



Genetic Consequences of Range Expansions

Author(s): Laurent Excoffier, Matthieu Foll and Rémy J. Petit

Source: *Annual Review of Ecology, Evolution, and Systematics*, Vol. 40 (2009), pp. 481-501

Published by: Annual Reviews

Stable URL: <http://www.jstor.org/stable/20744051>

Accessed: 06-03-2017 02:26 UTC

REFERENCES

Linked references are available on JSTOR for this article:

http://www.jstor.org/stable/20744051?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://about.jstor.org/terms>



Annual Reviews is collaborating with JSTOR to digitize, preserve and extend access to *Annual Review of Ecology, Evolution, and Systematics*

Genetic Consequences of Range Expansions

Laurent Excoffier,¹ Matthieu Foll,¹ and Rémy J. Petit²

¹Computational and Molecular Population Genetics Lab (CMPG), Institute of Ecology and Evolution, University of Berne, 3012 Berne, Switzerland and Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; email: laurent.excoffier@iee.unibe.ch

²INRA, UMR Biodiversity, Genes and Communities, F-33610 Cestas, France and Université de Bordeaux, UMR Biodiversity, Genes and Communities, F-33610 Cestas, France

Annu. Rev. Ecol. Evol. Syst. 2009. 40:481–501

First published online as a Review in Advance on September 2, 2009

The *Annual Review of Ecology, Evolution, and Systematics* is online at ecolsys.annualreviews.org

This article's doi:
10.1146/annurev.ecolsys.39.110707.173414

Copyright © 2009 by Annual Reviews.
All rights reserved

1543-592X/09/1201-0481\$20.00

Key Words

bioinvasion, coalescent simulations, introgression, long-range dispersal, range expansion, surfing

Abstract

Although range expansions have occurred recurrently in the history of most species, their genetic consequences have been little investigated. Theoretical studies show that range expansions are quite different from pure demographic expansions and that the extent of recent gene flow conditions expected patterns of molecular diversity within and between populations. Spatially explicit simulation studies have led to unexpected and fascinating results about genetic patterns emerging after a range expansion. For instance, spatial expansions can generate allele frequency gradients, promote the surfing of rare variants into newly occupied territories, induce the structuring of newly colonized areas into distinct sectors of low genetic diversity, or lead to massive introgression of local genes into the genome of an invading species. Interestingly, most of these patterns had been previously attributed to distinct selective processes, showing that taking into account the dynamic nature of a species range can lead to a paradigm shift in our perception of evolutionary processes.

INTRODUCTION

Range expansions have occurred repeatedly in the history of most if not all species, and they are presently occurring at an increasing rate owing to rapid climatic changes. In many cases, the range of a species can be considered to be highly dynamic, with periods of directional or isotropic growth, of range expansions, and of contractions followed by re-expansions (Davis & Shaw 2001; Hewitt 1996, 2000; Taberlet & Cheddadi 2002). It seems therefore important to understand and model the genetic consequences of these recurrent events.

The classic symposium volume *The Genetics of Colonizing Species* (Baker & Stebbins 1965) included one of the first discussions of these genetic consequences. One of the two organizers of this symposium, the botanist Herbert Baker, had been working on the invasion of plants into the British Isles using “genecological” approaches to complement the fossil record (Baker 1948, Baker & Stebbins 1965), thereby largely anticipating the interest in the genetic consequences of range shifts driven by ice ages that started in the 1980s (Critchfield 1984, Cwynar & MacDonald 1987, Hewitt 1993). Ingredients that finally prompted more attention on this topic were demographic studies of invasions (e.g., Hengeveld 1988) and detailed reconstructions of past distributions of species using fossil data (e.g., Huntley & Birks 1983). Interestingly, the ability of invasions to trigger introgression and to generate hybrid zones or admixture were identified early on (Ammerman & Cavalli-Sforza 1973, Baker 1948, Barton & Hewitt 1985, Gyllenstein & Wilson 1987, Hammer et al. 2001, Menozzi et al. 1978, Remington 1968), and invasions have been explicitly included in the study of speciation (Hewitt 1996). Another route toward the study of the genetic consequences of invasions emerged through the analysis of the variation of mitochondrial DNA in what came to be recognized as the field of phylogeography (Avise et al. 1987).

Overall, the long incubation time taken by the field of colonization genetics to develop might be attributed to the (enduring) tendency for researchers to focus on selection rather than history in their attempts to interpret spatial patterns of genetic structure and to the lack of appropriate tools to study invasions, including the lack of appropriate genetic markers and theoretical models. Although the spatial genetic diversity of many species is now being described and investigated (Storfer et al. 2007), realistic models explaining how this diversity has evolved are indeed still largely lacking. The dynamics and the stochasticity of spatial processes make them difficult to handle with conventional analytical methods, and the effects of spatial expansions have been somehow neglected until it was possible to perform realistic simulations. Although simulation studies are often criticized for their lack of generality, they have been particularly illuminating in the context of range expansions, mainly owing to their ability to display spatial patterns that are difficult if not impossible to describe by a few summary statistics.

In this review, we shall present the different models of demographic and range expansions that have been recently developed, including a few analytical and several simulation studies. We do not describe here the complex field of spatial population genetics, pioneered by the work of Wright and Malécot, among others (that is, Kimura 1953, Kimura & Weiss 1964, Malécot 1948, Maruyama 1969, Wright 1943), as it has been the object of several recent reviews (see e.g., Rousset 2007). We also do not review empirical studies of range expansion, but rather concentrate on the models and the processes involved in dynamic range expansion, underlining differences with models of pure demographic expansions but also some similarities with several selective processes.

GENETIC CONSEQUENCES OF PURE DEMOGRAPHIC EXPANSION

Spatial population expansions are usually accompanied by a large increase in the total number of individuals in the population, and one would expect a great similarity between the effects of spatial

and pure demographic expansions on genetic diversity. Hence, it appears useful to list some of the expected properties of expanding populations. In general, growing populations should show less genetic drift than stationary populations (Kimura & Crow 1963) and should have, therefore, relatively stable allele frequencies. Expanding populations should also show an excess of rare alleles and low frequency mutations, as well as an excess of homozygosity relative to the number of observed alleles (Maruyama & Fuerst 1984, 1985; Tajima 1989a; Watterson 1986). Finally, growing populations should have little linkage disequilibrium (LD) between closely linked loci (Pritchard & Przeworski 2001, Slatkin 1994), except if LD existed before the expansion (Slatkin 2008). However, as is made clear in the following sections, the observation of these patterns in spatially expanding populations depends on the levels of recent and past gene flow between subpopulations (Excoffier 2004, Ray et al. 2003).

Impact of Population Growth on Genetic Diversity

Coalescent theory provides a very intuitive framework to predict patterns of genetic diversity, which is why we briefly introduce it here. Coalescent approaches simply trace the genealogies of sampled lineages back in time until their Most Recent Common Ancestor (MRCA) (Kingman 1982). In **Figure 1-left**, we show two simulated gene genealogies for two independent loci. These two gene genealogies have quite different shapes and overall lengths owing to the high variability of the coalescent process. Although a full description of the properties of the coalescent process is out of the scope of the present review and can be found elsewhere (see e.g., Hudson 1990, Wakeley 2009), it may be enough to realize that, going backward in time, any two gene lineages can coalesce with probability $1/(2N)$ at each generation in a diploid population of size N . This probability is, thus, inversely proportional to the size of the population. The time it takes for two genes to coalesce will therefore, be longer in large rather than in small populations. This property is illustrated in **Figure 1-right**, where a small population has gone through a 100-fold expansion 500 generations ago. Although such an instantaneous expansion is quite unrealistic, it simply illustrates three general features of demographic expansions: (a) gene trees in expanding populations have long terminal branches, because coalescent events are much less likely in the currently large populations than in the smaller ancestral population, where they occur at a high rate and, thus, on a short timescale. As a consequence, they generally show a typical star-shape or comb-like structure (Rogers & Harpending 1992, Slatkin & Hudson 1991; see also **Figure 1-right**); (b) the topology and the size of gene genealogies in expanding populations are much more similar at unlinked loci than in cases of constant-size populations; and (c) most mutations will have accumulated on the terminal branches of the genealogy and, thus, will have small frequencies in the population. This last feature explains the excess of low-frequency variants in recently expanding populations noted above (Maruyama & Fuerst 1984, 1985) and introduces a typical skew toward singletons in the allele frequency spectrum (Harpending & Rogers 2000). The fact that most mutations occur on terminal branches in expanding populations also explains the quasi-absence of detectable linkage disequilibrium, even between tightly linked loci (Pritchard & Przeworski 2001, Slatkin 1994), because the classical LD coefficient $D = p_{ij} - p_i p_j$ is equal to $-1/n^2$ for singleton mutations (where $p_i = p_j = 1/n$ and $p_{ij} = 0$).

Other genetic properties of expanding population can be easily understood from coalescent trees. Early examination of mitochondrial DNA (mtDNA) sequence diversity showed that the distribution of the number of differences between pairs of sequences (also called mismatch distribution) was unimodal in most human populations (Di Rienzo & Wilson 1991). This is expected if a small population goes through a sudden and large demographic expansion (Li 1977, Rogers &

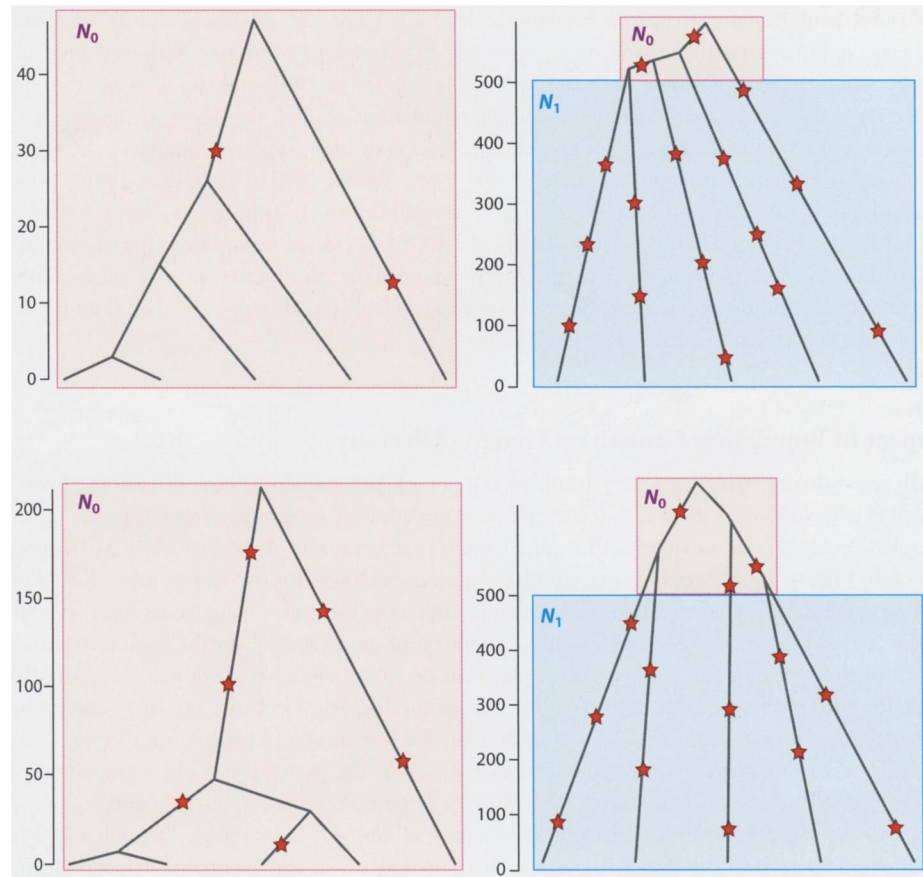


Figure 1

Random genealogies of five gene copies drawn from a stationary population (*left panels*) or from a population having gone through an instantaneous demographic expansion that occurred 500 generations ago (*right panels*). In our simulations, $N_0 = 100$ and $N_1 = 10,000$. Note that the timescales are in number of generations and that they are different for each genealogy. Red stars represent mutation events happening randomly along the genealogy according to a Poisson process.

Harpending 1992, Slatkin & Hudson 1991). Indeed, in such populations, most pairs of lineages should have a common ancestor living in the small ancestral population at about the same time (see **Figure 1**), and they will differ by about the same number of mutations. Therefore, the number of sites at which two DNA sequences differ is proportional to the expansion time (Rogers & Harpending 1992). This property was used to develop simple procedures to date the time of population expansions directly from mismatch distributions (Rogers & Harpending 1992, Schneider & Excoffier 1999). The accumulation of mutations on terminal branches affects the allele frequency spectrum. There is an excess of low-frequency mutations and a deficit of intermediate frequency mutations, which is precisely what would be expected after a recent selective sweep (Biswas & Akey 2006). Therefore, most tests of selective neutrality (Fu 1997, Fu & Li 1993, Tajima 1989b, Watterson 1978) should be significant in recently expanding populations because these tests also assume population stationarity.

GENETIC CONSEQUENCES OF SPATIAL EXPANSIONS

Many theoretical studies on range expansions have focused on the speed of the wave of advance, as pioneered by the work of Fisher (1937) and Skellam (1951), who showed that advantageous genes or invasive species should spread at a constant speed, unless (a) the environment is heterogeneous (e.g., Ackland et al. 2007, Garcia-Ramos & Rodriguez 2002, Kinezaki et al. 2003, Pialek & Barton 1997), (b) long-distance dispersal events occur relatively often (Clark 1998, Shigesada & Kawasaki 2002), or (c) there is interspecific competition (Aoki et al. 1996, Okubo et al. 1989, Shigesada et al. 1984). Given the difficulty of describing highly dynamic systems (e.g., Maruyama & Fuerst 1984, Nei et al. 1975), theoretical results about the effects of a range expansion on the genetic diversity of a subdivided population are still scarce, but are outlined below.

Genetic Diversity in Linear Habitats

Austerlitz et al. (1997) studied a simple colonization model of a linear stepping-stone habitat. They assumed that demes are colonized one after another from a single deme on one side of the range, that local deme density is logistically regulated with intrinsic rate of growth r and carrying capacity K , and that a given proportion m of migrants is sent to adjacent demes. Using a coalescent framework, they derived expected levels of diversity at different times of the colonization process in different portions of the range, as well as global levels of population differentiation as measured by F_{ST} . They found that genetic diversity decreases steadily along the range, owing to a series of consecutive bottlenecks or founder effects. As expected, the effect of these bottlenecks is smaller if demes have better recovery abilities ($r > 0.1$) (Nei et al. 1975), and if the number of migrants (Km) between demes is large. F_{ST} increases sharply during the colonization phase, and then decreases slowly, provided that $Km \geq 10$, but it remains larger than expected in the absence of a range expansion. Austerlitz et al. thus conclude that a steady decrease in genetic diversity along the expansion axis is a clear signature of a range expansion. This theory has been used to estimate the parameters of the expansion of modern humans from Eastern Africa using the observed steady decrease of heterozygosity with distance from Ethiopia (Liu et al. 2006, Prugnolle et al. 2005).

Le Corre & Kremer (1998) distinguished between the colonization process carried out by a given number (k) of founders to unoccupied areas and the migration process between occupied demes. This distinction was shown to affect the expected pattern of genetic diversity in metapopulations (Deshpande et al. 2009, Wade & McCauley 1988, Whitlock & McCauley 1990). They find that a reduced number of founders enhances the cumulative bottleneck effects and depletes genetic diversity, which drops sharply along the expansion axis, and, unless $k > 2Km = 8$, only populations located close to the source remain polymorphic. Hallatschek & Nelson (2008) further characterized the decay of genetic diversity during a one-dimensional colonization process and showed that it is decreasing exponentially at a rate $2/N_e$, where N_e is the effective population size on the wave front, which depends on the wave dynamics but grows sublinearly with carrying capacity K .

Infinite Island Approximation

In order to directly compare the effects of range expansions and pure demographic expansions in unsubdivided populations (Li 1977, Rogers & Harpending 1992, Slatkin & Hudson 1991), Excoffier (2004) described the expected pattern of diversity within and between a population having gone through an instantaneous range expansion under an infinite-island model approximation. This approach does not take into account recurrent founder effects occurring during a spatial range

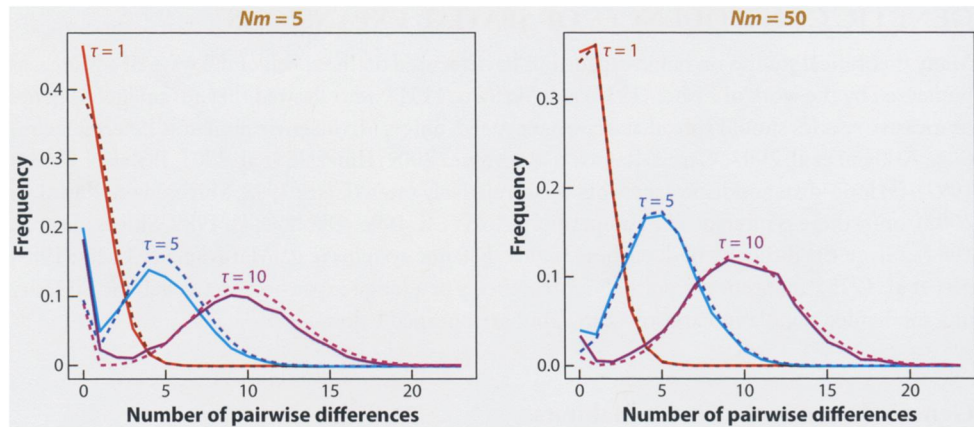


Figure 2

Mismatch distributions after an instantaneous range expansion. Solid lines show mismatch distributions expected under the infinite expansion model for different expansion times $\tau = 2\mu T = 1$ (red), 5 (blue), and 10 (purple). Dashed lines show empirical mismatch distributions obtained after an instantaneous expansion in a two-dimensional stepping-stone model for the same expansion times. The infinite-island model seems a very good approximation of the stepping-stone model, even though it underestimates the probability of gene identity, especially for low Nm values. Modified from figure 4 in Excoffier (2004).

expansion, but explicitly relates current patterns of diversity to the timing of the expansion and to migration rates between demes after the expansion. For relatively old expansions ($\tau = 2\mu T \geq 5$), the expected level of heterozygosity and the shape of the mismatch distribution are found to mainly depend on the number of genes (Nm) exchanged between demes (see **Figure 2**). In that case, the mismatch distributions are unimodal for large Nm values (say when $Nm > 10$), but bimodal for small Nm values. Indeed, for large Nm values and going backward in time, most genes migrate out of the sampled deme before they can coalesce, and most coalescent events are old and only occur in the deme at the origin of the expansion more than T generations ago. In contrast, for small Nm values many coalescent events are recent and occur in the sampled deme before genes migrate out, creating a second mode in the mismatch distribution. Note also that the infinite-island model provides a good approximation for the mismatch distribution expected under a spatially explicit two-dimensional stepping stone for large migration rates (see **Figure 2**). The expected level of population differentiation (F_{ST}) after an instantaneous spatial expansion quickly reaches its expected equilibrium value for large migration rates, but a much longer time ($\tau > 1$) is required to approach equilibrium when $Nm < 1$.

Interestingly, the relationship between expected patterns of within and between population diversity and the number of genes received from neighboring populations suggests that the parameter Nm could be estimated from single-population samples, contrasting with previous approaches estimating levels of gene flow from multiple population samples (see e.g., Beerli & Felsenstein 2001). Such a strategy has been successfully applied to estimate sex-specific migration rates after a range expansion in voles (Hamilton et al. 2005a) and humans (Hamilton et al. 2005b) using an approximate Bayesian computation (ABC) framework (Beaumont et al. 2002).

Expansions in Two Dimensions

Genetic diversity after a range expansion in a two-dimensional space has as yet only been explored by simulations. Using the SPLATCHE program (Curat et al. 2004), Ray et al. (2003) performed

explicit simulations of a progressive range expansion in a two-dimensional stepping-stone world and described patterns of DNA sequence diversity within and between populations. As in the instantaneous expansion case (Excoffier 2004), many aspects of molecular diversity, such as heterozygosity (H), the average number of pairwise differences (π), the number of segregating sites (S), or the shape of mismatch distributions are influenced by the raw number of migrants exchanged between neighboring demes (Nm or Km at carrying capacity). For a given expansion time τ , all measures of genetic diversity generally increase with larger Nm values. Moreover, several statistics commonly used to detect populations expansions or departure from selective neutrality, such as Tajima's D (Tajima 1989b) or Fu's F_S (Fu 1997), are only significant for large Nm values.

In other words, the ability to detect that a population has gone through a large range expansion is possible only if current populations exchange a large number of migrants per generation (typically when $Nm > 20$). This implies that a range expansion looks like a large demographic expansion only if there are many migrants exchanged between demes. This can happen for various combinations of migration rates and local population sizes. For instance, given similar dispersal abilities, populations with a large carrying capacity could show all the typical signals of a demographic expansion, like large negative Tajima's D and unimodal mismatch distributions, whereas smaller populations could show patterns perfectly compatible with population stationarity or even demographic decline, despite being part of a population having gone through a massive range expansion. This dependency on Nm could explain several unusual features such as the difference in patterns of mtDNA genetic diversity between human hunter-gatherer and food producer populations (Excoffier & Schneider 1999) or differences between sex-specific marker diversities in the case of sex-biased dispersal (Hamilton et al. 2005b).

Patterns of genetic diversity arising after a range expansion can be best understood by considering the coalescent trees arising in different situations. In **Figure 3**, we report such trees for different Nm values. On the one hand, when only a few migrants are exchanged between demes per generation (e.g., **Figure 3a**), several coalescent events occur within the sampled deme and, thus, are recent. On the other hand, for large migration rates (e.g., **Figure 3b,c**), gene lineages migrate out of the sampled deme before they can coalesce. This can be further visualized in **Figure 3**, where we report the spatial distributions of all coalescent events (*center row*) and the location of the MRCA (*bottom row*). For $Nm = 10$, most coalescent events occur in or close to the sampled deme, as well as along the axis of migration. Contrastingly, for larger migration rates, most coalescent events occur within a cone of colonization along the expansion axis, as well as close to or in the deme where the expansion started. These contrasting spatial distributions of coalescent events thus translate into bimodal mismatch distributions for low Nm values and into unimodal distributions for large Nm values (see **Figure 2**). The spatial distribution of the MRCA location, which is simply the location of the last coalescent event, is reported on the bottom row of **Figure 3**. MRCA's occur most often in the ancestral deme for large Nm values (91% and 99%, for $Nm = 100$ and 1000, respectively). By contrast, for small Nm values most coalescent events occur along the expansion axis (in 66% of the cases for $Nm = 10$), implying that all gene lineages sampled close to the edge of the settlement originated from a single gene after the onset of the expansion process (see section on Surfing below).

Spatial and Temporal Heterogeneity of the Environment

Most expansions do not occur in a homogeneous environment with constant carrying capacities. Wegmann et al. (2006) studied the effect of range expansions in a heterogeneous environment by performing simulations where the carrying capacity of different demes (K_i) were drawn from a Gamma distribution. In these simulations, patterns of genetic diversity was shown to depend on

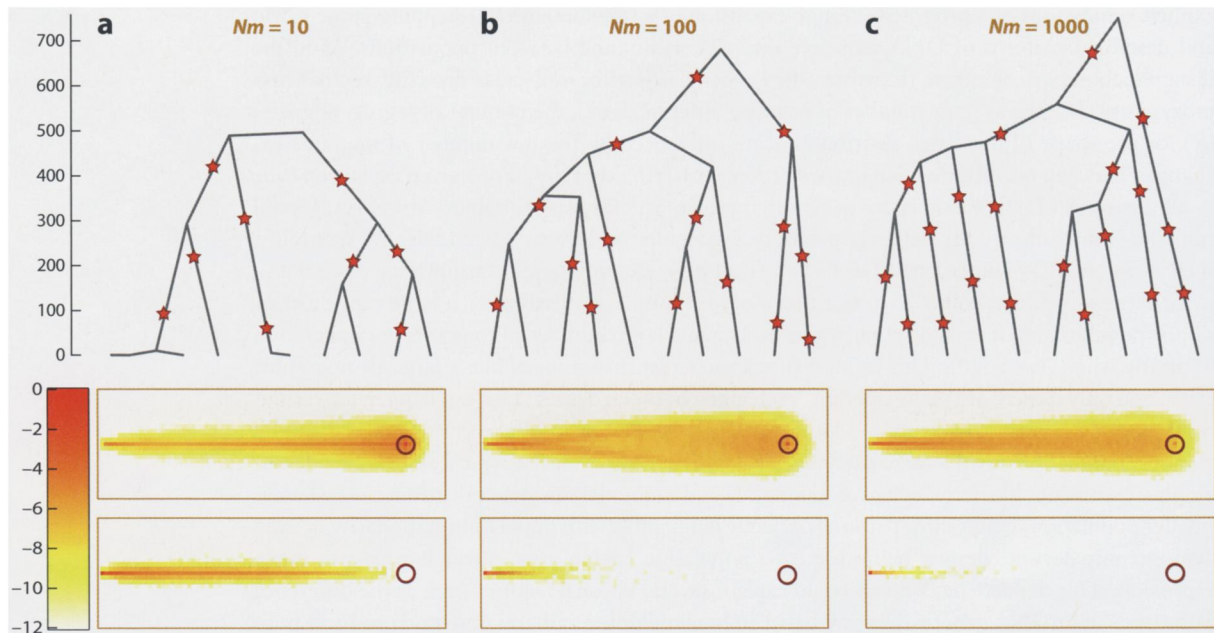


Figure 3

Distribution of coalescent events in a population that expanded recently. We simulated a spatial expansion (from the left to the right) in a world where demes are arranged on a lattice of 100×25 demes. The expansion starts from a deme located in the middle of the left side and takes about 500 generations to reach the right side. Demographic parameters: migration rate $m = 0.1$; growth rate $r = 0.8$; carrying capacity of all demes $K = N = 100$ (a), 1000 (b), and 10,000 (c). We sampled 10 genes in the deme located in the center of the circle on the right side of the world, and we show a random genealogy for each simulation condition on the top row. The empirical spatial density of coalescent events obtained from 100,000 simulations is shown on the central row, and the spatial distribution of the most recent common ancestor is on the bottom row. Note that all densities are color coded on a log scale.

the product $\bar{K}m$, where \bar{K} is the overall mean carrying capacity (Wegmann et al. 2006). For a given $\bar{K}m$ value, increasing environmental heterogeneity leads to lower levels of genetic diversity within a deme and increases the amount of differentiation between demes (F_{ST}). Higher environmental heterogeneity has, therefore, the same effect as a decrease of $\bar{K}m$ in homogeneous environments. This effect was found to be more pronounced for small \bar{K} values, implying that heterogeneity of the environment reduces genetic diversity more drastically in rare species than in abundant ones. The effect of temporal fluctuations of the environment was found to be less than that of spatial heterogeneity (Wegmann et al. 2006).

SPATIAL GENETIC PATTERNS AFTER A RANGE EXPANSION

Surfing

In a pioneering work, Edmonds et al. (2004) simulated the fate of new mutations arising on the edge of a wave of advance in two dimensions and showed that they would most often go extinct or stay at low frequencies around their place of origin. However, in a few cases, these new mutations could increase in frequency and be propagated by the wave, reaching very large frequencies and even fixation far away from their place of origin (see **Figure 4**). Thus, contrary to widespread beliefs, this suggests that the area with the highest frequency of a mutation is not necessarily its

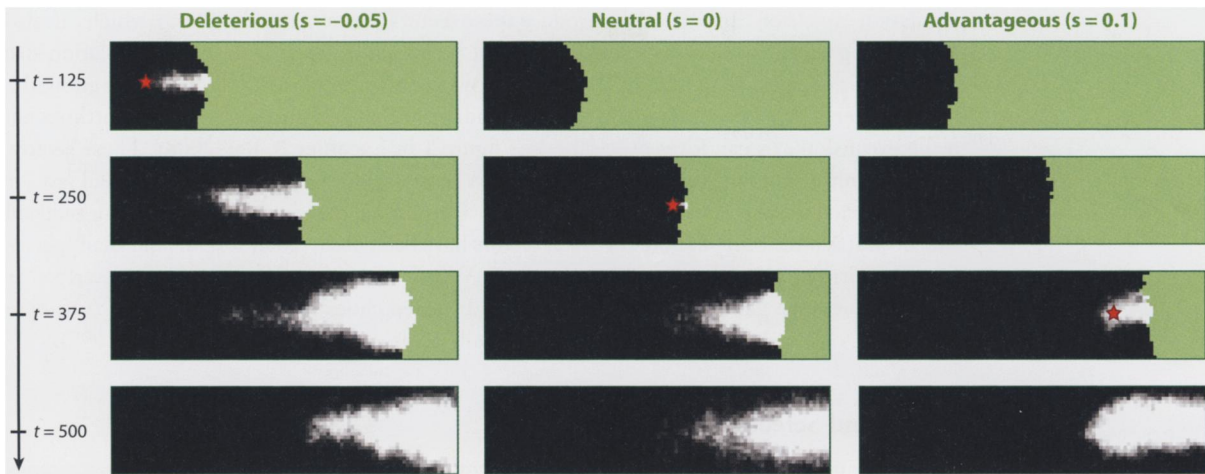


Figure 4

Illustration of the surfing of new mutations arising on the front of a wave of advance. The colonization of a world of 100×25 demes with K set to 100 starts from a single deme in the middle of the left border, and a wave of advance proceeds from left to right. At an arbitrary time t , a new mutation shown as a red star occurs on the edge of the wave, and it can surf on the expansion wave, reaching high frequencies (*lighter gray*) in newly colonized areas. We report here three random cases of successful surfing for a new mutation, that is either neutral, deleterious, or advantageous. The three types of mutations can surf, but their dynamics is quite different: deleterious mutations tend to stay close to the edge and are rapidly lost in the interior, whereas advantageous mutations can also travel back and propagate in the already occupied world. Note that despite different modes of selection, or different places and times of origin, the three types of mutations can show very similar spatial distributions at time $t = 500$.

place of origin. Klopstein et al. (2006) further examined this phenomenon they called surfing. They found that surfing occurs more often in small than in large populations and is favored in rapidly growing populations with limited dispersal. Also, when mutations surf in large populations (that is, when $K > 100$), they seldom reach very high frequencies, in contrast with mutations surfing on waves of small populations where the mutations can often be fixed in newly colonized areas. The surfing phenomenon is actually not restricted to new mutations occurring on the wave front, as it is also possible for standing variation. Actually, any gene present on the front can surf, but mutations or alleles being present right on the wave front have a much higher chance to surf than mutations occurring in already colonized areas (Klopstein et al. 2006, Travis et al. 2007). Mutations occurring on the extreme front of the wave have the highest chance to surf, but because these mutations are rare, the most likely area of origin of mutations having successfully surfed is located slightly away from the tip of the wave (Hallatschek & Nelson 2008). Finally, although frequent alleles are more likely to surf than rarer alleles, the surfing of rare alleles has more impact on the spatial distribution of allele frequencies and is, thus, more likely to be identified in studies of genetic differentiation (Hofer et al. 2008).

Sectors

Hallatschek et al. (2007) have provided the first experimental evidence of allele surfing. They monitored the growth of two fluorescently labeled strains of bacteria put in equal concentration in the middle of an agar plate. After about one day of growth, they observed the formation of stable sectors made up of a unique strain that were radiating from the center of the plate. These sectors consisted of microbes having originated from a single individual, which is fully analogous to the surfing of a new mutation occurring on the wave front. Computer simulations of a radiating

population expansion can perfectly reproduce these sectors (Excoffier & Ray 2008), which can also occur during expansions of species with higher dispersal abilities and with lower population size than bacteria or yeasts (Excoffier & Ray 2008). However, with increased interdeme migration, the limit between sectors becomes less clear, and gradients of allele frequencies that are orthogonal to the expansion axis can form over time (see figure 1 in Excoffier & Ray 2008). These sectors are, thus, only a transient structure in species with large dispersal abilities, but they could last for a longer period in species with limited motility or for genome components experiencing reduced gene flow. It is interesting to note that such sectors have been observed in several land snail species (Cain & Currey 1963, Goodhart 1963, Gould & Woodruff 1990), where they were described as “regions of relative genetic uniformity separated by steep clines in allele frequencies” (Ochman et al. 1983), or in oaks for seed-dispersed maternally transmitted markers (Petit et al. 1997).

Surfing and Selection

The surfing phenomenon is not restricted to neutral variants: Deleterious as well as advantageous mutations can surf. Travis et al. (2007) showed that even though deleterious mutations occurring on the wave front have less chance to survive than neutral mutations, the few surviving deleterious mutations are very likely to have surfed. In other words, surfing is the main cause of survival for deleterious mutations. As expected, the probability of survival and surfing increases with the selective advantage of new beneficial mutations. However, the fraction of surviving advantageous mutations having surfed is lower than for neutral mutations, because beneficial mutations can perfectly survive while remaining located close to their place of origin.

Unless they are recessive, deleterious mutations can only survive close to the wave front while surfing, and then should disappear once the colonization is over (**Figure 4**). Advantageous mutations surviving in a range expansion have usually a more complex spatial distribution, which can be bimodal, with a mode on the wave front, and another mode close to their place of origin from where they spread radially in all directions, as predicted by Fisher (1937). However, without any information on its age and origin, the spatial distribution of an allele does not seem sufficient to decide on the nature of the selection process. This is because different types of selection can lead to very similar spatial distributions of allele frequencies, as shown, for example, in **Figure 4** for $t = 500$.

Range expansions could be a propitious time for the spread and appearance of selected traits. Burton & Travis (2008a) have shown that multilocus maladapted genotypes could persist by surfing during range expansions, be propagated to new environments, and thus promote peak shift in the adaptive landscape. Additionally, dispersal itself could be favored during expansions (Travis & Dytham 2002). A good example of such a process has been described in the Australian cane toad, for which the current speed of invasion has increased fivefold since its original introduction, possibly owing to the selection of toads with longer legs moving faster at the wave front (Phillips et al. 2006). It would thus be worth investigating how much of current phenotypic diversity was selected during the expansion of successful species.

Gradients

Although gradients of allele frequencies are often interpreted as being determined by the environment (Endler 1973, Novembre et al. 2005, Ochman et al. 1987) or as resulting from the incomplete spread of favorably selected alleles (Livingstone 1989), early simulation studies of population expansions have confirmed that range expansions can create gradients of allele frequencies. These gradients arise either as a consequence of demic diffusion (Cavalli-Sforza et al.

1993, Rendine et al. 1986), by recurrent founder effects without admixture (Barbujani et al. 1995, Currat & Excoffier 2005), or by kin structured migrations of small groups (Fix 1997). However, significant gradients of allele frequencies are generally found for frequent alleles, which are more likely to surf (Klopfstein et al. 2006, Travis et al. 2007) and are rarely observed for low-frequency alleles (that is, when $p < 5\%$; Currat & Excoffier 2005). A recent analysis of the spatial patterns emerging from principal component analysis (PCA) of allele frequencies revealed that the first few PCA components consistently show clinal patterns in the absence of any population expansion (Novembre & Stephens 2008). This suggests that spatial PCA gradients can emerge under pure isolation by distance and not only during population expansions (Cavalli-Sforza et al. 1993) or through selection driven by environmental gradients (Piazza et al. 1981).

Spatial Constraints

The environment is never completely homogeneous. Hence, some colonization routes might be more likely than others owing to geographic or ecological barriers (Shigesada & Kawasaki 1997), the contours of the continents (Prugnolle et al. 2005, Ray et al. 2005), mountains (Travis et al. 2007), or ecovegetation types (Ackland et al. 2007, Ray et al. 2005). The effect of heterogeneous migration rates are analogous to the effects of heterogeneous population densities induced by the environment discussed above (Wegmann et al. 2006), because patterns of coalescence in structured populations depend on Nm values rather than on N or m separately. Therefore, one can expect that demes receiving fewer migrants than others will show lower levels of diversity (Wakeley 2001). The expected relationship between Nm and genetic diversity has been verified in the case of human populations showing opposite sex-biased dispersal (Hamilton et al. 2005b, Oota et al. 2001), and it can be intuitively understood by coalescent reasoning. Going backward in time, genes will migrate out of the sampled demes before they coalesce if Nm is large, but in geographically constrained environment, they will still coalesce early if they often migrate to the same location. Therefore, the probability of recent coalescent events is much larger in constrained than in large open environment, leading to overall lower genetic diversity.

The effect of range contractions or spatial bottlenecks on surfing has been further studied by simulations (Burton & Travis 2008b, Currat et al. 2006, Klopfstein et al. 2006, Ray et al. 2008, Travis et al. 2007), showing that coalescent events accumulate in and around regions of spatial bottlenecks (Burton & Travis 2008b, Travis et al. 2007, and see **Figure 5**), often corresponding to narrow migration corridors such as land bridges, rivers, or coastlines (Ray et al. 2008). Distributions of coalescent events during an expansion going through an artificial spatial bottleneck are shown in **Figure 5** for $Nm = 10$ and $Nm = 1000$. As compared to an absence of spatial bottleneck (**Figure 3**), there is a much higher density of coalescent events occurring within and on both sides of the spatial bottleneck, especially for $Nm = 10$. The distribution of MRCAs is also deeply affected, as 97% of the MRCAs now occur during the expansion for $Nm = 10$ (as compared to 66% without bottleneck in **Figure 3**), most of them being located around the spatial bottleneck. Spatial bottlenecks have less effect when $Nm = 1000$, as only 4% of the MRCAs occur during the expansion (as compared to less than 1% without bottleneck; **Figure 3**). These simulations suggest that spatial bottlenecks lead to a much larger probability of surfing during a range expansion, because all MRCAs not occurring in the deme at the origin of the expansion are indicative of surfing events. Interestingly, spatial constraints also increase the probability for deleterious mutations to surf (Burton & Travis 2008b). Another implication is that spatial bottleneck should highly structure phylogenetic trees, with a high density of coalescent events occurring at the time of the bottleneck, which might be (correctly) taken as evidence for past demographic contraction (Drummond et al. 2005). Moreover, spatial bottlenecks should similarly affect the phylogeny and the genetic diversity

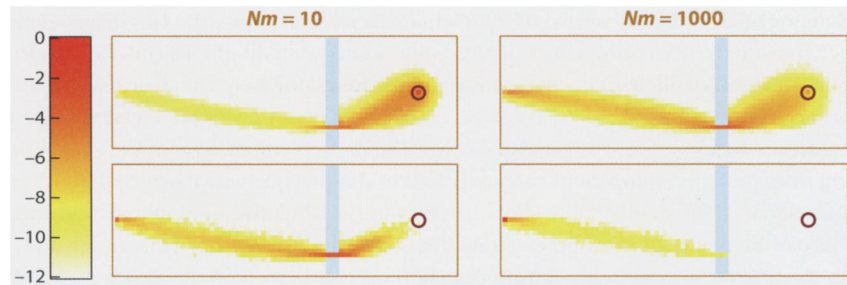


Figure 5

Distribution of coalescent events in a population that recently expanded through a spatial bottleneck. The simulation's conditions are similar to those of **Figure 3**, but there is an obstacle (shown in *pale blue*) on the path of the expansion, blocking the wave front, and obliging the wave to go through a small migration corridor in the lower part of the barrier. For each simulation we sampled 10 genes in the deme located in the center of the circle on the right side of the world. We report the empirical spatial density of coalescent events obtained from 100,000 simulations on the central row, and the spatial distribution of the most recent common ancestor on the bottom row. Note that all densities are color coded on a log scale.

of several independent genes, leading to drastic genetic differences between populations located on the two sides of the spatial bottleneck, which could be (wrongly) taken as evidence of adaptation to different environments (Excoffier & Ray 2008, Hofer et al. 2008).

Patches Generated by Long-Distance Dispersal

Very early on, Reid (1899) postulated that long-distance dispersal (LDD) events could account for the high speed of plant dispersal in open habitats after the last ice age. This intuition was subsequently confirmed using analytical and simulation approaches (Clark 1998, Kawasaki et al. 2006, Skellam 1951). LDD events also affect the genetic structure of an invading species, which typically shows a more patchy distribution of allele frequencies than species propagating only by short-range dispersal (Bialozyt et al. 2006, Ibrahim et al. 1996, LeCorre et al. 1997, Nichols & Hewitt 1994). A close examination of expansions with LDD reveals a nonmonotonic relationship between the frequency of LDD and the amount of genetic diversity preserved during colonization. Surprisingly, the occurrence of very rare LDD migration events leads to much lower levels of diversity than with short-range migrations (Bialozyt et al. 2006). This phenomenon, called embolism, is caused by the founding, well ahead of the wave front, of isolated populations having typically very little or no genetic diversity. The growth of these pioneer population generates a second and genetically uniform wave front, which is blocking the first wave (Bialozyt et al. 2006). As for surfing events, spatial constraints are important, and embolism is most likely to occur when colonization takes place in narrow corridors. With larger LDD levels, genetic diversity persists across the landscape because the homogenizing effect of surfing is prevented (Le Corre & Kremer 1998).

Phylogenetic Patterns Generated by Expansions

Despite the success of phylogeography (Avice et al. 1987), explicit and rigorous analyses of the phylogenetic structure expected after a range expansion are rare. Colonization routes are often reconstructed by simply mapping particular variants and by inferring the direction of the spread from other types of evidence, such as paleoenvironmental (Hewitt 1999, 2000; Taberlet et al. 1998)

or fossil data (Dumolin-Lapègue et al. 1997, Wijsman & Cavalli-Sforza 1984). Yet, coalescent considerations suggest that the most recently colonized populations should not only be characterized by increasingly divergent allelic frequencies but also by increasingly derived sequences due to mutations having accumulated on external branches. For instance, in humans, which are known to have experienced serial founder effects during their worldwide expansion, a recent large genomic study revealed a highly asymmetric population tree, and recently founded populations show more derived alleles (Li et al. 2008). Such molecular signatures of colonization remain to be further investigated and integrated into a rigorous inferential statistical framework. Although phylogenetic trees are not necessarily the best way to test evolutionary scenarios (Beaumont & Panchal 2008, Smouse 1998), they could help visualize patterns and infer past dynamics, as in cases of rapidly evolving viruses and other microorganisms (see e.g., Biek et al. 2007), where individuals sampled from successive epidemics form separate clades with branch lengths correlated with their time or place of sampling (Grenfell et al. 2004, Holmes 2004).

INTROGRESSION DURING A RANGE EXPANSION

Most range expansions do not occur in completely uninhabited areas, and the propagation of a wave of advance thus depends on intraspecific factors as well as on the patterns of interspecific interactions (Ackland et al. 2007, Aoki et al. 1996, Barton 2000, Rendine et al. 1986, Shigesada & Kawasaki 1997). Moreover, if interbreeding is possible between a local and an invading species, a moving hybrid zone develops during the expansion (Buggs 2007), and it will have quite different properties and consequences than static hybrid zones (e.g., Barton & Hewitt 1985).

Massive Introgression into the Invading Species

Simulation studies show that if the invading and the local species can interbreed and the fitness (γ) of the hybrids is not too low (that is, when $\gamma > 10\%$), the genome of the invading species should be massively introgressed by that of the local species, and the local species should show little sign of introgression (Currat & Excoffier 2004, Currat et al. 2008). When the fitness of the hybrids is very low (that is, when $0.1\% < \gamma < 10\%$), more complex patterns can emerge, like gradients of introgression (Currat et al. 2008, Rendine et al. 1986) or bidirectional introgression close to the area of introduction of the invasive species. Nevertheless, the expected massive and asymmetrical introgression patterns hold true for a wide range of demographic conditions, including when the invading species has a much higher density than the local species or when there is competition between the two species driving the local species to extinction (Currat & Excoffier 2004, Currat et al. 2008).

This relatively counterintuitive result about the asymmetry and extent of introgression is due to three separate effects illustrated in **Figure 6**. First, there is a simple progressive dilution of the invader gene pool owing to recurrent interbreeding with the local species, making the colonizers progressively more admixed along the expansion axis. A second source of asymmetry is the differential demography of the local and invading species in newly occupied territories: A successful invading species grows in newly occupied areas, whereas the local species either stays at its carrying capacity or declines when there is competition. This implies that a gene introgressing into the invading species on the wave front has a lower chance of getting lost by drift if the newly established population is growing fast. Finally, a gene introgressing on the wave front can occasionally surf in the invading population and, thus, reach very high frequencies in new colonized areas. Although the first effect simply describes the process of demic diffusion (Cavalli-Sforza et al. 1993), the two others are directly linked to the wave dynamics and have remained largely unnoticed so far.

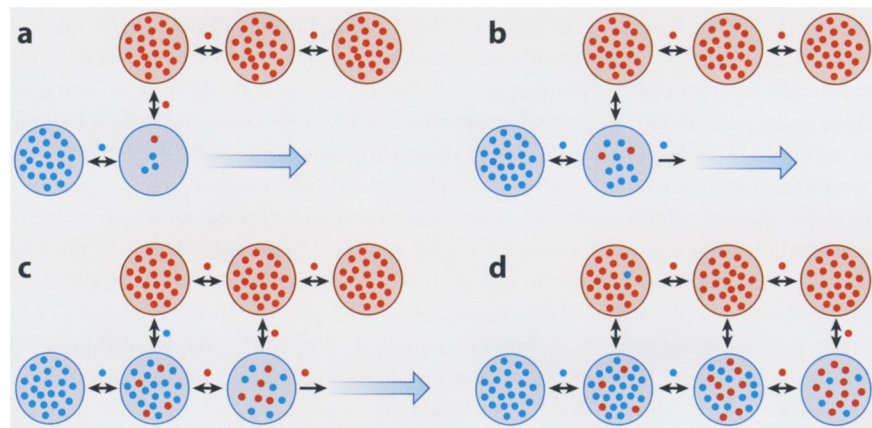


Figure 6

Illustration of the asymmetric introgression occurring during range expansions. We assume here that there is no competition between the two species. (a) A previously allopatric population (*blue*) invades the territory of a local population (*red*). The two populations can interbreed and a red gene introgresses the blue population on the wave front where the density is low. (b) The blue population grows logistically and the introgressed red gene is duplicated on the wave front. The blue population continues its progression into the red territory, sending mainly blue but also some red genes to unoccupied demes. (c) A new introgression occurs at the wave front, contributing to increasing the relative frequency of the red genes on the wave front, which now begins to send more red genes to unoccupied demes. Note that some blue genes can also introgress the red population in the wake of the wave where the blue population has reached its carrying capacity, but these blue genes remain at relatively low frequencies. (d) The colonization of new demes by red genes and further introgression events make the blue population more and more introgressed on the wave front. The continuation of this process should lead to a massive asymmetric introgression of the blue population by local red genes in newly colonized areas.

Simulations actually suggest that it is only when introgression events are too rare to occur on the wave front that introgression rates between the local and the invading species are symmetrical (yet very limited) (Currat et al. 2008).

The predicted dynamics of introgression can serve as a neutral model, allowing one to identify genome regions that could be under selection in the invading or in the local species or to identify mechanisms promoting or preventing introgression in one of the species. For instance, massive introgression into the local species is not expected according to this neutral model. If observed, some form of selection might be inferred—for instance, differential survival of the hybrids following Haldane's rule (Haldane 1922) or a selective advantage of the invader genes in the local background (e.g., Kim et al. 2008). In contrast, a lack of introgression into the invader genome could either indicate selection against the hybrids (Barton & Bengtsson 1986, Currat & Excoffier 2004, Gavrillets & Cruzan 1998) or selection against a local gene in the invader's genetic background. Such situations should be common because successful invading species have not been introgressed at genes controlling the phenotypic difference allowing them to invade. Hence, the examination of patterns of introgression at the genomic level should allow one to discover islands of differentiation where potential speciation genes could lie.

Empirical Examination of Introgression Patterns

An empirical survey of the literature confirmed the validity and generality of the asymmetrical introgression model (Currat et al. 2008). Among 44 cases of introgression that occurred after a

range expansion where the sequence of arrival of closely related species was known and where markers were available to test the direction of introgression, 82% of the studies showed massive introgression from the resident toward the invading species (Currat et al. 2008). This trend is also expected when the local species has completely disappeared and is well illustrated by the case of the Iberian hare that harbors, in Northern Spain, the mitochondrial genes from the mountain hare now extinct in Spain (Melo-Ferreira et al. 2007). Hence, when massive introgression of genes into an invading species is observed, it appears neither necessary nor parsimonious to invoke mechanisms such as selection (Payseur 2009, Rieseberg et al. 2007), disassortative mating (Wirtz 1999), or unequal species density (Cianchi et al. 2003).

Effect of Intraspecific Gene Flow

Simulations generally suggest that introgression levels in the invading species should be higher for genes characterized by limited levels of intraspecific gene flow (Currat & Excoffier 2004, Currat et al. 2008). This is because genes introgressing from the local species can be swamped out by genes migrating from the wake of the wave, when intraspecific migrations are high. The negative effect of gene flow on introgression might seem counterintuitive because it predicts that gene flow between species should be higher for genes with low levels of gene flow within species. However, it could explain why organelle genes often show higher levels of introgression than nuclear genes, owing to their generally lower Nm values (Currat et al. 2008). Two recent studies attempted to test the prediction that genome components experiencing less gene flow should introgress more, as a consequence of their different modes of inheritance and of sex-biased dispersal. Du et al. (2009) showed that in Conifers, maternally inherited mtDNA (subject to limited gene flow by seeds) introgresses more readily than paternally inherited chloroplast DNA (subject to high levels of gene flow, especially through wind-dispersed pollen). The pattern of introgression was also studied for markers with different modes of inheritance in wild-animal species with strong sex-biased dispersal (Petit & Excoffier 2009). In most cases, markers associated with the least dispersing sex showed more introgression than those associated with the most dispersing sex. In particular, nuclear markers are generally less introgressed than mtDNA markers in species with male-biased dispersal (18/19 cases; most mammals and some insect species), whereas the reverse is true in species with female-biased dispersal (16/16 cases; most bird and some insect species).

CONCLUSIONS

It is important to understand the effect of range expansions on the genetic diversity of populations because they have occurred in the recent history of most species. Range expansions generally lead to a loss of genetic diversity along the expansion axis owing to recurrent bottleneck effects. However, a large amount of migration between neighboring populations preserves genetic diversity and makes range expansions look more like large demographic expansions. Explicit consideration of spatial processes provides profound and often unexpected insights, revealing that the consequences of range expansions are conditioned by the properties and the dynamics of the wave front, a place where surfing events and asymmetric introgression originate. It is also important to realize that the patterns of genetic diversity generated by surfing are similar to those generated by selective sweeps and could, thus, be mistakenly interpreted as adaptive events (Currat et al. 2006). Like selection and unlike most other demographic effects, surfing generally does not affect all loci, and thus seems especially difficult to distinguish from directional selection. Asymmetrical introgression is another process that is often interpreted as a footprint of selection. Studies of range expansions have thus dramatically changed our view of what neutral diversity should look like, which should

allow us to better uncover selective and adaptive processes. Another logical step in this direction is to study the behavior of non-neutral variants during expansions. Preliminary studies indicate that, as a consequence of the often intense genetic drift taking place at the expanding species edge, both negatively and positively selected variants can surf. Expansion phases should therefore be considered as transition periods where genetic revolutions can take place.

FUTURE ISSUES

1. Integration of realistic spatial range expansions into a statistical inferential framework for the estimation of demographic, mutation, and selection parameters.
2. Disentangling signatures of spatial expansions from signatures of selection, and designing new ways to detect loci under selection.
3. Study the genetic consequences of past range contractions and range shifts, which are at least as prevalent as range expansions.
4. Study of the effect of surfing on nearby genomic diversity and on patterns of linkage disequilibrium.
5. Detect genes controlling phenotypic differences between invading and local interfertile populations from full genomes resequencing.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

M.F. was supported by a Swiss NSF grants, No. 3100-112072 and 3100-126074, to L.E.; and R.J.P. was supported by the network of excellence EVOLTREE. We are grateful to Nicolas Ray, Valérie Le Corre, Pauline Garnier-Géré and Gerald Heckel for discussions and their helpful comments on the manuscript.

LITERATURE CITED

- Ackland GJ, Signitzer M, Stratford K, Cohen MH. 2007. Cultural hitchhiking on the wave of advance of beneficial technologies. *Proc. Natl. Acad. Sci. USA* 104:8714–19
- Ammerman AJ, Cavalli-Sforza LL. 1973. A population model for the diffusion of early farming in Europe. In *The Explanation of Culture Change*, ed. C Renfrew, pp. 343–57. London: Duckworth
- Aoki K, Shida M, Shigesada N. 1996. Travelling wave solutions for the spread of farmers into a region occupied by hunter-gatherers. *Theor. Popul. Biol.* 50:1–17
- Austerlitz F, Jung-Muller B, Godelle B, Gouyon P-H. 1997. Evolution of coalescence times, genetic diversity and structure during colonization. *Theor. Popul. Biol.* 51:148–64
- Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, et al. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu. Rev. Ecol. Syst.* 18:489–522
- Baker H. 1948. Stages in invasion and replacement demonstrated by species of *Melandrium*. *J. Ecol.* 36:96–119
- Baker H, Stebbins G. 1965. *The Genetics of Colonizing Species*. New York: Academic
- Barbujani G, Sokal RR, Oden NL. 1995. Indo-European origins: a computer-simulation test of five hypotheses. *Am. J. Phys. Anthropol.* 96:109–32

- Barton N, Bengtsson BO. 1986. The barrier to genetic exchange between hybridising populations. *Heredity* 57(Pt 3):357–76
- Barton NH. 2000. Genetic hitchhiking. *Philos. Trans. R. Soc. London Ser. B* 355:1553–62
- Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* 16:113–48
- Beaumont MA, Panchal M. 2008. On the validity of nested clade phylogeographical analysis. *Mol. Ecol.* 17:2563–65
- Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genetics* 162:2025–35
- Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA* 98:4563–68
- Bialozyt R, Ziegenhagen B, Petit RJ. 2006. Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *J. Evol. Biol.* 19:12–20
- Biek R, Henderson JC, Waller LA, Rupprecht CE, Real LA. 2007. A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. *Proc. Natl. Acad. Sci. USA* 104:7993–98
- Biswas S, Akey JM. 2006. Genomic insights into positive selection. *Trends Genet.* 22:437–46
- Buggs RJ. 2007. Empirical study of hybrid zone movement. *Heredity* 99:301–12
- Burton OJ, Travis JM. 2008a. The frequency of fitness peak shifts is increased at expanding range margins due to mutation surfing. *Genetics* 179:941–50
- Burton OJ, Travis JM. 2008b. Landscape structure and boundary effects determine the fate of mutations occurring during range expansions. *Heredity* 101:329–40
- Cain AJ, Currey JD. 1963. Area effects in *Cepaea*. *Philos. Trans. R. Soc. London Ser. B* 246:1–81
- Cavalli-Sforza LL, Menozzi P, Piazza A. 1993. Demic expansions and human evolution. *Science* 259:639–46
- Cianchi R, Ungaro A, Marini M, Bullini L. 2003. Differential patterns of hybridization and introgression between the swallowtails *Papilio machaon* and *P. hospiton* from Sardinia and Corsica islands (Lepidoptera, Papilionidae). *Mol. Ecol.* 12:1461–71
- Clark JS. 1998. Why trees migrate so fast: confronting theory with dispersal biology and the paleorecord. *Am. Nat.* 152:204–24
- Critchfield W. 1984. *Impact of the Pleistocene on the genetic structure of North American conifers*. Presented at 8th North Am. For. Biol. Workshop, Utah State Univ., Logan
- Currat M, Excoffier L. 2004. Modern humans did not admix with neanderthals during their range expansion into Europe. *PLoS Biol.* 2:2264–74
- Currat M, Excoffier L. 2005. The effect of the Neolithic expansion on European molecular diversity. *Proc. R. Soc. London Ser. B* 272:679–88
- Currat M, Excoffier L, Maddison W, Otto SP, Ray N, 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* 313:172; author reply 72
- Currat M, Ray N, Excoffier L. 2004. SPLATCHE: a program to simulate genetic diversity taking into account environmental heterogeneity. *Mol. Ecol. Notes* 4:139–42
- Currat M, Ruedi M, Petit RJ, Excoffier L. 2008. The hidden side of invasions: massive introgression by local genes. *Evolution* 62:1908–20
- Cwynar LC, MacDonald GM. 1987. Geographical variation of lodgepole pine in relation to population history. *Am. Nat.* 129: 463–69
- Davis MB, Shaw RG. 2001. Range shifts and adaptive responses to quaternary climate change. *Science* 292:673–79
- Deshpande O, Batzoglou S, Feldman MW, Cavalli-Sforza LL. 2009. A serial founder effect model for human settlement out of Africa. *Proc. R. Soc. London Ser. B* 276:291–300
- Di Rienzo A, Wilson AC. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 88:1597–601
- Drummond AJ, Rambaut A, Shapiro B, Pybus OG. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* 22:1185–92
- Du F, Petit R, Liu J. 2009. More introgression with less gene flow: chloroplast versus mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other Conifers. *Mol. Ecol.* 18:1396–407

- Dumolin-Lapègue S, Demesure B, Fineschi S, Le Corre V, Petit RJ. 1997. Phylogeographic structure of white oaks throughout the European continent. *Genetics* 146:1475–87
- Edmonds CA, Lillie AS, Cavalli-Sforza LL. 2004. Mutations arising in the wave front of an expanding population. *Proc. Natl. Acad. Sci. USA* 101:975–79
- Endler JA. 1973. Gene flow and population differentiation. *Science* 179:243–50
- Excoffier L. 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Mol. Ecol.* 13:853–64
- Excoffier L, Ray N. 2008. Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.* 23:347–51
- Excoffier L, Schneider S. 1999. Why hunter-gatherer populations do not show sign of Pleistocene demographic expansions. *Proc. Natl. Acad. Sci. USA* 96:10597–602
- Fisher RA. 1937. The wave of advance of advantageous genes. *Ann. Eug.* 7:355–69
- Fix AG. 1997. Gene frequency clines produced by kin-structured founder effects. *Hum. Biol.* 69:663–73
- Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–25
- Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. *Genetics* 133:693–709
- Garcia-Ramos G, Rodriguez D. 2002. Evolutionary speed of species invasions. *Evolution* 56:661–8
- Gavrilets S, Cruzan MB. 1998. Neutral gene flow across single locus clines. *Evolution* 52:1277–84
- Goodhart CB. 1963. Area effects and non-adaptive variation between populations of *Cepaea* (Mollusca). *Heredity* 18:459–65
- Gould SJ, Woodruff DS. 1990. History as a cause of area effects—an illustration from *Cerion* on Great Inagua, Bahamas. *Biol. J. Linn. Soc.* 40:67–98
- Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, et al. 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303:327–32
- Gyllenstein U, Wilson A. 1987. Interspecific mitochondrial DNA transfer and the colonization of Scandinavia by mice. *Genet. Res.* 49:25–29
- Haldane J. 1922. Sex ratio and unisexual sterility in hybrid animals. *J. Genet.* 12:101–09
- Hallatschek O, Hersen P, Ramanathan S, Nelson DR. 2007. Genetic drift at expanding frontiers promotes gene segregation. *Proc. Natl. Acad. Sci. USA* 104:19926–30
- Hallatschek O, Nelson DR. 2008. Gene surfing in expanding populations. *Theor. Popul. Biol.* 73:158–70
- Hamilton G, Currat M, Ray N, Heckel G, Beaumont M, Excoffier L. 2005a. Bayesian estimation of recent migration rates after a spatial expansion. *Genetics* 170:409–17
- Hamilton G, Stoneking M, Excoffier L. 2005b. Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. *Proc. Natl. Acad. Sci. USA* 102:7476–80
- Hammer MF, Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, et al. 2001. Hierarchical patterns of global human Y-chromosome diversity. *Mol. Biol. Evol.* 18:1189–203
- Harpending H, Rogers A. 2000. Genetic perspectives on human origins and differentiation. *Annu. Rev. Genom. Hum. G.* 1:361–85
- Hengeveld R. 1988. Mechanisms of biological invasions. *J. Biogeogr.* 15:819–28
- Hewitt G. 1993. Postglacial distribution and species substructure: lessons from pollen, insects and hybrid zones. In *Evolutionary patterns and processes*, ed. D Lees, D Edwards, pp. 97–123. London: Academic
- Hewitt GM. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58:247–76
- Hewitt GM. 1999. Post-glacial re-colonization of European biota. *Biol. J. Linn. Soc.* 68:87–112
- Hewitt GM. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405:907–13
- Hofer T, Ray N, Wegmann D, Excoffier L. 2008. 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
- Holmes EC. 2004. The phylogeography of human viruses. *Mol. Ecol.* 13:745–56
- Hudson RR. 1990. Gene genealogies and the coalescent process. In *Oxford Surveys in Evolutionary Biology*, ed. DJ Futuyma, JD Antonovics, pp. 1–44. New York: Oxford Univ. Press
- Huntley B, Birks HJB. 1983. *An atlas of past and present pollen maps for Europe: 0–13000 B.P.* Cambridge: Cambridge Univ. Press. 667 pp.

- Ibrahim KM, Nichols RA, Hewitt GM. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* 77:282–91
- Kawasaki K, Takasu F, Caswell H, Shigesada N. 2006. How does stochasticity in colonization accelerate the speed of invasion in a cellular automaton model? *Ecol. Res.* 21:334–45
- Kim M, Cui ML, Cubas P, Gillies A, Lee K, et al. 2008. Regulatory genes control a key morphological and ecological trait transferred between species. *Science* 322:1116–9
- Kimura M. 1953. “Stepping stone” model of population. *Ann. Rep. Nat. Inst. Genet. Japan* 3:62–63
- Kimura M, Crow JF. 1963. Measurement of effective population number. *Evolution* 17:279–88
- Kimura M, Weiss WH. 1964. The stepping stone model of genetic structure and the decrease of genetic correlation with distance. *Genetics* 49:561–76
- Kinezaki N, Kawasaki K, Takasu F, Shigesada N. 2003. Modeling biological invasions into periodically fragmented environments. *Theor. Popul. Biol.* 64:291–302
- Kingman JFC. 1982. On the genealogy of large populations. *J. Appl. Prob.* 19A:27–43
- Klopfstein S, Currat M, Excoffier L. 2006. The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* 23:482–90
- Le Corre V, Kremer A. 1998. Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *J. Evol. Biol.* 11:495–512
- LeCorre V, Machon N, Petit RJ, Kremer A. 1997. Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. *Genet. Res.* 69:117–25
- Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, et al. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319:1100–04
- Li WH. 1977. Distribution of nucleotide differences between two randomly chosen cistrons in a finite population. *Genetics* 85:331–37
- Liu H, Prugnolle F, Manica A, Balloux F. 2006. A geographically explicit genetic model of worldwide human settlement history. *Am. J. Hum. Genet.* 79:230–7
- Livingstone FB. 1989. Simulation of the diffusion of the beta-globin variants in the Old World. *Hum. Biol.* 61:297–309
- Malécot G. 1948. *Les Mathématiques de l'Hérédité*. Paris: Masson
- Maruyama T. 1969. Genetic correlation in the stepping stone model with non-symmetrical migration rates. *J. Appl. Prob.* 6:463–77
- Maruyama T, Fuerst PA. 1984. Population bottlenecks and nonequilibrium models in population genetics. I. Allele numbers when populations evolve from zero variability. *Genetics* 108:745–63
- Maruyama T, Fuerst PA. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* 111:675–89
- Melo-Ferreira J, Boursot P, Randi E, Kryukov A, Suchentrunk F, 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.* 16:605–18
- Menozzi P, Piazza A, Cavalli-Sforza LL. 1978. Synthetic maps of human gene frequencies in Europeans. *Science* 201:786–92
- Nei M, Maruyama T, Chakraborty R. 1975. The bottleneck effect and genetic variability in populations. *Evolution* 29:1–10
- Nichols RA, Hewitt GM. 1994. The genetic consequences of long-distance dispersal during colonization. *Heredity* 72:312–17
- Novembre J, Glavani A, Slatkin M. 2005. The geographic spread of the CCR5 D32 HIV-resistance allele. *PLoS Biol.* 3:e339
- Novembre J, Stephens M. 2008. Interpreting principal component analyses of spatial population genetic variation. *Nat. Genet.* 40:646–9
- Ochman H, Jones JS, Selander RK. 1983. Molecular area effects in *Cepaea*. *Proc. Natl. Acad. Sci. USA* 80:4189–93
- Ochman H, Jones JS, Selander RK. 1987. Large scale patterns of genetic differentiation at enzyme loci in the land snails *Cepaea nemoralis* and *Cepaea hortensis*. *Heredity* 58(Pt 1):127–38
- Okubo A, Maini PK, Williamson MH, Murray JD. 1989. On the spatial spread of the grey squirrel in Britain. *Proc. R. Soc. London Ser. B* 238:113–25

- Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilineal versus patrilineal residence. *Nat. Genet.* 29:20–21
- Payseur BA. 2009. Y not introgress? Insights into the genetics of speciation in European rabbits. *Mol. Ecol.* 18:23–24
- Petit RJ, Excoffier L. 2009. Gene flow and species delimitation. *Trends Ecol. Evol.* 24:386–93
- Petit RJ, Pineau E, Demesure B, Bacilieri R, Ducouso A, Kremer A. 1997. Chloroplast DNA footprints of postglacial recolonization by oaks. *Proc. Natl. Acad. Sci. USA* 94:9996–10,001
- Phillips BL, Brown GP, Webb JK, Shine R. 2006. Invasion and the evolution of speed in toads. *Nature* 439:803
- Pialek J, Barton NH. 1997. The spread of an advantageous allele across a barrier: the effects of random drift and selection against heterozygotes. *Genetics* 145:493–504
- Piazza A, Menozzi P, Cavalli-Sforza LL. 1981. Synthetic gene frequency maps of man and selective effects of climate. *Proc. Natl. Acad. Sci. USA* 78:2638–42
- Pritchard JK, Przeworski M. 2001. Linkage disequilibrium in humans: models and data. *Am. J. Hum. Genet.* 69:1–14
- Prugnolle F, Manica A, Balloux F. 2005. Geography predicts neutral genetic diversity of human populations. *Curr. Biol.* 15: R159–60
- Ray N, Currat M, Berthier P, Excoffier L. 2005. Recovering the geographic origin of early modern humans by realistic and spatially explicit simulations. *Genome Res.* 15:1161–67
- Ray N, Currat M, Excoffier L. 2003. Intra-deme molecular diversity in spatially expanding populations. *Mol. Biol. Evol.* 20:76–86
- Ray N, Currat M, Excoffier L. 2008. Incorporating environmental heterogeneity in spatially-explicit simulations of human genetic diversity. In *Simulations, Genetics and Human Prehistory—A Focus on Islands*, ed. S Matsumura, P Forster, C Renfrew, pp. 103–17. Cambridge: McDonald Inst. Archaeol. Res.
- Reid C. 1899. *The Origin of the British Flora*. London: Dulau
- Remington C. 1968. Suture-zones of hybrid interaction between recently joined biotas. In *Evolutionary Biology*, ed. T Dobzhansky, M Hecht, W Steere, pp. 321–428. New York: Plenum Press
- Rendine S, Piazza A, Cavalli-Sforza LL. 1986. Simulation and separation by principal components of multiple demic expansions in Europe. *Am. Nat.* 128:681–706
- Rieseberg LH, Kim SC, Randell RA, Whitney KD, Gross BL, et al. 2007. Hybridization and the colonization of novel habitats by annual sunflowers. *Genetica* 129:149–65
- Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9:552–69
- Rousset F. 2007. Inferences from spatial population genetics. In *Handbook of Statistical Genetics*, ed. D Balding, M Bishop, C Cannings, pp. 945–79. Chichester, England: Wiley & Sons
- Schneider S, Excoffier L. 1999. Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152:1079–89
- Shigesada N, Kawasaki K. 1997. *Biological Invasions: Theory and Practice*. Oxford: Oxford Univ. Press. 205 pp.
- Shigesada N, Kawasaki K. 2002. Invasion and the range expansion of species: effects of long-distance dispersal. In *Dispersal Ecology. 42nd Symp. Br. Ecol. Soc.*, ed. JM Bullock, RE Kenward, RS Hails, pp. 350–73. Cambridge: Cambridge Univ. Press
- Shigesada N, Kawasaki K, Teramoto E. 1984. The effects of interference competition on stability, structure and invasion of a multi-species system. *J. Math. Biol.* 21:97–113
- Skellam JG. 1951. Random dispersal in theoretical populations. *Biometrika* 38:196–218
- Slatkin M. 1994. Linkage disequilibrium in growing and stable populations. *Genetics* 137:331–36
- Slatkin M. 2008. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genet.* 9:477–85
- Slatkin M, Hudson RR. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129:555–62
- Smouse PE. 1998. To tree or not to tree. *Mol. Ecol.* 7:399–412
- Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, et al. 2007. Putting the “landscape” in landscape genetics. *Heredity* 98:128–42

- Taberlet P, Cheddadi R. 2002. Ecology. Quaternary refugia and persistence of biodiversity. *Science* 297:2009–10
- Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* 7:453–64
- Tajima F. 1989a. The effect of change in population size on DNA polymorphism. *Genetics* 123:597–601
- Tajima F. 1989b. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–95
- Travis JM, Munkemüller T, Burton OJ, Best A, Dytham C, Johst K. 2007. Deleterious mutations can surf to high densities on the wave front of an expanding population. *Mol. Biol. Evol.* 24:2334–43
- Travis MJ, Dytham C. 2002. Dispersal evolution during invasions. *Evol. Ecol. Res.* 4:1119–29
- Wade MJ, McCauley DE. 1988. Extinction and recolonization—their effects on the genetic differentiation of local populations. *Evolution* 42:995–1005
- Wakeley J. 2001. The coalescent in an island model of population subdivision with variation among demes. *Theor. Popul. Biol.* 59:133–44
- Wakeley J. 2009. *Coalescent Theory: An Introduction*. Greenwood Village, Colorado: Roberts & Co. Publ.
- Watterson GA. 1978. The homozygosity test of neutrality. *Genetics* 88:405–17
- Watterson GA. 1986. The homozygosity test after a change in population size. *Genetics* 112:899–907
- Wegmann D, Currat M, Excoffier L. 2006. Molecular diversity after a range expansion in heterogeneous environments. *Genetics* 174:2009–20
- Whitlock MC, McCauley DE. 1990. Some population genetic consequences of colony formation and extinction—genetic correlations within founding groups. *Evolution* 44:1717–24
- Wijsman EM, Cavalli-Sforza LL. 1984. Migration and genetic population-structure with special reference to humans. *Annu. Rev. Ecol. Syst.* 15:279–301
- Wirtz P. 1999. Mother species-father species: unidirectional hybridization in animals with female choice. *Anim. Behav.* 58:1–12
- Wright S. 1943. Isolation by distance. *Genetics* 28:114–38