

Review

# Selecting among Alternative Scenarios of Human Evolution by Simulated Genetic Gradients

Catarina Branco and Miguel Arenas \* 

Department of Biochemistry, Genetics and Immunology, University of Vigo, 36310 Vigo, Spain; caraujo@uvigo.es

\* Correspondence: marenas@uvigo.es; Tel.: +34-986-130047

Received: 13 September 2018; Accepted: 16 October 2018; Published: 18 October 2018



**Abstract:** Selecting among alternative scenarios of human evolution is nowadays a common methodology to investigate the history of our species. This strategy is usually based on computer simulations of genetic data under different evolutionary scenarios, followed by a fitting of the simulated data with the real data. A recent trend in the investigation of ancestral evolutionary processes of modern humans is the application of genetic gradients as a measure of fitting, since evolutionary processes such as range expansions, range contractions, and population admixture (among others) can lead to different genetic gradients. In addition, this strategy allows the analysis of the genetic causes of the observed genetic gradients. Here, we review recent findings on the selection among alternative scenarios of human evolution based on simulated genetic gradients, including pros and cons. First, we describe common methodologies to simulate genetic gradients and apply them to select among alternative scenarios of human evolution. Next, we review previous studies on the influence of range expansions, population admixture, last glacial period, and migration with long-distance dispersal on genetic gradients for some regions of the world. Finally, we discuss this analytical approach, including technical limitations, required improvements, and advice. Although here we focus on human evolution, this approach could be extended to study other species.

**Keywords:** human genetic gradients; human evolution; model selection; range expansion; range contraction; last glacial maximum; long-distance dispersal; allele surfing

## 1. Introduction

The evolutionary history of our species persists as a hot topic of research due to the curiosity about our past and the continuous interesting findings from both genetic and archeological data, despite the fact that these findings are sometimes contradictory e.g., [1–3]. Indeed, knowledge about human genetic variation may help us to understand the causes and effects of some human diseases, like those presenting variable behaviour among ethnic groups or populations e.g., [4,5]. Conveniently, the genetic material of current humans still presents signatures of past evolutionary events, allowing us to investigate aspects about our origins. However, the interpretation of these findings is not always straightforward, because different evolutionary processes can lead to similar results. A clear example is the interpretation of the genetic gradients (clines) of modern humans by Cavalli-Sforza et al. [6–8]. These gradients were initially explained as genetic signatures of specific migrations. For example, Cavalli-Sforza et al. interpreted the European southeast–northwest (SE–NW) gradient of genetic variation as the result of the demic diffusion of early Neolithic farmers during their expansion from the Near East [9,10]. Posterior studies suggested that such genetic gradients could be caused or influenced by other processes such as range contractions or population admixture, i.e., hence, not necessarily attributed to a particular range expansion [11–15]. Interestingly, applying spatially-explicit computer simulations, François et al. [12] and Arenas et al. [14] showed that genetic gradients can present a direction perpendicular to the direction of the expansion as a consequence

of *allele surfing* [16,17], where mutations occurring on the wave of advance of the range expansion generate highly-differentiated genetic sectors aligned perpendicular to the direction of the expansion. Allele surfing is more detectable in recent expansions of small populations and under low migration rates, where sectors were not yet removed through homogenization [17]. Genetic gradients can also present the direction of the range expansion if the genetic signatures of allele surfing are lower than the genetic signatures of other genetic processes, such as isolation by distance (IBD). For example, Branco et al. [18] recently studied the influence of different evolutionary scenarios on American genetic gradients of modern humans through extensive spatially-explicit computer simulations. They found that at the continental level, the genetic gradient presented a direction following that of the range expansion under any studied evolutionary scenario (see Section 3), which was explained as IBD (similarly to the interpretations by Cavalli-Sforza et al. [7]), but this gradient varied when studied in smaller geographic regions, suggesting that the influence of different genetic processes on genetic gradients can vary with the geographic features of the landscape.

Since recent studies showed that genetic gradients can vary with different evolutionary processes, one can perform a selection among alternative evolutionary scenarios with data simulated under each scenario, followed by a fitting between simulated and real data based on genetic gradients. This strategy is not new in population genetics; for instance, the approximate Bayesian computation (ABC) approach [19,20] is frequently used to evaluate alternative scenarios of human evolution e.g., [21–25]. A goal of ABC is that it provides a quantitative evaluation of the fitting between real and simulated data; however, on the other hand, it usually requires a huge number of computer simulations (from many thousands to millions, although they can run in parallel) to obtain results with an acceptable level of accuracy and precision. Concerning the analyses based on a comparison between real and simulated genetic gradients, the most recent studies only required hundreds of simulations to identify the best fitting scenario, but these comparisons were mainly qualitative (direction of genetic gradients).

Here, we review the application of genetic gradients simulated under spatially-explicit computer simulations to distinguish between alternative evolutionary scenarios of modern humans by their fitting with real genetic gradients. First, we present the commonly-used methodologies to perform this selection among alternative scenarios, including the simulation of genetic data and estimation of genetic gradients. Next, we describe previous studies applying this approach to investigate the influence of human range expansions, range contractions followed by range re-expansions (processes that can be induced by glacial periods), population admixture and migration with long-distance dispersal, among others, on genetic gradients of some regions of the world, and to perform a selection among alternative evolutionary scenarios. Finally, we discuss advantages and limitations of those studies, and we provide recommendations based on our experience.

## 2. Simulation of Genetic Gradients

The simulation of genetic gradients usually consists of two main steps, namely: the simulation of genetic data under a given evolutionary scenario, and the estimation of the genetic gradient from the simulated data. Next, we describe the most frequently-used methodologies to perform both steps.

### 2.1. Simulation of Genetic Data under Diverse Evolutionary Scenarios of Human Evolution

A variety of approaches exist to simulate genetic data in population genetics, and they can be roughly classified in two types concerning the kind of simulation: (i) simulation of the evolutionary history of a sample, and (ii) simulation of genetic data upon a given evolutionary history.

Concerning the simulation of the evolutionary history of a sample, a number of approaches have been developed. The most commonly-used approaches are the coalescent [26], the birth-death approach [27], and the forward-time approach [28]. Basically, the coalescent simulates the evolutionary history of a sample of alleles from the present to the past until their most recent common ancestor (MRCA). The birth-death approach simulates the evolution of a sample considering birth and death

rates, which drive the amount of variability (branching) in the simulated history. By contrast, the forward-time approach simulates the evolution of a whole population from the past to the present. Despite the fact that the forward-time approach incorporates more evolutionary processes than the other approaches (i.e., interactions among individuals e.g., [29], population admixture e.g., [30], complex selection e.g., [29,31], and complex migration models e.g., [30,32,33]), computer simulations performed under this approach are computationally-intensive because of the simulation of many individuals (although progress is being made in this respect e.g., [34]). The simulation of the evolutionary history under a birth-death approach is much faster (similarly to the coalescent) but requires prior knowledge about birth and death rates. The coalescent is possibly the most commonly-implemented approach to be applied in population genetics (including studies on human evolution e.g., [23,35]), probably because of its rapid computation, its similarity with population genetics processes by modeling evolution based on the population size, and because it is capable of taking into account additional evolutionary processes such as demographics [36], recombination [37,38], population structure and migration [39–41], or selection e.g., [42–45]. Indeed, because of its rapid simulation and realistic population genetics modeling, the coalescent is a very useful approach when extensive simulations are required, for example in studies based on ABC or Bayesian approaches. For further details about approaches and frameworks to simulate evolutionary histories, we recommend the following reviews [46–48]. Interestingly, the forward-time and coalescent approaches were combined into the simulator *SPLATCHE*, allowing a rapid simulation of the evolutionary history of a sample accounting for evolutionary processes acting at the whole population level [49,50]. Basically, this framework simulates a spatial and temporal evolution of the whole population followed by the reconstruction of the evolutionary history of a given sample that is embedded in the previously-simulated population [50]; further details are shown later. This technical innovation made this framework well established in population genetics studies of terrestrial species, including humans [51].

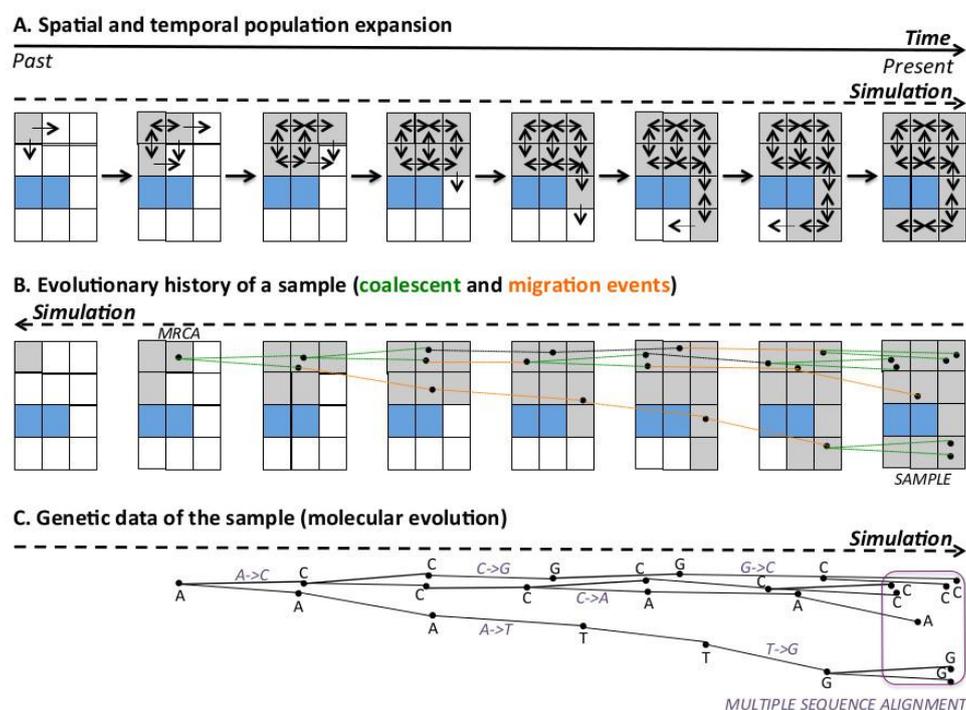
Once the evolutionary history of the sample is simulated (i.e., a simulated phylogenetic tree), it is possible to simulate molecular evolution upon such evolutionary history to obtain genetic sequences for all the internal and tip nodes (note that the set of simulated sequences of the tip nodes can compose a multiple sequence alignment) [46,52,53]. The traditional procedure is based on the following two steps: First, a genetic sequence (random or devised by the researcher) must be assigned to the MRCA node. Second, the MRCA sequence is evolved, from the past to the present, over the evolutionary history to obtain a sequence for every internal and tip (sample) node (an illustrative example is presented in the following subsection). The number of simulated substitutions depends on the branch length, while the type of simulated substitutions depends on the specified substitution model of evolution [46,52,54].

### Spatially-Explicit Computer Simulations

It is known that the consideration of a 2-dimensional (2D) landscape with its particular geography may result in simulations which are more realistic than those obtained with models of a lower number of dimensions [28,55]. This is because the real processes are often influenced by spatial constraints that may also vary over time, leading to the need for spatially-explicit models of evolution [56,57]. Despite some computer simulators implementing spatially-explicit models [49,50,58–60], unfortunately, several of them are not available to the public (i.e., the tool developed by Rendine et al. [61] applied to simulate an European Paleolithic and Neolithic expansion with admixture and the tool developed by Rasteiro et al. [30] applied to simulate human sex-biased migration). Other spatially-explicit computer simulators (i.e., *KERNELPOP* [59], *IBDSim* [60] and *CDMetaPOP* [62]) have not been yet widely applied to the study of human evolution, but they are potentially applicable for that purpose (see [28,51] for comparisons among different Spatially-explicit computer simulators). Next, the spatially-explicit computer simulator *SPLATCHE* [50] and its second version *SPLATCHE2* [49] have been largely used to study human evolution, perhaps because of their variety of implemented capabilities and their

graphical user interface. Hence, hereafter we focus on this simulator, which is the simulator used in the studies presented in the following sections of this review.

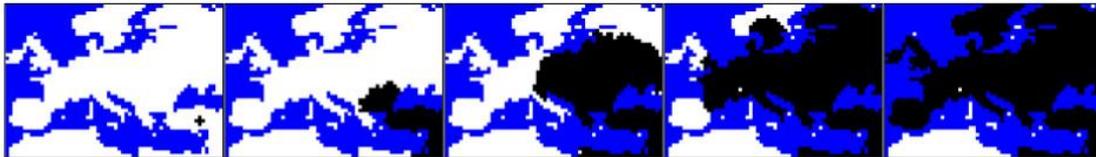
Spatially-explicit computer simulations with *SPLATCHE2* require a 2D landscape/map, which, for various regions of the world, can be imported from a Geographical Information System (GIS) [63]. This map can be split into a grid of small areas (demes) with a given deme size. *SPLATCHE2* simulates samples of genetic data by three main steps (Figure 1): (i) A forward-in-time simulation of the evolutionary history of the entire population accounting for spatial and demographic information (Figure 1A). Here, a deme must be chosen as a point of origin to start an expansion over the space and time. Next, migration events occur towards neighboring demes under the 2D stepping-stone migration model [64]. Each deme can be modeled with particular environmental conditions such as a particular carrying capacity (a measure of the resources available in the deme) and friction (capacity to move through the deme), and these parameters can vary over time to mimic periods with different resources. Next, the variation of the population size over time for each deme depends on the population growth rate and the specific environmental parameters of the deme. Indeed, the number of migration events from each deme depends on the migration rate and population size of the deme [50]. The simulation occurs during a user-specified number of generations that should be higher than the time to the MRCA (TMRCA) of the sample. An illustrative example of simulation of spatial and temporal expansion of European modern humans is shown in the Figure 2. The next steps consist of: (ii) the application of the coalescent to reconstruct the evolutionary history of a sample (which is embedded in the history of the entire population) (Figure 1B) and, (iii) a simulation of molecular (sequence) evolution over the evolutionary history of the sample to obtain genetic data for the sample (Figure 1C). *SPLATCHE2* can simulate genetic sequences with diverse molecular markers, including DNA, single nucleotide polymorphism (SNP), and short tandem repeat (STR).



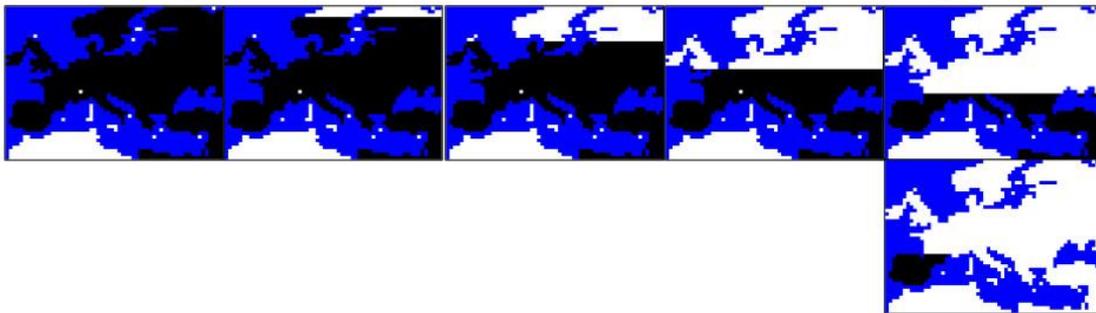
**Figure 1.** Illustrative example of a spatially-explicit simulation of a range expansion according to a 2-dimensional (2D) stepping-stone migration model [64], followed by the reconstruction of the evolutionary history of the sample and the simulation of genetic data. (A): Population range expansion, from the past to the present. It starts from the upper-left deme (origin), and migrants are sent to neighboring demes. Colonized demes (gray) can send/receive individuals to/from the neighboring demes, while non-colonized demes (white) can only receive individuals. We included a region representing a sea that cannot be colonized (blue), constituting a spatial barrier to migration.

(B): Reconstruction of the evolutionary history of a sample of 7 individuals (present). Going backwards in time, coalescence (green) and migration (orange) events occur until the most recent common ancestor (MRCA) of the sample is reached, which does not necessarily correspond to the origin (time and place) of the range expansion. (C): Simulation of genetic data for the sample. A random sequence (for simplicity, in this example, it is just 1 nucleotide, (A)) is evolved forward in time, incorporating substitutions along branches (violet), until reaching the sample (present). At the end of the simulation, a multiple sequence alignment is obtained by combining all the sequences of the sample. Note that the spatial barrier can affect the shape of the evolutionary history of the sample, and consequently, the genetic information of the sample.

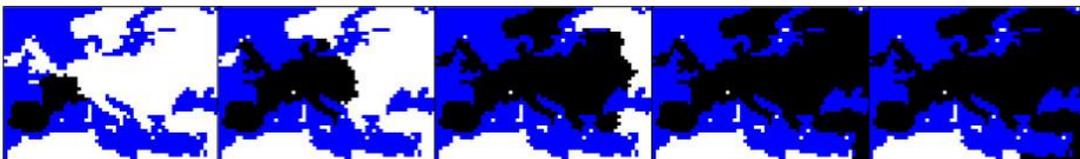
#### A. Paleolithic range expansion



#### B. Paleolithic range contraction towards the Iberian Peninsula induced by the last glacial period



#### C. Paleolithic range re-expansion from the Iberian Peninsula



**Figure 2.** Illustrative example of the simulation of spatial and temporal expansion, contraction, and re-expansion of Paleolithic Europeans. The figure presents snapshots obtained with the program *SPLATCHE2* for an example of: (A) simulation of a Paleolithic range expansion over Europe, (B) simulation of a Paleolithic range contraction towards the Iberian Peninsula induced by the last glacial maximum (LGM), and (C) simulation of a Paleolithic range re-expansion from the Iberian Peninsula after the LGM. To perform this simulation, we applied settings similar to those specified in [14]. Note that the time moves from the left to the right and the range expansion starts from the bottom-right corner of Europe (Middle East). Snapshots are taken each 50 generations. White demes indicate empty regions and black demes indicate colonized regions. Note that after this Paleolithic expansion, contraction and re-expansion, a Neolithic expansion (also from the Middle East) could be simulated with or without admixture with Paleolithic populations.

## 2.2. Estimation of Genetic Gradients in Studies of Human Evolution

Nowadays, several approaches allow the estimation of a genetic gradient from a dataset of genetic sequences. The traditionally-applied method to estimate genetic gradients is the principal component analysis (PCA). PCA identifies orthogonal axes (principal components, PCs) where objects show the highest variance of the information present in the original data. In population genetics, PCs provide an acceptable approximation of the covariance pattern among individuals of a given dataset [12]. They were largely used to study human evolution by Cavalli-Sforza [65], with the estimation of genetic gradients of European populations from allele-frequency data, and posteriorly used to estimate

genetic gradients for other worldwide human populations [6,7]. Nowadays, PCA remains a very useful and powerful technique to estimate genetic gradients [11,66] because it properly summarizes information present in large genetic data [67,68]. Recent studies that used PCA to obtain genetic gradients [12,14,18] applied the “*prcomp*” function of the *R* software environment. Indeed, the studies by Arenas et al. [14] and Branco et al. [18], which performed a high number of computer simulations per analyzed evolutionary scenario, estimated a genetic gradient for each simulated dataset. Next, they connected the geographical centroids of the positive and negative coordinates for every gradient to obtain a line representing the direction of the gradient. Finally, in order to summarize all the simulated genetic gradients obtained from each evolutionary scenario, they computed the median of the lines (slope and intercept) of the simulated gradients per scenario.

Recently, more complex methods have been developed to estimate genetic gradients and population structures for a given landscape [69–71]. These methods apply the Bayesian approach to infer genetic variation by modeling genetic distances between populations as a function of their geographic distance e.g., [72–74]. A limitation of these methodologies is that they may require long computer times to obtain convergence among MCMC chains, and, from our experience (unpublished), can generate artifacts when inferring genetic gradients in non-sampled regions (extrapolation of a genetic gradient). Rapid estimation with PCA is convenient for studies based on a high number of genetic datasets, like those presented in the following sections of this review involving many computer simulations. We believe that comprehensive comparisons of performance among the new Bayesian methods, and also including the PCA methods, should be investigated (for example by computer simulations).

### 3. Selecting among Alternative Scenarios of Population Admixture through Simulated Genetic Gradients

The European settlement by Paleolithic and Neolithic populations has been generally proposed with little admixture, yet studies still disagree. The estimated level of admixture varied with the applied genetic marker and the type of analyses performed, with Neolithic contributions below 25% [75], near 50% [76], and above 50% [9,77,78]. Assuming that the level of admixture could affect genetic gradients, a few studies investigated the amount of admixture by fitting genetic gradients simulated under different levels of admixture with the observed (real) genetic gradients obtained by Cavalli-Sforza et al. [6,7]. The gradients found by Cavalli-Sforza et al. present a SE-NW orientation, and were originally interpreted by these authors as a consequence of a demic diffusion process of Neolithic farmers from the Near East [6–8,79]. A first study exploring the influence of Paleolithic-Neolithic admixture through spatially-explicit computer simulations was performed by Currat and Excoffier [78]. They always found a gradient with a direction following the Neolithic expansion (SE-NW). Later, François et al. [12] repeated the study analyzing several levels of admixture. Under a high proportion (>20%) of Neolithic ancestry, they found a genetic gradient with a SW-NE direction, which is perpendicular to the direction of the range expansion from the Middle East. This gradient was interpreted as a consequence of allele surfing (see Introduction). Under lower levels of Neolithic ancestry, the gradient presented a direction following the direction of the range expansion (fitting with the gradient obtained from real data by Cavalli-Sforza et al.); this gradient was interpreted as a Paleolithic introgression along the direction of the Neolithic expansion. However, François et al. [12] only performed 10 simulations per studied evolutionary scenario and ignored some evolutionary processes such as range contractions induced by the last glacial maximum (LGM) period (Section 4). In a posterior study, Arenas et al. [14] repeated the analyses with more sophisticated evolutionary scenarios (including a variety of levels of Paleolithic-Neolithic admixture and range contractions modeling the effect of the LGM, as discussed in Section 4), and increased to 100 the number of simulations per studied evolutionary scenario. They verified that under high levels of Neolithic ancestry (>20%), the genetic gradients follow a direction perpendicular to the direction of the range expansion (allele surfing). By contrast, under low levels of Neolithic ancestry, the genetic

gradient followed the direction of the range expansion. However, they also found that the LGM could also affect the genetic gradients, leading to a more complex system that we present in Section 4.

The genetic gradient in the Americas based on real data follows the direction of the expansion (NW-SE) from Bering [6,7,80]. A similar gradient was also obtained from the analysis of the geographic distribution of linguistic families and subfamilies in this continent [7]. Recently, we and coauthors tried to investigate the level of admixture between the first Amerindian populations by applying computer simulations spatially [18]. We simulated two hypothetical Amerindian expansions from current Alaska: the first at 18 thousand years ago (kya) (ending the LGM) [81] and the second at 11 kya (beginning the Holocene) [82]. We investigated several levels of admixture between both populations, including a 100% and 0% contribution of the second population to the final genetic pool. We also simulated other evolutionary scenarios such as ice-sheets derived from the LGM and migration with long-distance dispersal (LDD) events; these are presented in Sections 4 and 5. The main finding was a simulated genetic gradient with a NW-SE direction throughout the entire continent, which was very similar to the genetic gradient obtained from real data. Importantly, we found that this genetic gradient was invariable with the level of population admixture. We interpreted this gradient as a consequence of IBD caused by the long NW-SE distance of the American continent, and where allele surfing could exist but in a lower extent. That result was for the analysis of the entire continent. Next, we separately analyzed North America to find, for any level of admixture, a gradient with direction NE-SW, perpendicular to the direction of the expansion, that we interpreted as a consequence of allele surfing. This gradient did not fit with the gradients derived from real data. However, as indicated above, we simulated additional evolutionary scenarios (LGM and LDD) to find that the gradient derived from real data in North America can be obtained only if those scenarios are considered (Sections 4 and 5). The findings suggested that genetic processes such as allele surfing, serial founder events, or IBD, which drive the direction of genetic gradients, can differ among the regions of a landscape.

#### 4. Selecting among Alternative Scenarios of Presence and Absence of the Last Glacial Period through Simulated Genetic Gradients

A factor that has been frequently ignored in interpretations of the SE-NW European genetic gradient is the last ice age that occurred at 29–13 kya [83]. During that period, European hunter-gatherer populations probably migrated towards the south through a range contraction, and next, re-expand north to recolonize the areas after the glacial period [84]. Arenas et al. [14] evaluated the influence of the last glacial period on the direction of the European genetic gradient. They performed spatially-explicit computer simulations under the following evolutionary scenarios: (i) absence of the last glacial period, (ii) presence of the last glacial period through the modeling of a range contraction towards all Southern Europe, followed by a period of time at refugia in all Southern Europe and a posterior re-expansion to recolonize the north and, (iii) presence of the last glacial period through the modeling of a range contraction towards only the Iberian Peninsula, followed by a period of time at refugia in only the Iberian Peninsula and a posterior re-expansion to recolonize the north (Figure 2). Note that the scenarios (ii) and (iii) present a different direction of range re-expansion: a re-expansion with direction S-N in (ii) and a re-expansion with direction SW-NE in (iii). The range contractions were simulated by a series of progressive contraction events during which demes located in the most northern areas became uninhabitable by setting its carrying capacity to zero [85,86]. In addition, the range contraction was simulated accounting for isotropic and anisotropic migration; the latter was designed to mimic humans who were aware about the glacial period, and that promotes a higher migration towards the south [84]. They found that both range re-expansions produced genetic gradients perpendicular to their direction: the re-expansion S-N led to a genetic gradient with direction E-W and the re-expansion SW-NE led to a genetic gradient with direction NW-SE, but only if the Paleolithic contribution to the final genetic pool was large enough (>80%). It is expected that the last glacial period affects genetic gradients only for large Paleolithic ancestry, because this period occurred during the Paleolithic. The simulated gradients were interpreted as a consequence of allele surfing derived from the range re-expansion, probably

because this expansion was recent. Altogether, Arenas et al. [14] found two evolutionary scenarios that fitted better with the real genetic gradients: (i) a scenario based on a large Paleolithic ancestry (>95%) and absence of any range contraction, and (ii) a scenario with some Paleolithic ancestry that considered a range contraction towards the Iberian Peninsula caused by the last ice period. In contrast, pure Neolithic expansions (without admixture and without genetic signatures from the last ice period) produced genetic gradients that did not fit with the genetic gradients estimated from real data.

Branco et al. [18] studied the influence of ice sheets caused by the last glacial period on the genetic gradients of the entire American continent and North America. It is known that as a consequence of the last glacial period, North America presented two large ice sheets (Laurentide and Cordilleran) that could have affected the entry and settlement of the first modern humans in this continent [87,88]. Concerning the entry to the Americas, two main routes have been proposed (and highly discussed): a coastal route through the North Pacific coastline, and an inland route (ice-free corridor) at the eastern side of the Rocky Mountains [89–91]. Indeed, the ice sheets could lead to temporary ice-free refugia in southern regions of North America and posterior expansions to colonize northern regions after melting [92]. In Branco et al. [18], we simulated the colonization of the entire continent and North America considering and ignoring ice sheets derived from the LGM [87]. Following previous works, scenarios with ice sheets were simulated by specifying carrying capacity of the demes covered by ice to zero [85] from 18 kya to 10 kya, a period that considers the duration of ice sheets, frozen grounds, and subsequent inundations [89]. Indeed, the coastal and inland corridors of entry into the Americas were simulated allowing a north to south passage without ice of 1–2 demes (100–200 km) width. At the entire continental level, we found that considering or ignoring the last glacial period does not alter the NW-SE genetic gradient, which was similar to that obtained from real data [6,7,80]. However, in North America, we found that the simulated genetic gradient in absence of the LGM presents a NE-SW direction (which does not fit with the real genetic gradient), while in presence of the LGM it presents a NW-SE direction (similar to the real genetic gradient). We concluded that at the continental level the NW-SE genetic gradient (which was invariable with population admixture and presence/absence of ice sheets) was mainly caused by a strong IBD, probably favored by the long north-south distance of this continent. However, in North America, the ice sheets must be considered to obtain the NW-SE gradient observed from real data. In addition, we also found that migration, including a proportion of long-distance dispersal (LDD) events, favors the simulation of the NW-SE genetic gradient (Section 5.3). Again, these findings suggest that the genetic processes driving the direction of genetic gradients can differ among regions of a landscape.

## 5. Selecting among Alternative Scenarios of Other Evolutionary Processes through Simulated Genetic Gradients

In addition to population admixture and the last glacial period, some other processes were investigated for testing their influence on genetic gradients. In this Section, we also briefly present the application of PC2 and PC3 to identify genetically isolated regions.

### 5.1. Influence of a Paleolithic Expansion from the Iberian Peninsula on the European Genetic Gradient

François et al. [12] investigated a European Paleolithic expansion from the Iberian Peninsula (instead of from the Middle East) followed by a Neolithic expansion from the Middle East. They found that if the simulated Paleolithic ancestry is large (>80%), scenarios with Paleolithic expansion from the Iberian Peninsula lead to genetic gradients similar to those from scenarios with Paleolithic expansion from the Middle East, suggesting that the origin of the Paleolithic expansion does not alter the SE-NW genetic gradient. Because of this, they concluded that the real genetic gradient (SE-NW) was caused by a Paleolithic introgression along the direction of Neolithic expansion instead of just by a Paleolithic range expansion from the Middle East.

### 5.2. Influence of Varying Evolutionary Parameters on Genetic Gradients

Arenas et al. [14] investigated the influence of several evolutionary parameters on European genetic gradients. They found that the genetic gradients were invariable to realistic changes of the ancestral population size, growth rate, and the carrying capacity of Neolithic populations (similar findings were found for American genetic gradients [18]). The only parameter that altered the gradient generated by the Neolithic population was the simulated number of generations. If the number of generations of the simulated Neolithic population is similar to the number of generations of the simulated Paleolithic population (both expansions starting at 40 kya), then the Neolithic population generates a gradient similar to that from the Paleolithic population, supporting the hypothesis that allele surfing could be the cause of the Neolithic genetic gradient (if the expansion is not recent, genetic sectors are lost by homogenization).

### 5.3. Influence of Long-Distance Dispersal on Genetic Gradients

Some studies suggested that the expansion of modern humans throughout the world could present LDD events, for example traveling by boats [93]. Actually, a recent study on the colonization of Eurasia by modern humans found that evolutionary scenarios based on LDD better fitted real data than evolutionary scenarios ignoring LDD [21]. Considering this aspect, in the study of American genetic gradients by Branco et al. [18], we investigated the influence of a proportion of migration through LDD on the genetic gradients. We performed spatially-explicit computer simulations under the LDD model developed by Ray and Excoffier [32], following a LDD distribution estimated from human data [94], a LDD proportion of 5% [21,33], and considering 1,000 km as a maximum distance of dispersal per generation [21]. We found that considering or ignoring LDD does not alter the NW-SE genetic gradient simulated along the entire continent (which is similar to the real genetic gradient [6,7,80]). This again supported the interpretation of strong genetic signatures caused by IBD along the entire American continent. However, in the specific analysis of North America, LDD generated the NW-SE genetic gradient (similar to the real gradient) if there is any genetic contribution from the first (more ancestral) population. This suggested that LDD events that occurred from the first population promoted a homogenization of genetic diversity [33] leading to the gradient that follows the longest geographic distance (an scenario of IBD), while LDD events in only the second expansion would require more time to obtain such homogenization. These findings suggest that LDD events could have occurred in the Americas from the first expansion, explaining the rapid colonization of this continent; this is in agreement with the presence of LDD in previous expansions throughout Eurasia [21].

### 5.4. Evolutionary Information from the Second and Third Principal Components

The first few PCs from a PCA are often used to explore the structure and variance of the data. In the analysis of a genetic sample, the first PC (PC1) map provides a spatial genetic gradient and the following PC maps (especially PC2 and PC3 because the amount of information of the original data is reduced by increasing the PC number) indicates genetically isolated regions. PC2 and PC3 maps were estimated in analyses of European populations [12,14] to highlight Scandinavia and the British Islands as genetically-isolated regions. Concerning the Americas, the inferred PC2 maps showed several regions with genetic isolation: Alaska, the Labrador Peninsula, Central America and Patagonia [18]. All these estimations were in agreement with the findings from real data [6,7,80] and the genetic isolation was mainly explained as a consequence of geographic isolation.

## 6. Conclusions and Future Prospects

Comparisons between simulated and real genetic gradients showed that spatially-explicit computer simulations provide good approximations of real processes and can be used to perform selection among alternative evolutionary scenarios. However, so far all the studies testing alternative scenarios of human evolution through simulated genetic gradients only performed qualitative

comparisons with real genetic gradients [12,14,18,78]. In those studies, the fitting between simulated and real gradients was performed just by a visual inspection of their overlapping, and we believe that it is likely that future studies could present situations requiring a quantitative evaluation. As previously indicated, Arenas et al. [14] and Branco et al. [18] performed a high number of computer simulations per studied evolutionary scenario; for each simulated dataset, they estimated a genetic gradient and obtained its direction by connecting the geographical centroids (positive and negative coordinates) to finally compute the median among all the simulated gradients of the evolutionary scenario. We believe that future studies could estimate, in addition to the median, the variance of the simulated gradient for each scenario, and these statistics could be used to perform a quantitative fitting between simulated and real genetic gradients (i.e., with ABC).

Another important aspect in this strategy, as in any analytical strategy based on computer simulations, is that the computer simulations should be as realistic as possible.

The studies discussed in previous sections analyzed genetic gradients of modern humans, ignoring some geographical barriers such as rivers and mountain ranges. We believe that these assumptions may not cause relevant biases when investigating large world regions (as done in such studies), but they could be crucial when investigating small regions. Moreover, the simulations should consider not only the current geographic landscape, but also its evolution from the beginning of the simulated evolutionary history (i.e., accounting for past vegetation maps [95]). Of course, some studies considered the last ice period [14,18], but still, the available resources may vary over time at any region of the landscape, and it was found that a temporal variation of environmental heterogeneity can induce a loss of genetic diversity within demes and increase the population differentiation among demes [96], which we believe could also affect genetic gradients.

Another way to generate more realistic computer simulations is by improving the modeling of human evolution. The aforementioned studies performed computer simulations based on evolutionary parameters (i.e., time and population size at the onset of the expansions, population growth rate, migration rate, LDD proportion, mutation rate, etc) estimated in previous works. However, the real processes were probably more complex, presenting multiple expansion waves and admixtures (e.g., in Europe the Roman and muslim expansions [97,98], or in the Americas, the admixture with non-American populations after the European contact [99,100]), complex demographics, where the population growth rate can vary over time (i.e., caused by population bottlenecks [101]), variation of migration rates over time (which could depend on the lifestyle and technology; for example it was found that Neolithic populations did not expand more rapidly than Paleolithic populations [102], perhaps because of their more sedentary lifestyle, or, as another example, the expansion throughout the Americas was faster than previous expansions throughout other regions [103]), or spatial and temporal selection [104]. Moreover, serial/longitudinal sampling should also be implemented in spatially-explicit computer simulators to analyse the increasing quantity of available ancient genetic data e.g., [105]. In all, the researcher is often forced to identify and apply only those parameters and capabilities implemented in the simulator that could better mimic a desired evolutionary scenario. Hopefully these complex processes will be incorporated into current and future spatially-explicit computer simulators.

In summary, simulated genetic gradients can be useful to perform selections among alternative evolutionary scenarios of modern humans, and we believe that they could also be applied to study other species with similar migration patterns. It is clear that the methods used so far can be improved, especially with more realistic computer simulations (based on high resolution maps and more realistic environmental and evolutionary conditions), and with the application of robust statistical methods for quantitatively evaluating the fitting between simulated and real genetic gradients. We believe that the application of genetic gradients for testing among alternative scenarios will increase in interest and use in the coming years.

**Funding:** This research was funded by [Ministerio de Economía y Competitividad of the Spanish Government] grant number [RYC-2015-18241].

**Acknowledgments:** We thank the Guest Editors of the special issue “Tools for Population and Evolutionary Genetics” for finding suitable our study for the special issue.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Novembre, J.; Stephens, M. Response to Cavalli-Sforza interview [Human Biology 82(3):245-266 (June 2010)]. *Hum. Biol.* **2010**, *82*, 469–470. [[CrossRef](#)] [[PubMed](#)]
2. Relethford, J.H. Genetic evidence and the modern human origins debate. *Heredity (Edinb.)* **2008**, *100*, 555–563. [[CrossRef](#)] [[PubMed](#)]
3. Lopez, S.; van Dorp, L.; Hellenthal, G. Human Dispersal Out of Africa: A Lasting Debate. *Evol. Bioinform. Online* **2015**, *11*, 57–68. [[CrossRef](#)] [[PubMed](#)]
4. Perez-Losada, M.; Posada, D.; Arenas, M.; Jobes, D.V.; Sinangil, F.; Berman, P.W.; Crandall, K.A. Ethnic differences in the adaptation rate of HIV gp120 from a vaccine trial. *Retrovirology* **2009**, *6*, 67. [[CrossRef](#)] [[PubMed](#)]
5. Wiencke, J.K. Impact of race/ethnicity on molecular pathways in human cancer. *Nat. Rev. Cancer* **2004**, *4*, 79–84. [[CrossRef](#)] [[PubMed](#)]
6. Cavalli-Sforza, L.L.; Menozzi, P.; Piazza, A. Demic expansions and human evolution. *Science* **1993**, *259*, 639–646. [[CrossRef](#)] [[PubMed](#)]
7. Cavalli-Sforza, L.L.; Menozzi, P.; Piazza, A. *The History and Geography of Human Genes*; Princeton University Press: Princeton, NJ, USA, 1994.
8. Piazza, A.; Rendine, S.; Minch, E.; Menozzi, P.; Mountain, J.; Cavalli-Sforza, L.L. Genetics and the origin of European languages. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5836–5840. [[CrossRef](#)] [[PubMed](#)]
9. Chikhi, L.; Nichols, R.A.; Barbujani, G.; Beaumont, M.A. Y genetic data support the Neolithic demic diffusion model. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11008–11013. [[CrossRef](#)] [[PubMed](#)]
10. Sokal, R.R.; Menozzi, P. Spatial Autocorrelations of HLA Frequencies in Europe Support Demic Diffusion of Early Farmers. *Am. Nat.* **1982**, *119*, 1–17. [[CrossRef](#)]
11. Novembre, J.; Stephens, M. Interpreting principal component analyses of spatial population genetic variation. *Nat. Genet.* **2008**, *40*, 646–649. [[CrossRef](#)] [[PubMed](#)]
12. François, O.; Currat, M.; Ray, N.; Han, E.; Excoffier, L.; Novembre, J. Principal component analysis under population genetic models of range expansion and admixture. *Mol. Biol. Evol.* **2010**, *27*, 1257–1268. [[CrossRef](#)] [[PubMed](#)]
13. McVean, G. A genealogical interpretation of principal components analysis. *PLoS Genet.* **2009**, *5*, e1000686. [[CrossRef](#)] [[PubMed](#)]
14. Arenas, M.; François, O.; Currat, M.; Ray, N.; Excoffier, L. Influence of admixture and paleolithic range contractions on current European diversity gradients. *Mol. Biol. Evol.* **2013**, *30*, 57–61. [[CrossRef](#)] [[PubMed](#)]
15. Reich, D.; Price, A.L.; Patterson, N. Principal component analysis of genetic data. *Nat. Genet.* **2008**, *40*, 491–492. [[CrossRef](#)] [[PubMed](#)]
16. Edmonds, C.A.; Lillie, A.S.; Cavalli-Sforza, L.L. Mutations arising in the wave front of an expanding population. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 975–979. [[CrossRef](#)] [[PubMed](#)]
17. Excoffier, L.; Ray, N. Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol. Evol.* **2008**, *23*, 347–351. [[CrossRef](#)] [[PubMed](#)]
18. Branco, C.; Velasco, M.; Benguigui, M.; Currat, M.; Ray, N.; Arenas, M. Consequences of diverse evolutionary processes on american genetic gradients of modern humans. *Heredity* **2018**, in press. [[CrossRef](#)] [[PubMed](#)]
19. Beaumont, M.A.; Zhang, W.; Balding, D.J. Approximate Bayesian computation in population genetics. *Genetics* **2002**, *162*, 2025–2035. [[PubMed](#)]
20. Beaumont, M.A. Approximate Bayesian computation in evolution and ecology. *Annu. Rev. Ecol. Evol. Syst.* **2010**, *41*, 379–405. [[CrossRef](#)]
21. Alves, I.; Arenas, M.; Currat, M.; Sramkova Hanulova, A.; Sousa, V.C.; Ray, N.; Excoffier, L. Long-distance dispersal shaped patterns of human genetic diversity in Eurasia. *Mol. Biol. Evol.* **2016**, *33*, 946–958. [[CrossRef](#)] [[PubMed](#)]
22. Pimenta, J.; Lopes, A.M.; Comas, D.; Amorim, A.; Arenas, M. Evaluating the Neolithic Expansion at Both Shores of the Mediterranean Sea. *Mol. Biol. Evol.* **2017**, *34*, 3232–3242. [[CrossRef](#)] [[PubMed](#)]

23. Fagundes, N.J.; Ray, N.; Beaumont, M.; Neuenschwander, S.; Salzano, F.M.; Bonatto, S.L.; Excoffier, L. Statistical evaluation of alternative models of human evolution. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 17614–17619. [[CrossRef](#)] [[PubMed](#)]
24. Ray, N.; Wegmann, D.; Fagundes, N.J.; Wang, S.; Ruiz-Linares, A.; Excoffier, L. A statistical evaluation of models for the initial settlement of the american continent emphasizes the importance of gene flow with Asia. *Mol. Biol. Evol.* **2010**, *27*, 337–345. [[CrossRef](#)] [[PubMed](#)]
25. Gamba, C.; Fernandez, E.; Tirado, M.; Deguilloux, M.F.; Pemonge, M.H.; Utrilla, P.; Edo, M.; Molist, M.; Rasteiro, R.; Chikhi, L.; et al. Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. *Mol. Ecol.* **2012**, *21*, 45–56. [[CrossRef](#)] [[PubMed](#)]
26. Kingman, J.F.C. The coalescent. *Stoch. Process. Appl.* **1982**, *13*, 235–248. [[CrossRef](#)]
27. Kendall, D.G. On the Generalized “Birth-and-Death” Process. *Ann. Math. Stat.* **1948**, *19*, 1–15. [[CrossRef](#)]
28. Epperson, B.K.; McRae, B.H.; Scribner, K.; Cushman, S.A.; Rosenberg, M.S.; Fortin, M.J.; James, P.M.; Murphy, M.; Manel, S.; Legendre, P.; et al. Utility of computer simulations in landscape genetics. *Mol. Ecol.* **2010**, *19*, 3549–3564. [[CrossRef](#)] [[PubMed](#)]
29. Peng, B.; Amos, C.I.; Kimmel, M. Forward-time simulations of human populations with complex diseases. *PLoS Genet.* **2007**, *3*, e47. [[CrossRef](#)] [[PubMed](#)]
30. Rasteiro, R.; Bouttier, P.A.; Sousa, V.C.; Chikhi, L. Investigating sex-biased migration during the Neolithic transition in Europe, using an explicit spatial simulation framework. *Proc. Biol. Sci.* **2012**, *279*, 2409–2416. [[CrossRef](#)] [[PubMed](#)]
31. Calafell, F.; Grigorenko, E.L.; Chikhanian, A.A.; Kidd, K.K. Haplotype evolution and linkage disequilibrium: A simulation study. *Hum. Hered.* **2001**, *51*, 85–96. [[CrossRef](#)] [[PubMed](#)]
32. Ray, N.; Excoffier, L. A first step towards inferring levels of long-distance dispersal during past expansions. *Mol. Ecol. Resour.* **2010**, *10*, 902–914. [[CrossRef](#)] [[PubMed](#)]
33. Mona, S.; Ray, N.; Arenas, M.; Excoffier, L. Genetic consequences of habitat fragmentation during a range expansion. *Heredity* **2014**, *112*, 291–299. [[CrossRef](#)] [[PubMed](#)]
34. Padhukasahasram, B.; Marjoram, P.; Wall, J.D.; Bustamante, C.D.; Nordborg, M. Exploring population genetic models with recombination using efficient forward-time simulations. *Genetics* **2008**, *178*, 2417–2427. [[CrossRef](#)] [[PubMed](#)]
35. Laval, G.; Patin, E.; Barreiro, L.B.; Quintana-Murci, L. Formulating a historical and demographic model of recent human evolution based on resequencing data from noncoding regions. *PLoS ONE* **2010**, *5*, e10284. [[CrossRef](#)] [[PubMed](#)]
36. Slatkin, M. Simulating genealogies of selected alleles in a population of variable size. *Genet. Res.* **2001**, *78*, 49–57. [[CrossRef](#)] [[PubMed](#)]
37. Hudson, R.R. Properties of a neutral allele model with intragenic recombination. *Theor. Popul. Biol.* **1983**, *23*, 183–201. [[CrossRef](#)]
38. Arenas, M. The importance and application of the ancestral recombination graph. *Front. Genet.* **2013**, *4*, 206. [[CrossRef](#)] [[PubMed](#)]
39. Hudson, R.R. Island models and the coalescent process. *Mol. Ecol.* **1998**, *7*, 413–418. [[CrossRef](#)]
40. Arenas, M.; Posada, D. Recodon: Coalescent simulation of coding DNA sequences with recombination, migration and demography. *BMC Bioinform.* **2007**, *8*, 458. [[CrossRef](#)] [[PubMed](#)]
41. Arenas, M.; Posada, D. Simulation of genome-wide evolution under heterogeneous substitution models and complex multispecies coalescent histories. *Mol. Biol. Evol.* **2014**, *31*, 1295–1301. [[CrossRef](#)] [[PubMed](#)]
42. Hudson, R.R.; Kaplan, N.L. The coalescent process in models with selection and recombination. *Genetics* **1988**, *120*, 831–840. [[PubMed](#)]
43. Arenas, M.; Posada, D. Coalescent simulation of intracodon recombination. *Genetics* **2010**, *184*, 429–437. [[CrossRef](#)] [[PubMed](#)]
44. Ewing, G.; Hermisson, J. MSMS: A coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics* **2010**, *26*, 2064–2065. [[CrossRef](#)] [[PubMed](#)]
45. Arenas, M. Applications of the Coalescent for the Evolutionary Analysis of Genetic Data. In *Reference Module in Life Sciences*; Elsevier: Amsterdam, The Netherlands, 2019; Volume 2, pp. 746–758.
46. Arenas, M. Simulation of Molecular Data under Diverse Evolutionary Scenarios. *PLoS Comput. Biol.* **2012**, *8*, e1002495. [[CrossRef](#)] [[PubMed](#)]

47. Hoban, S.; Bertorelle, G.; Gaggiotti, O.E. Computer simulations: Tools for population and evolutionary genetics. *Nat. Rev. Genet.* **2012**, *13*, 110–122. [[CrossRef](#)] [[PubMed](#)]
48. Arenas, M. Computer programs and methodologies for the simulation of DNA sequence data with recombination. *Front. Genet.* **2013**, *4*, 9. [[CrossRef](#)] [[PubMed](#)]
49. Ray, N.; Currat, M.; Foll, M.; Excoffier, L. SPLATCHE2: A spatially explicit simulation framework for complex demography, genetic admixture and recombination. *Bioinformatics* **2010**, *26*, 2993–2994. [[CrossRef](#)] [[PubMed](#)]
50. Currat, M.; Ray, N.; Excoffier, L. SPLATCHE: A program to simulate genetic diversity taking into account environmental heterogeneity. *Mol. Ecol. Notes* **2004**, *4*, 139–142. [[CrossRef](#)]
51. Benguigui, M.; Arenas, M. Spatial and temporal simulation of human evolution. Methods, frameworks and applications. *Curr. Genom.* **2014**, *15*, 245–255. [[CrossRef](#)] [[PubMed](#)]
52. Yang, Z. *Computational Molecular Evolution*; Oxford University Press: Oxford, UK, 2006.
53. Arenas, M.; Posada, D. Simulation of coding sequence evolution. In *Codon Evolution*; Cannarozzi, G.M., Schneider, A., Eds.; Oxford University Press: Oxford, UK, 2012; pp. 126–132.
54. Arenas, M. Trends in substitution models of molecular evolution. *Front. Genet.* **2015**, *6*, 319. [[CrossRef](#)] [[PubMed](#)]
55. Dunning, J.B.; Stewart, D.J.; Danielson, B.J.; Noon, B.R.; Root, T.L.; Lamberson, R.H.; Stevens, E.E. Spatially explicit population models: Current forms and future uses. *Ecol. Appl.* **1995**, *5*, 3–11. [[CrossRef](#)]
56. Excoffier, L.; Foll, M.; Petit, R.J. Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* **2009**, *40*, 481–501. [[CrossRef](#)]
57. Ray, N.; Excoffier, L. Inferring past demography using spatially explicit population genetic models. *Hum. Biol.* **2009**, *81*, 141–157. [[CrossRef](#)] [[PubMed](#)]
58. Landguth, E.L.; Cushman, S.A. CDPOP: A spatially explicit cost distance population genetics program. *Mol. Ecol. Resour.* **2010**, *10*, 156–161. [[CrossRef](#)] [[PubMed](#)]
59. Strand, A.E.; Niehaus, J.M. KERNELPOP, a spatially explicit population genetic simulation engine. *Mol. Ecol. Notes* **2007**, *7*, 969–973. [[CrossRef](#)]
60. Leblois, R.; Estoup, A.; Rousset, F. IBDSim: A computer program to simulate genotypic data under isolation by distance. *Mol. Ecol. Resour.* **2009**, *9*, 107–109. [[CrossRef](#)] [[PubMed](#)]
61. Rendine, S.; Piazza, A.; Cavalli-Sforza, L.L. Simulation and separation by principal components of multiple demic expansions in europe. *Am. Nat.* **1986**, *128*, 681–706. [[CrossRef](#)]
62. Landguth, E.L.; Bearlin, A.; Day, C.C.; Dunham, J. CDMetaPOP: An individual-based, eco-evolutionary model for spatially explicit simulation of landscape demogenetics. *Methods Ecol. Evol.* **2016**, *8*, 4–11. [[CrossRef](#)]
63. Leempoel, K.; Duruz, S.; Rochat, E.; Widmer, I.; Orozco-terWengel, P.; Joost, S. Simple rules for an efficient use of geographic information systems in molecular ecology. *Front. Ecol. Evol.* **2017**, *5*, 33. [[CrossRef](#)]
64. Kimura, M.; Weiss, G.H. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **1964**, *49*, 561–576. [[PubMed](#)]
65. Cavalli-Sforza, L.L. Population structure and human evolution. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **1966**, *164*, 362–379. [[CrossRef](#)]
66. Jakobsson, M.; Scholz, S.W.; Scheet, P.; Gibbs, J.R.; VanLiere, J.M.; Fung, H.C.; Szpiech, Z.A.; Degnan, J.H.; Wang, K.; Guerreiro, R.; et al. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* **2008**, *451*, 998–1003. [[CrossRef](#)] [[PubMed](#)]
67. Novembre, J.; Ramachandran, S. Perspectives on human population structure at the cusp of the sequencing era. *Annu. Rev. Genom. Hum. Genet.* **2011**, *12*, 245–274. [[CrossRef](#)] [[PubMed](#)]
68. Patterson, N.; Price, A.L.; Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2006**, *2*, e190. [[CrossRef](#)] [[PubMed](#)]
69. Petkova, D.; Novembre, J.; Stephens, M. Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.* **2016**, *48*, 94–100. [[CrossRef](#)] [[PubMed](#)]
70. Bradburd, G.S.; Ralph, P.L.; Coop, G.M. A spatial framework for understanding population structure and admixture. *PLoS Genet.* **2016**, *12*, e1005703. [[CrossRef](#)] [[PubMed](#)]
71. Duforet-Frebourg, N.; Blum, M.G. Nonstationary patterns of isolation-by-distance: Inferring measures of local genetic differentiation with Bayesian kriging. *Evolution* **2014**, *68*, 1110–1123. [[CrossRef](#)] [[PubMed](#)]

72. Messina, F.; Finocchio, A.; Akar, N.; Loutradis, A.; Michalodimitrakis, E.I.; Brdicka, R.; Jodice, C.; Novelletto, A. Spatially Explicit Models to Investigate Geographic Patterns in the Distribution of Forensic STRs: Application to the North-Eastern Mediterranean. *PLoS ONE* **2016**, *11*, e0167065. [[CrossRef](#)] [[PubMed](#)]
73. Jeong, C.; Peter, B.M.; Basnyat, B.; Neupane, M.; Beall, C.M.; Childs, G.; Craig, S.R.; Novembre, J.; Di Rienzo, A. A longitudinal cline characterizes the genetic structure of human populations in the Tibetan plateau. *PLoS ONE* **2017**, *12*, e0175885.
74. Uren, C.; Kim, M.; Martin, A.R.; Bobo, D.; Gignoux, C.R.; van Helden, P.D.; Moller, M.; Hoal, E.G.; Henn, B.M. Fine-Scale Human Population Structure in Southern Africa Reflects Ecogeographic Boundaries. *Genetics* **2016**, *204*, 303–314. [[CrossRef](#)] [[PubMed](#)]
75. Richards, M. The Neolithic Invasion of Europe. *Annu. Rev. Anthropol.* **2003**, *32*, 135–162. [[CrossRef](#)]
76. Lazaridis, I.; Patterson, N.; Mittnik, A.; Renaud, G.; Mallick, S.; Kirsanow, K.; Sudmant, P.H.; Schraiber, J.G.; Castellano, S.; Lipson, M.; et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **2014**, *513*, 409–413. [[CrossRef](#)] [[PubMed](#)]
77. Dupanloup, I.; Bertorelle, G.; Chikhi, L.; Barbujani, G. Estimating the impact of prehistoric admixture on the genome of Europeans. *Mol. Biol. Evol.* **2004**, *21*, 1361–1372. [[CrossRef](#)] [[PubMed](#)]
78. Currat, M.; Excoffier, L. The effect of the Neolithic expansion on European molecular diversity. *Proc. Biol. Sci.* **2005**, *272*, 679–688. [[CrossRef](#)] [[PubMed](#)]
79. Sokal, R.R.; Oden, N.L.; Wilson, C. Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature* **1991**, *351*, 143–145. [[CrossRef](#)] [[PubMed](#)]
80. Salas, A.; Lovo-Gomez, J.; Alvarez-Iglesias, V.; Cerezo, M.; Lareu, M.V.; Macaulay, V.; Richards, M.B.; Carracedo, A. Mitochondrial echoes of first settlement and genetic continuity in El Salvador. *PLoS ONE* **2009**, *4*, e6882. [[CrossRef](#)] [[PubMed](#)]
81. Dillehay, T.D. Probing deeper into first American studies. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 971–978. [[CrossRef](#)] [[PubMed](#)]
82. Forster, P.; Harding, R.; Torroni, A.; Bandelt, H.J. Origin and evolution of Native American mtDNA variation: A reappraisal. *Am. J. Hum. Genet.* **1996**, *59*, 935–945. [[PubMed](#)]
83. Straus, L.G. Southwestern Europe at the Last Glacial Maximum. *Curr. Anthropol.* **1991**, *32*, 189–199. [[CrossRef](#)]
84. Barbujani, G.; Bertorelle, G. Genetics and the population history of Europe. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 22–25. [[CrossRef](#)] [[PubMed](#)]
85. Arenas, M.; Ray, N.; Currat, M.; Excoffier, L. Consequences of range contractions and range shifts on molecular diversity. *Mol. Biol. Evol.* **2012**, *29*, 207–218. [[CrossRef](#)] [[PubMed](#)]
86. Arenas, M.; Mona, S.; Trochet, A.; Sramkova Hanulova, A.; Currat, M.; Ray, N.; Chikhi, L.; Rasteiro, R.; Schmeller, D.S.; Excoffier, L. The scaling of genetic diversity in a changing and fragmented world. In *Scaling in Ecology and Biodiversity Conservation*; Henle, K., Potts, S.G., Kunin, W.E., Matsinos, Y.G., Similä, J., Pantis, J.D., Grobelnik, V., Penev, L., Settele, J., Eds.; Pensoft Publishers: Sofia, Bulgaria, 2014; pp. 55–60.
87. Ray, N.; Adams, J.M. A GIS-Based Vegetation Map of the World at the Last Glacial Maximum (25,000–15,000 BP). *Internet Archaeol.* **2001**, *11*. [[CrossRef](#)]
88. Marshall, S.J.; James, T.S.; Clarke, G.K.C. North American Ice Sheet reconstructions at the Last Glacial Maximum. *Quat. Sci. Rev.* **2002**, *21*, 175–192. [[CrossRef](#)]
89. Bodner, M.; Perego, U.A.; Huber, G.; Fendt, L.; Rock, A.W.; Zimmermann, B.; Olivieri, A.; Gomez-Carballa, A.; Lancioni, H.; Angerhofer, N.; et al. Rapid coastal spread of First Americans: Novel insights from South America's Southern Cone mitochondrial genomes. *Genome Res.* **2012**, *22*, 811–820. [[CrossRef](#)] [[PubMed](#)]
90. Fagundes, N.J.; Kanitz, R.; Eckert, R.; Valls, A.C.; Bogo, M.R.; Salzano, F.M.; Smith, D.G.; Silva, W.A., Jr.; Zago, M.A.; Ribeiro-dos-Santos, A.K.; et al. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am. J. Hum. Genet.* **2008**, *82*, 583–592. [[CrossRef](#)] [[PubMed](#)]
91. Pedersen, M.W.; Ruter, A.; Schweger, C.; Friebe, H.; Staff, R.A.; Kjeldsen, K.K.; Mendoza, M.L.; Beaudoin, A.B.; Zutter, C.; Larsen, N.K.; et al. Postglacial viability and colonization in North America's ice-free corridor. *Nature* **2016**, *537*, 45–49. [[CrossRef](#)] [[PubMed](#)]
92. Rogers, R.A.; Rogers, L.A.; Hoffmann, R.S.; Martin, L.D. Native american biological diversity and the biogeographic influence of ice age refugia. *J. Biogeogr.* **1991**, *18*, 623–630. [[CrossRef](#)]
93. Balme, J. Of boats and string: The maritime colonisation of Australia. *Quat. Int.* **2013**, *285*, 68–75. [[CrossRef](#)]

94. Novembre, J.; Galvani, A.P.; Slatkin, M. The geographic spread of the CCR5 Delta32 HIV-resistance allele. *PLoS Biol.* **2005**, *3*, e339. [[CrossRef](#)] [[PubMed](#)]
95. Binney, H.; Edwards, M.; Macias-Fauria, M.; Lozhkin, A.; Anderson, P.; Kaplan, J.O.; Andreev, A.; Bezrukova, E.; Blyakharchuk, T.; Jankovska, V.; et al. Vegetation of Eurasia from the last glacial maximum to present: Key biogeographic patterns. *Quat. Sci. Rev.* **2017**, *157*, 80–97. [[CrossRef](#)]
96. Wegmann, D.; Currat, M.; Excoffier, L. Molecular diversity after a range expansion in heterogeneous environments. *Genetics* **2006**, *174*, 2009–2020. [[CrossRef](#)] [[PubMed](#)]
97. Zalloua, P.A.; Platt, D.E.; El Sibai, M.; Khalife, J.; Makhoul, N.; Haber, M.; Xue, Y.; Izaabel, H.; Bosch, E.; Adams, S.M.; et al. Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *Am. J. Hum. Genet.* **2008**, *83*, 633–642. [[CrossRef](#)] [[PubMed](#)]
98. Nebel, A.; Landau-Tasseron, E.; Filon, D.; Oppenheim, A.; Faerman, M. Genetic evidence for the expansion of Arabian tribes into the Southern Levant and North Africa. *Am. J. Hum. Genet.* **2002**, *70*, 1594–1596. [[CrossRef](#)] [[PubMed](#)]
99. Hunley, K.; Healy, M. The impact of founder effects, gene flow, and European admixture on native American genetic diversity. *Am. J. Phys. Anthropol.* **2011**, *146*, 530–538. [[CrossRef](#)] [[PubMed](#)]
100. Lindo, J.; Huerta-Sanchez, E.; Nakagome, S.; Rasmussen, M.; Petzelt, B.; Mitchell, J.; Cybulski, J.S.; Willerslev, E.; DeGiorgio, M.; Malhi, R.S. A time transect of exomes from a Native American population before and after European contact. *Nat. Commun.* **2016**, *7*, 13175. [[CrossRef](#)] [[PubMed](#)]
101. O’Fallon, B.D.; Fehren-Schmitz, L. Native Americans experienced a strong population bottleneck coincident with European contact. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20444–20448. [[CrossRef](#)] [[PubMed](#)]
102. Ammerman, A.J.; Cavalli-Sforza, L.L. *The Neolithic Transition and the Genetics of Populations in Europe*; Princeton University Press: Princeton, NJ, USA, 1984.
103. Regueiro, M.; Alvarez, J.; Rowold, D.; Herrera, R.J. On the origins, rapid expansion and genetic diversity of Native Americans from hunting-gatherers to agriculturalists. *Am. J. Phys. Anthropol.* **2013**, *150*, 333–348. [[CrossRef](#)] [[PubMed](#)]
104. Weaver, T.D.; Roseman, C.C. New developments in the genetic evidence for modern human origins. *Evol. Anthropol. Issues News Rev.* **2008**, *17*, 69–80. [[CrossRef](#)]
105. Schlebusch, C.M.; Malmstrom, H.; Gunther, T.; Sjodin, P.; Coutinho, A.; Edlund, H.; Munters, A.R.; Vicente, M.; Steyn, M.; Soodyall, H.; et al. Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science* **2017**, *358*, 652–655. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).