanderthals and 24,000-year-old anatomically modern Europeans. *Proc. Natl. Acad. Sci. USA* 100(11):6593–6597.

Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza. 1994. *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.

Chamberlain, A. 2009. Archaeological demography. *Hum. Biol.* 81(2–3):275–286 (this issue).

Chan, Y. L., C. N. K. Anderson, and E. A. Hadly. 2006. Bayesian estimation of the timing and severity of a population bottleneck from ancient DNA. *PLoS Genet.* 2(4):e59.

Currat, M., and L. Excoffier. 2004. Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol*. 2(12):2264–2274.

Currat, M., and L. Excoffier. 2005. The effect of the Neolithic expansion on European molecular diversity. *Proc. Biol. Sci*. 272(1564):679–688.

Currat, M., L. Excoffier, W. Maddison et al. 2006. Comment on "Ongoing Adaptive Evolution of ASPM, a Brain Size Determinant in *Homo sapiens*" and "Microcephalin, a Gene Regulating Brain Size, Continues to Evolve Adaptively in Humans." *Science* 313(5784):172.

Currat, M., N. Ray, and L. Excoffier. 2004. SPLATCHE: A program to simulate genetic diversity taking into account environmental heterogeneity. *Mol. Ecol. Notes* 4(1):139–142.

Currat, M., M. Ruedi, R. J. Petit et al. 2008. The hidden side of invasions: Massive introgression by local genes. *Evolution* 62(8):1908–1920.

Dobson-Stone, C., J. M. Gatt, S. A. Kuan et al. 2007. Investigation of *MCPH1* G37995C and *ASPM* A44871G polymorphisms and brain size in a healthy cohort. *Neuroimage* 37(2):394–400.

Edmonds, C. A., A. S. Lillie, and L. L. Cavalli-Sforza. 2004. Mutations arising in the wave front of an expanding population. *Proc. Natl. Acad. Sci. USA* 101(4):975–979.

Estoup, A., S. J. E. Baird, N. Ray et al. 2009. Combining genetic, historical, and geographic data to reconstruct the dynamics of the bioinvasion of cane toad *Bufo marinus*. Unpublished.

Estoup, A., M. Beaumont, F. Sennedot et al. 2004. Genetic analysis of complex demographic scenarios: Spatially expanding populations of the cane toad, *Bufo marinus*. *Evol. Intl. J. Org. Evol.* 58(9):2021–2036.

Eswaran, V. 2002. A diffusion wave out of Africa: The mechanism of the modern human revolution? *Curr. Anthropol.* 43(5):749–774.

Eswaran, V., H. Harpending, and A. R. Rogers. 2005. Genomics refutes an exclusively African origin of humans. *J. Hum. Evol.* 49(1):1–18.

Evans, P. D., N. Mekel-Bobrov, E. J. Vallender et al. 2006. Evidence that the adaptive allele of the brain size gene microcephalin introgressed into *Homo sapiens* from an archaic *Homo* lineage. *Proc. Natl. Acad. Sci. USA* 103(48):18,178–18,183.

Excoffier, L. 2002. Human demographic history: Refining the recent African origin model. *Curr. Opin. Genet. Dev.* 12(6):675–682.

Excoffier, L. 2006. Neanderthal genetic diversity: A fresh look from old samples. *Curr. Biol.* 16(16): R650–R652.

Excoffier, L., A. Estoup, and J. M. Cornuet. 2005. Bayesian analysis of an admixture model with mutations and arbitrarily linked markers. *Genetics* 169(3):1727–1738.

Excoffier, L., and N. Ray. 2008. Surfing during population expansions promotes genetic revolutions and structuration. *Tr. Ecol. Evol.* 23(7):347–351.

Fagundes, N. J., N. Ray, M. Beaumont et al. 2007. Statistical evaluation of alternative models of human evolution. *Proc. Natl. Acad. Sci. USA* 104(45):17,614–17,619.

François, O., M. G. B. Blum, M. Jakobsson et al. 2008. Demographic history of European populations of *Arabidopsis thaliana*. *PLoS Genet.* 4(5):e1000075.

Gonder, M. K., H. M. Mortensen, F. A. Reed et al. 2007. Whole-mtDNA genome sequence analysis of ancient African lineages. *Mol. Biol. Evol.* 24(3):757–768.

Green, R. E., J. Krause, S. E. Ptak et al. 2006. Analysis of one million base pairs of Neanderthal DNA. *Nature* 444(7117):330–336.

Hallatschek, O., P. Hersen, S. Ramanathan et al. 2007. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci. USA* 104(50):19,926–19,930.

Hallatschek, O., and D. R. Nelson. 2008. Gene surfing in expanding populations. *Theor. Popul. Biol.* 73(1):158–170.

Hamilton, G., M. Currat, N. Ray et al. 2005a. Bayesian estimation of recent migration rates after a spatial expansion. *Genetics* 170(1):409–417.

Hamilton, G., M. Stoneking, and L. Excoffier. 2005b. Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. *Proc. Natl. Acad. Sci. USA* 102(21):7476–7480.

Handley, L. J. L., A. Manica, J. Goudet et al. 2007. Going the distance: Human population genetics in a clinal world. *Tr. Genet.* 23(9):432–439.

Harding, R. M., and G. McVean. 2004. A structured ancestral population for the evolution of modern humans. *Curr. Opin. Genet. Dev.* 14(6):667–674.

Hazelwood, L., and J. Steele. 2004. Spatial dynamics of human dispersals: Constraints on modeling and archaeological validation. *J. Archaeol. Sci.* 31:669–679.

Hewitt, G. M. 1999. Postglacial recolonization of European biota. *Biol. J. Linn. Soc.* 68:87–112.

Hewitt, G. M. 2001. Speciation, hybrid zones, and phylogeography—or seeing genes in space and time. *Mol. Ecol.* 10(3):537–549.

Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R. Soc. Lond. B* 359(1442):183–195.

Hickerson, M. J., E. A. Stah, and H. A. Lessios. 2006. Test for simultaneous divergence using approximate Bayesian computation. *Evolution* 60(12):2435–2453.

Hofer, T., N. Ray, D. Wegmann et al. 2009. Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Ann. Hum. Genet.* 73:95–108.

Hudson, R. R. 1983. Testing the constant-rate neutral allele model with protein sequence data. *Evolution* 37:203–217.

Ibrahim, K. M., R. A. Nichols, and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* 77:282–291.

Jouzel, J., V. Masson-Delmotte, O. Cattani et al. 2007. Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science* 317(5839):793–796.

Kingman, J. F. C. 1982a. The coalescent. *Stoch. Proc. Their Appl.* 13:235–248.

Kingman, J. F. C. 1982b. On the genealogy of large populations. *J. Appl. Probab.* 19A:27–43.

Klein, E. K., C. Lavigne, and P. H. Gouyon. 2006. Mixing of propagules from discrete sources at long distance: Comparing a dispersal tail to an exponential. *BMC Ecol.* 6(1):3.

Klopfstein, S., M. Currat, and L. Excoffier. 2006. The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* 23(3):482–490.

Kuhner, M. K. 2009. Coalescent genealogy samplers: Windows into population history. *Tr. Ecol. Evol.* 24(2):86–93.

Kuhner, M. K., J. Yamato, and J. Felsenstein. 1998. Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* 149(1):429–434.

Lewontin, R. C., and J. Krakauer. 1973. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* 74(1):175–195.

Manica, A., F. Prugnolle, and F. Balloux. 2005. Geography is a better determinant of human genetic differentiation than ethnicity. *Hum. Genet.* 118(3–4):366–371.

Marjoram, P., J. Molitor, V. Plagnol et al. 2003. Markov chain Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. USA* 100(26):15,324–15,328.

Marjoram, P., and S. Tavare. 2006. Modern computational approaches for analyzing molecular genetic variation data. *Nat. Rev. Genet.* 7(10):759–770.

Martino, L. A., A. Osella, C. Dorso et al. 2007. Fisher equation for anisotropic diffusion: Simulating South America human dispersals. *Phys. Rev. E* 76:1–10.

Mekel-Bobrov, N., S. L. Gilbert, P. D. Evans et al. 2005. Ongoing adaptive evolution of *ASPM*, a brain size determinant in *Homo sapiens*. *Science* 309(5741):1720–1722.

Mekel-Bobrov, N., D. Posthuma, S. L. Gilbert et al. 2007. The ongoing adaptive evolution of *ASPM* and microcephalin is not explained by increased intelligence. *Hum. Mol. Genet.* 16(6):600–608.

Melo-Ferreira, J., P. Boursot, E. Randi et al. 2007. The rise and fall of the mountain hare (*Lepus timidus*) during Pleistocene glaciations: Expansion and retreat with hybridization in the Iberian peninsula. *Mol. Ecol.* 16(3):605–618.

Miller, N., A. Estoup, S. Toepfer et al. 2005. Multiple transatlantic introductions of the Western Corn Rootworm. *Science* 310(5750):992.

Neuenschwander, S., C. R. Largiader, N. Ray et al. 2008. Colonization history of the Swiss Rhine basin by the bullhead (*Cottus gobio*): Inference under a Bayesian spatially explicit framework. *Mol. Ecol.* 17(3):757–772.

Nichols, R. A., and G. M. Hewitt. 1994. The genetic consequences of long-distance dispersal during colonization. *Heredity* 72:312–317.

Nielsen, R., I. Hellmann, M. Hubisz et al. 2007. Recent and ongoing selection in the human genome. *Nat. Rev. Genet.* 8(11):857–868.

Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: A Markov chain Monte Carlo approach. *Genetics* 158(2):885–896.

Novembre, J., A. P. Galvani, and M. Slatkin. 2005. The geographic spread of the *CCR5* Delta32 HIV-resistance allele. *PLoS Biol*. 3(11):e339.

Novembre, J., and M. Stephens. 2008. Interpreting principal component analyses of spatial population genetic variation. *Nat. Genet.* 40(5):646–649.

Pascual, M., M. P. Chapuis, F. Mestres et al. 2007. Introduction history of *Drosophila subobscura* in the New World: A microsatellite-based survey using ABC methods. *Mol. Ecol.* 16(15):3069–3083.

Pinhasi, R., J. Fort, and A. J. Ammerman. 2005. Tracing the origin and spread of agriculture in Europe. *PLoS Biol.* 3(12):e410.

Pritchard, J. K., M. T. Seielstad, A. Perez-Lezaun et al. 1999. Population growth of human Y chromosomes: A study of Y-chromosome microsatellites. *Mol. Biol. Evol.* 16(12):1791–1798.

Prugnolle, F., A. Manica, and F. Balloux. 2005. Geography predicts neutral genetic diversity of human populations. *Curr. Biol.* 15(5):R159–R160.

Ramachandran, S., O. Deshpande, C. C. Roseman et al. 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc. Natl. Acad. Sci. USA* 102(44):15,942–15,947.

Ray, N., M. Currat, P. Berthier et al. 2005. Recovering the geographic origin of early modern humans by realistic and spatially explicit simulations. *Genome Res.* 15(8):1161–1167.

Ray, N., M. Currat, and L. Excoffier. 2003. Intrademe molecular diversity in spatially expanding populations. *Mol. Biol. Evol.* 20(1):76–86.

Ray, N., M. Currat, and L. Excoffier. 2008. Incorporating environmental heterogeneity in spatially explicit simulations of human genetic diversity. In *Simulations, Genetics, and Human Prehistory*, S. Matsumura, P. Forster, and C. Renfrew, eds. Cambridge, U.K.: McDonald Institute for Archaeological Research, 103–117.

Reich, D. E., and D. B. Goldstein. 1998. Genetic evidence for a Paleolithic human population expansion in Africa. *Proc. Natl. Acad. Sci. USA* 95(14):8119–8123.

Reich, D., A. L. Price, and N. Patterson. 2008. Principal component analysis of genetic data. *Nat. Genet.* 40(5):491–492.

Riede, F. 2009. Climate and demography in early prehistory: Using calibrated ^{14}C dates as population proxies. *Hum. Biol.* 81(2–3):309–337 (this issue).

Rosenberg, N. A., J. K. Pritchard, J. L. Weber et al. 2002. Genetic structure of human populations. *Science* 298(5602):2381–2385.

Shriver, M. D., G. C. Kennedy, E. J. Parra et al. 2004. The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. *Hum. Genom.* 1(4):274–286.

Sisson, S. A., Y. Fan, and M. M. Tanaka. 2007. Sequential Monte Carlo without likelihoods. *Proc. Natl. Acad. Sci. USA* 104(6):1760–1765.

Sokal, R. R., N. L. Oden, and C. Wilson. 1991. Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature* 351(6322):143–145.

Stajich, J. E., and M. W. Hahn. 2005. Disentangling the effects of demography and selection in human history. *Mol. Biol. Evol.* 22(1):63–73.

Steele, J., J. Adams, and T. Sluckin. 1998. Modeling Paleo-Indian dispersals. *World Archaeol.* 30(2):286–305.

Storz, J. F., B. A. Payseur, and M. W. Nachman. 2004. Genome scans of DNA variability in humans reveal evidence for selective sweeps outside of Africa. *Mol. Biol. Evol.* 21(9):1800–1811.

Taberlet, P., L. Fumagalli, A. G. Wust-Saucy et al. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* 7(4):453–464.

Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105:437–460.

Tanaka, M., A. Francis, F. Luciani et al. 2006. Estimating tuberculosis transmission parameters from genotype data using approximate Bayesian computation. *Genetics* 173:1511–1520.

Tavare, S., D. J. Balding, R. C. Griffiths et al. 1997. Inferring coalescence times from DNA sequence data. *Genetics* 145(2):505–518.

Templeton, A. R. 2002. Out of Africa again and again. *Nature* 416:45–51.

Tenenhaus, M., J.-P. Gauchi, and C. Ménardo. 1995. Regression PLS and applications. *Rev. Stat. Appl.* 43(1):7–63.

Timpson, N., J. Heron, G. D. Smith et al. 2007. Comment on papers by Evans et al. and Mekel-Bobrov et al. on evidence for positive selection of *MCPH1* and *ASPM*. *Science* 317(5841):1036.

Travis, J. M. J., T. Munkemuller, O. J. Burton et al. 2007. Deleterious mutations can surf to high densities on the wave front of an expanding population. *Mol. Biol. Evol.* 24(10):2334–2343.

Wang, S., M. Jakobsson, J. C. M. Lewis et al. 2007. Genetic variation and population structure in Native Americans. *PLoS Genet.* 3(11):e185.

Wegmann, D., M. Currat, and L. Excoffier. 2006. Molecular diversity after a range expansion in heterogeneous environments. *Genetics* 174:2009–2020.

Wegmann, D., L. Leuenberger, and L. Excoffier. 2009. Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics* (in press).

Williamson, S. H., R. Hernandez, A. Fledel-Alon et al. 2005. Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proc. Natl. Acad. Sci. USA* 102(22):7882–7887.

Wilson, I. J., and D. J. Balding. 1998. Genealogical inference from microsatellite data. *Genetics* 150(1):499–510.

Wolpoff, M. H., J. Hawks, and R. Caspari. 2000. Multiregional, not multiple origins. *Am. J. Phys. Anthropol.* 112(1):129–136.

Woods, R. P., N. B. Freimer, J. A. De Young et al. 2006. Normal variants of microcephalin and *ASPM* do not account for brain size variability. *Hum. Mol. Genet.* 15(12):2025–2029.

Wu, C. I., and C. T. Ting. 2004. Genes and speciation. *Nat. Rev. Genet.* 5(2):114–122.

Yu, F., R. S. Hill, S. F. Schaffner et al. 2007. Comment on "Ongoing Adaptive Evolution of *ASPM*, a Brain Size Determinant in *Homo sapiens*." *Science* 316(5823):370.