Genetic simulations of population interactions during past human expansions in Europe

Currat, Mathias; Excoffier, Laurent Georges Louis; Ray, Nicolas

Part II

What Simulations Tell Us About Complex European (Pre-)histories

Chapter 4

Genetic Simulations of Population Interactions during Past Human Expansions in Europe

Mathias Currat, Laurent Excoffier & Nicolas Ray

Two prehistoric demographic events profoundly affected human species in Europe. The first one is the replacement of Neanderthals by early modern humans from around 45,000 to 30,000 years ago and the second one corresponds to the Neolithic transition around 10,000 years ago. Both events entail to the arrival of new individuals coming from the East and the successive disappearance of pre-existing local people. In both cases the main question is how large was the contribution of pre-existing populations to the genetic pool of contemporary Europeans. In order to answer this question, we developed a new approach allowing the simulation of interactions between two populations (competition, admixture and assimilation) during a colonization phase and the resulting genetic structure of the remaining population. We applied this methodology to both the Neanderthal disappearance and the Neolithic transition by testing many different demographic scenarios. Our results showed that there were probably no genetic exchanges between Neanderthals and modern humans and that they therefore may have constituted two different species. We also showed no genetic evidence for a large replacement of European pre-existing hunter-gatherer populations during the Neolithic transition. Finally we point out the importance of taking in consideration the type of genetic markers used for the analysis as they can reveal very different patterns.

Range expansions and contractions, as well as interactions between populations, have been part of the complex history of humans. These demographic events played a very important role in shaping the genetic structure of populations. However, it is quite difficult to define the genetic signatures left by different demographic events mainly due to the confounding effects of demography and other evolutionary factors. If simple demographic events, such as population growth (i.e. King *et al.* 2000; Rogers & Harpending 1992; Slatkin & Hudson 1991; Tajima 1989), have been studied in detail, more realistic scenarios are difficult to handle analytically. Computer simulations offer an alternative to the analytical approach. We have therefore implemented a realistic and spatially explicit simulation framework to model past human demography and its impact on resulting genetic diversity (Currat *et al.* 2004). We have previously studied the genetic consequences of a range expansion in an empty area (Ray *et al.* 2003; 2005), but in the case of European prehistory, the two main range expansions occurred in areas that were already settled. Indeed, Europe was colonized by early modern humans between *c.* 45,000 and 30,000 years ago, at a time when Neanderthals already occupied the continent (e.g. Mellars 2004). At a later stage, between 10,000 and 5000 years ago, the change from hunting and gathering to farming techniques, the so-called Neolithic transition, occurred in an already settled European continent (e.g. Mazurié de Keroualin 2003). In both cases, new immigrants coming from the East spread toward the West and met pre-existing local people. The exact nature of these contacts is not well understood yet, but it seems that some competition and/or assimilation occurred as the pre-existing populations (Neanderthals and hunter-gatherers) disappeared in both cases. Genetic admixture might also have occurred between new incoming populations and local ones but the proportion of genetic exchanges

is still unknown. In the case of Neanderthals, it seems that admixture was very limited, as no Neanderthal lineages are observed in early and contemporary modern European populations (Krings *et al.* 1997; 1999; Ovchinnikov *et al.* 2000; Serre *et al.* 2004). In the Neolithic case, the level at which European hunter-gatherers adopted the new farming techniques coming from the Levant is still under investigation (Barbujani & Bertorelle 2001; Richards 2003; Richards *et al.* 2000). We describe here a simulation framework allowing the study of the genetic consequences of the invasion of an immigrant population into an area already occupied, by including competition and admixture. Using our method, we study the effects of interbreeding during the arrival of early modern humans and the Neolithic transition.

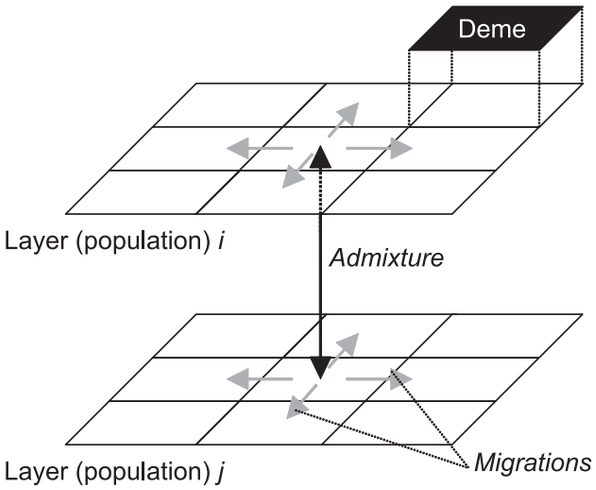

Figure 4.1. *Structure of the simulations: two layers of demes are superimposed, each one representing one population* i *or* j *(for example Neanderthals/Modern humans or Hunter-gatherers/Neolithic farmers). Each layer of demes is arranged as a 2D stepping-stone such that migrations can occur between the four adjacent demes at a rate* m. *Admixture can only occur between two demes from different populations that are present in the same cell.*

Methodology

In order to study the genetic consequences of the colonization of an occupied area by an invasive population, we modified the software SPLATCHE (Currat *et al.* 2004) by allowing the simulation of the demography of two interacting populations and their resulting genetic structure. This program first performs a forward demographic simulation followed by a backward simulation of genetic data based on a coalescent approach.

Simulating demography

A digital map of Europe was divided into a regular grid of 2500 km^2 cells (50 × 50 km). In each cell, there are two demes (population sub-units) representing the two populations of interest, which can be either Neanderthals (HN), modern humans (HS), hunter-gatherers (HG) or Neolithic farmers (NF) (Fig. 4.1). We thus simulated the demography of the subdivided populations in two superimposed layers of demes where movements were constrained by geographical contours. To each deme, we associated a given carrying capacity (K), which represents the maximum number of individuals that can subsist within the deme. K depends on the amount of resource of the deme and is identical for all demes belonging to the same population. However, values of K can be different for both populations, which represent the differences in their ability to exploit the resources of the environment.

Each generation, three demographic phases occur in each deme.

1. *Admixture:* gene flow can occur at a rate $A_{ij} = \gamma_{ij}(2N_iN_j)/(N_i + N_j)^2$ between two geographically identical demes from two different populations i and j, where N are the densities and γ_{ij} is an index of admixture. If $\gamma_{ij} = 0$, there is no genetic exchange between populations and if $\gamma_{ij} = 1$, reproduction is done at random between individuals from different populations (layers). An admixture event is supposed to lead to the birth of a child in population i whose parents originated from two different populations. Note that A_{ij} could result either by admixture (Neanderthal/Modern humans) or by assimilation (Hunter-gatherers/Farmers). Following admixture, population densities are then updated as $N_i' = N_i\,[1 - A_{ji}] + A_{ij}N_j$.

2. *Demographic regulation including density-dependent competition:* our model of density regulation incorporating competition is based on the Lotka-Volterra interspecific competition model, which is an extension of the logistic growth model (Lotka 1932; Volterra 1926). For each population, a new density N_i'' is calculated from the former density as $N_i'' = N_i'\left(1 = r_i(K_i - N_i' - \alpha_{ij}N_j'\right)/K_i$ where r_i is the intrinsic growth rate of the i-th population, K_i is its carrying capacity, and α_{ij} is an asymmetric competition coefficient. The rate of competition exerted by the population j over the population i is calculated as $\alpha_{ij} = N_j'/(N_i' + N_j')$, reflecting the fact that the influence of the members of a population on the other population increases with its density.

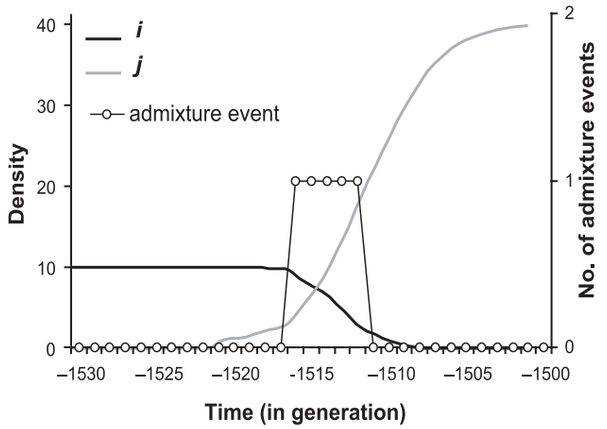

Figure 4.2. *Example of the evolution of two populations in competition (*i *and* j*) within a cell.* j *has a competitive edge over* i *due to its larger carrying capacity (*$K_j > K_i$*). In this example, the population* i *is already present in the cell at time –1530 (0 = present) with a density equal to* K_i *(10 effective genes). The cell is colonized by population* j *at time –1520. Then* j *replaces* i *gradually because of competition and admixture, the latter taking place during the cohabitation period with* $\gamma_{ij} = 0.4$*. In the current example, 5 effective genes migrate from deme* i *to deme* j *(thin black line with white circles).*

3. *Migration phase:* in each deme, individuals are sent outward at rate m and distributed equally among the four neighbouring demes. The density within the deme is thus regulated as $N_i''' = N_i''[1 - m] + I_i$, where I_i is the number of immigrants received from neighbouring demes. Figure 4.2 illustrates the evolution of population densities within a single cell. See Currat & Excoffier (2004; 2005) as well as Currat *et al.* (2004) for more details about this methodology.

Simulating genetic data

During the demographic simulation, all densities and migrations are stored into a data base in memory. These values (in number of effective genes) are then used to reconstruct the genealogy of a given number of genes sampled in the population, using a coalescent approach (Hudson 1990; Kingman 1982; Nordborg 2001). This genetic phase includes the following steps.

1. Sampled cells are chosen according to the geographical position of real samples. A number of genes n corresponding to the number of sampled genes in the real samples are placed in each of these cells.
2. Going backward in time generation after generation, genes can either migrate to adjacent demes or can coalesce with another gene located in the same deme at the same time. The probability of these two kinds of event (migration and coalescence) is calculated using the data base generated during the demographic phase. Each gene belonging to a deme i has a probability m_{ij} to migrate in the neighbouring deme j equal to $m_{ij} = I_{ji}/N_i$, where I_{ji} is the number of immigrants from deme j to deme i and N_i is the deme size. The probability of a coalescent event in a deme is equal to $n_i(n_i - 1)/(2N_i)$, where n_i is equal to the number of gene lineages present in deme i and N_i is the deme size in number of effective genes. This process continues until the MRCA (Most Recent Common Ancestor) of all sampled genes is reached.
3. Once the coalescent tree has been reconstructed, mutations are distributed over the branches of the tree in order to generate molecular diversity. The way these mutations are added on the coalescent tree depends on the mutation model, but various molecular markers can be generated: DNA sequences, Single Nucleotide Polymorphisms (SNPs), Restriction Fragment Length Polymorphisms (RFLPs) or Short Tandem Repeats (STRs).

Results and discussion

We present here the main results obtained using our approach concerning the two main human expansions in Europe, namely the arrival of early modern humans and the Neolithic transition.

Figure 4.3 shows the typical simulation of two consecutive expansion waves. We first simulated the arrival of early modern humans starting 40,000 years ago (1600 generations assuming a 25-year generation time) from the Near East (point P on Fig. 4.3) and the replacement of Neanderthals that were previously occupying the dark gray area in Figure 4.3 A and B. We then simulated a second expansion wave starting 10,000 years ago (400 generations) from the Near East (point N on Fig. 4.3). This latter expansion represents the diffusion of Neolithic techniques in Europe and the disappearance of the hunter-gatherer way of subsistence. During both expansions, we vary the admixture index γ between the populations. We simulated scenarios as realistic as possible using parameter values drawn from the literature. As there is a large uncertainty for the values of some parameters, we tried many different scenarios by varying demographic parameters such as growth and migration rates, carrying capacities, initial population sizes or expansion starting points. We only present and discuss here the main results, but details about the various tested scenarios can be found elsewhere (Currat & Excoffier 2004; 2005).

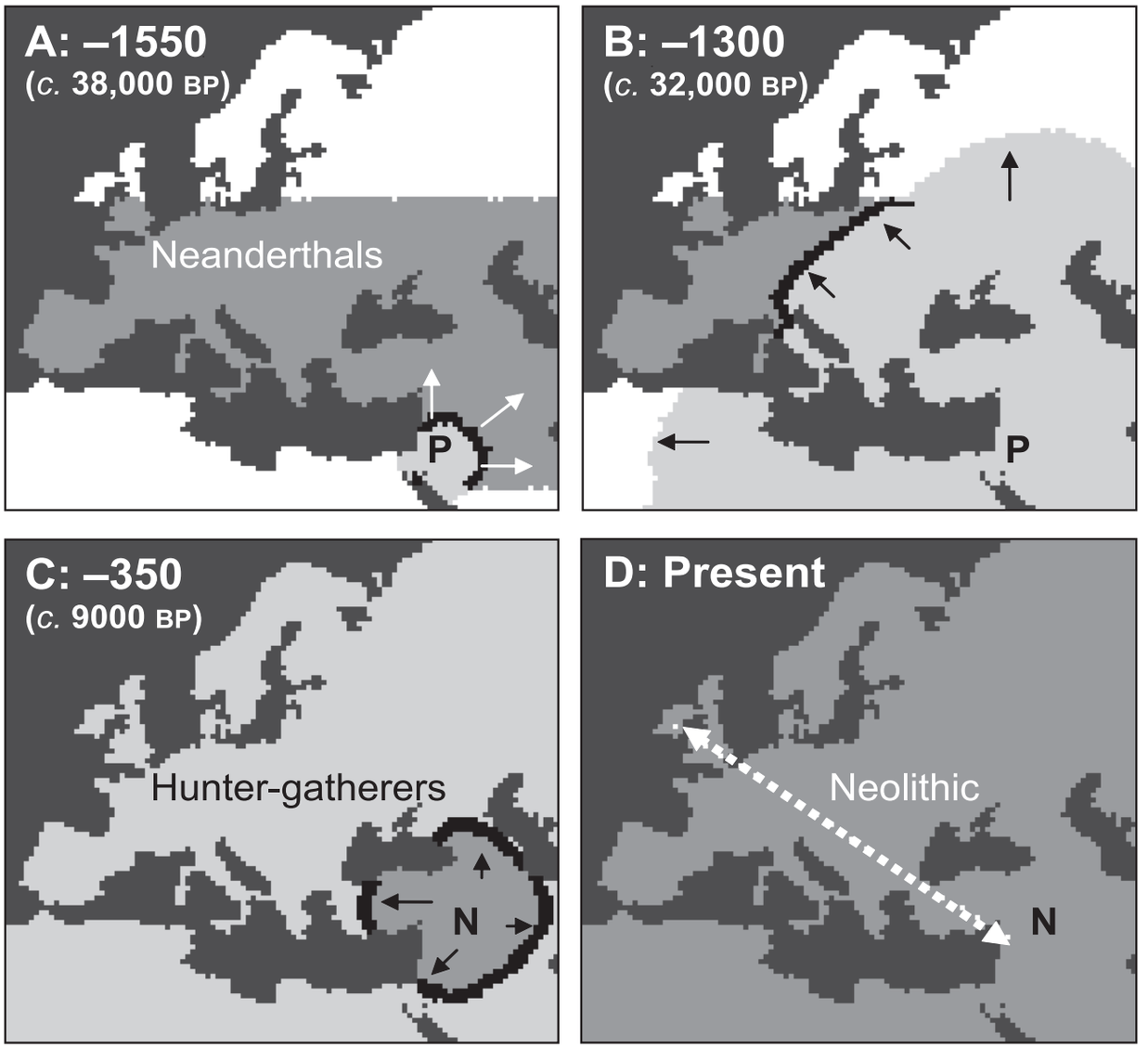

Figure 4.3. *Example of the simulation of the Palaeolithic expansion of early modern humans (in light grey) in Europe from the Near East and the replacement of Neanderthals (dark grey on panels A and B). P is the starting point for the early modern human expansion. The Neolithic expansion (in dark grey on panels C and D) starts 400 generations before the present from the origin N and expands over the Palaeolithic/Mesolithic substrata. During each invasion, competition occurs between the two populations, and the invaded population (Neanderthals then Hunter-gatherers) disappears. The black bands represent the cohabitation zone where admixture can potentially occur if $\gamma > 0$. The dashed arrow on panel D represents the axis along which 20 genetic samples are simulated.*

Introgression

The main phenomenon revealed by our simulations is a progressive introgression of the lineages of the invaded population into the genome of the invasive populations if admixture occurs. As shown in Figure 4.4, even a minute amount of interbreeding between the two populations is sufficient for the lineages of the invaded population to be strongly represented in the final gene pool. For example, an admixture index γ as small as 10 per cent between Mesolithic hunter-gatherers (HG) and the Neolithic farmers (NF) is enough for the European gene pool to be composed of only 30 per cent of genes descended from the initial Neolithic population from the Near East (what we will call 'Neolithic genes'). This γ value of 10 per cent corresponds to the incorporation into the Neolithic

population of only two HG lineages, on average per deme during the whole cohabitation period between Mesolithic and Neolithic populations (lasting between 140 and 200 years per cell).

The genetic consequences of this introgression phenomenon during the two main European expansions could be the following ones:

- If admixture occurs between Neanderthals (HN) and early modern humans (HS), then HN lineages are still expected to be present in the current European gene pool. The sequencing of Neanderthal mtDNA (Krings *et al.* 1997; 1999; Ovchinnikov *et al.* 2000; Schmitz *et al.* 2002; Serre *et al.* 2004) has revealed a specific type of sequence unambiguously distinct from any modern human type. Given the fact that no sequence of Neanderthal type is currently observed among 12,000 European mtDNA (Peter Forster pers. comm.; see also Handt *et al.* 1998; Richards *et al.* 1996), we can expect the genetic input of Neanderthals into modern humans to be very low, if not absent. However, previous analytical methods have not allowed one to exclude an initial Neanderthal contribution smaller than 25 per cent to the current European genetic pool (Nordborg 1998; Serre *et al.* 2004), as all initial HN lineages would have potentially disappeared by drift since the admixture event (Hagelberg 2003; Relethford 2001). These previous estimations were made using very simple demographic models where population sizes were constant or exponential within a unique deme and admixture was instantaneous. We used a much more realistic model that took population subdivisions into account as well as progressive admixture and competition between HS and HN, and we obtained a maximum estimate of the HN contribution to the European gene pool equal to 0.1 per cent (Currat & Excoffier 2004). This value is 250 times smaller than the previous estimates obtained with less realistic models (Nordborg 1998; Serre *et al.* 2004). Our calculations imply that Neanderthals and early modern humans probably did not interbreed. Whether this absence of genetic exchange was due to inter-sterility, selection against hybrids, or an absence of ecological competition remains to be demonstrated, but a cultural separation or absence of contacts due to low densities seems to be excluded since technological exchanges between Neanderthals and Cro-Magnons are documented (e.g. Hublin *et al.* 1996).
- Concerning the Neolithic transition, it has been postulated that a gradient of allele frequency observed from southeast to northwest Europe was the result of the dispersion of the lineages of the first farmers from the Near East (the 'Neolithic lineages') into Europe (Barbujani & Pilastro 1993; Chikhi *et al.* 1998; Menozzi *et al.* 1978; Rosser *et al.* 2000; Sokal *et al.* 1991). We estimated that in order to get a gradient of 'Neolithic lineages' over the whole Europe, the γ index between HG and NF has to be smaller than 10 per cent. Such small γ values correspond to less than 2 HG lineages incorporated in the communities of dispersing farmers at each location (cell). The γ index must even be smaller than 5 per cent in order to get a proportion of 'Neolithic lineages' larger than 50 per cent over the whole Europe, as estimated for certain markers (Barbujani & Dupanloup 2002; Chikhi 2002; Dupanloup *et al.* 2004). This means that only an extremely low admixture between HG and NF is compatible with an European gene pool constituted of a majority of lineages brought by early farmers during the Neolithic (less than one HG lineage incorporated into the Neolithic population on average per deme: see Fig. 4.4).

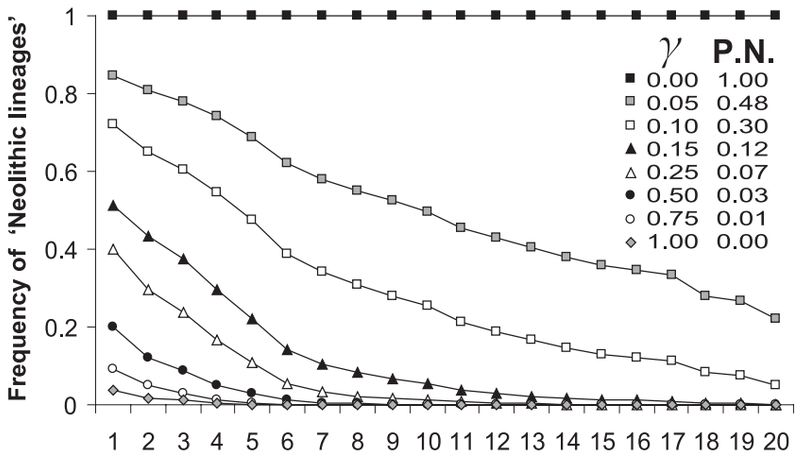

Figure 4.4. *Proportion of Near Eastern 'Neolithic genes' (see text) in a series of samples distributed regularly along the axis between Lebanon and Ireland (see arrow on Fig. 4.3D). These proportions are given for various admixture indices γ ranging from 0 to one. The average proportion of these 'Neolithic genes' over all samples on the transect is also reported as P.N.*

Two processes can jointly explain a rapid introgression of genes from the invaded population into the genome of the invasive population. First, at each step of its expansion, the invasive population incorporates

a given proportion of genes of the invaded population. The proportion of genes of the initial population in the wave front therefore decreases generation after generation, such that the larger the number of migration steps involved during the invasive processes, the stronger the dilution phenomenon. Second, the invaded lineages are incorporated into the invasive population during the cohabitation period, when the density of invaders is still low (see Fig. 4.2). It means that invaded lineages are included into the subsequent local population growth and are then amplified by the phase of logistic growth.

Molecular diversity

As the Neanderthal contribution to the current European gene pool was estimated to be almost zero (Currat & Excoffier 2004), we did not take Neanderthal lineages into account when studying the Neolithic transition. We looked at the following molecular signatures characteristic of range expansions: unimodal mismatch distribution (distribution of the number of differences between all pairs of genes within a sample) and observation of Allele Frequency Clines (AFC) along a southeast to northwest transect. Note that these typical signatures can also result from other phenomenon (e.g selection or bottlenecks), which do not directly concern the present paper. We present here the results obtained for two extreme models of the Neolithic transition. The Diffusion-Replacement model (DR) consists in a complete replacement of HG by NF without any genetic exchanges ($\gamma = 0$). Under the opposite model of Cultural Diffusion (CD), HG are adopting Neolithic technologies from invading Neolithic populations and reproduction is done randomly between the two populations ($\gamma = 1$). Note that this last model is not, strictly speaking, a purely cultural diffusion as it also implies population movements. The reality is certainly somewhere between these two models (Arias 1999; Gronenborg 1999; Mazurié de Keroualin 2003), but we focus here on their genetic consequences in order to study what pattern of diversity are expected under these two contrasting hypotheses, to which real data can be compared. Note that the classical 'Demic Diffusion' model proposed by Ammerman & Cavalli-Sforza (1984) accounts for a very limited local HG incorporation, and is thus close to our DR model.

Genealogical trees

First of all, it is necessary here to describe the genealogical trees obtained under both scenarios (Fig. 4.5A). Under DR, most of the coalescent events occur during the Neolithic expansion and lead to short terminal branches with only few long internal branches. Contrastingly, under the CD model coalescent events occur more regularly through time and there is an almost equal proportion of short and long branches. In both cases, the final coalescent events occur at the time of the Palaeolithic expansion.

Mismatch distributions

The difference in the shape of the trees is reflected in the mismatch distributions inferred from DNA sequences of 300 bp length (corresponding approximately to typical HV1 mtDNA sequences). DR mismatches are often bimodal with a first mode centred on 0 differences whereas almost all CD mismatch distributions are unimodal and smooth (Fig. 4.5B). The mismatch distributions simulated using 22 SNPs (corresponding to Y-chromosome data published by Semino *et al*. 2000) reveal approximately the same patterns but with a much larger variance (Fig. 4.5C). Note that these 22 SNPs were simulated without any ascertainment bias (see below).

Allele frequency clines

One of the typical molecular signatures of a demographic expansion is the occurrence of AFCs along the axis of the expansion, due to successive founder effects (Austerlitz *et al.* 2000; Barbujani *et al.* 1995). The observation of AFCs between the Near East and the northern part of Europe has been considered as a support for a Demic Diffusion of Neolithic farmers from the Near East and consequently for a large replacement of previous European HG populations (Ammerman & Cavalli-Sforza 1984). However, we show here that AFCs could also have been created during the expansion of early modern humans in Europe as proposed by some authors (Barbujani & Bertorelle 2001; Richards *et al.* 1996). Indeed, Neolithic populations followed the same routes than the early modern Europeans did about 30,000 years earlier (Mellars 2004). Under both extreme models (RD and CD) as well as under any intermediate case ($0 < \gamma < 1$), the proportion of randomly simulated mutations showing a cline from the Near East to the northern part of Europe is always smaller than 5 per cent (Table 4.1, column 3). The presence of clines is thus not a support for a large replacement of pre-existing HG people, as it is almost independent of the intensity of HG assimilation reflected by the index γ.

Ascertainment bias

One important aspect of current genetic data is the way molecular markers are selected. Except for mitochondrial DNA, there are only few loci where complete DNA sequences are available for a representative set of populations. Most of the large data sets consist

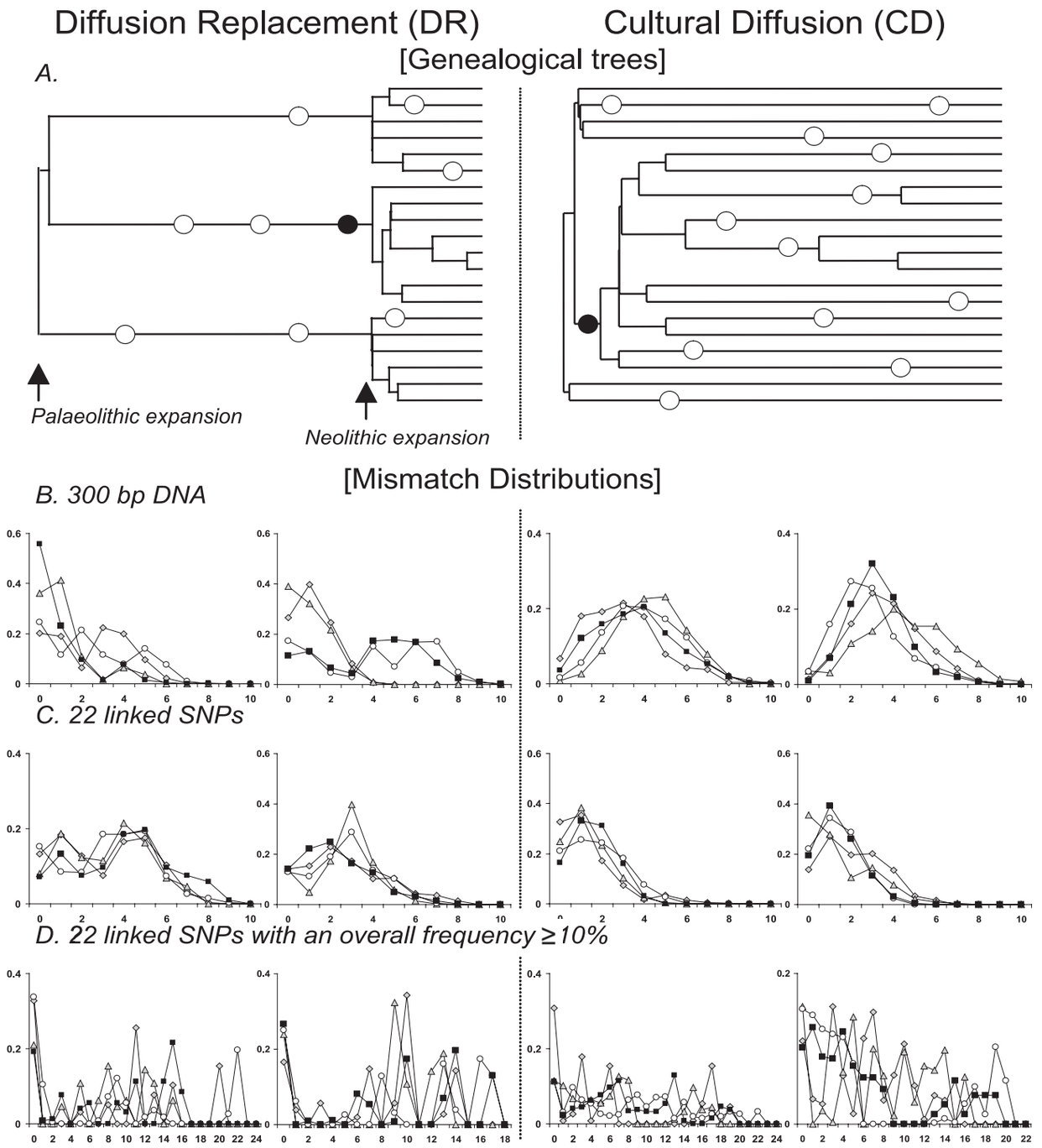

Figure 4.5. *Typical genealogical trees of 20 genes obtained under the two extreme simulated models. DR implements a complete replacement of Mesolithic hunter-gatherers by Neolithic farmers coming from the Near East and CD simulates free genetic exchanges between the two populations. For each model, we show eight random mismatch distributions replicates (*x *axis = number of differences,* y *axis = frequency). Simulations were performed with a large current* Nm *(= 800), where* m *is the migration rate and* N *the density, as estimated for post-Neolithic European populations (Excoffier 2004). White circles on panel A represent random mutations and black circle frequent mutations (see text).*

Table 4.1. *Proportion of simulations (over 10,000) that show a significant AFC at the 5 per cent significance level for various amount of gene flow between HG and NF demes (no gene flow: γ = 0, maximum: γ = 1).* R^2 *= average determination coefficient for the significant AFCs. Proportions are shown for random mutations but also for mutations for which the minor allele has a frequency of at least 5 per cent or 10 per cent (see text).*

Palaeolithic contribution		Allele Frequency Clines					
		All mutations		Bias (f. ≥ 5%)		Bias (f. ≥ 10%)	
γ	No. genes	Prop.	R^2	Prop.	R^2	Prop.	R^2
0.00	0	0.03	0.50	0.57	0.60	0.56	0.62
0.05	1	0.03	0.47	0.48	0.54	0.45	0.58
0.10	2	0.03	0.45	0.50	0.56	0.51	0.63
0.15	3	0.04	0.42	0.51	0.58	0.78	0.70
0.25	5	0.03	0.42	0.66	0.59	0.86	0.71
0.50	10	0.02	0.43	0.71	0.58	0.82	0.68
0.75	15	0.02	0.40	0.70	0.58	0.82	0.67
1.00	20	0.02	0.40	0.68	0.59	0.80	0.63

of SNPs, STRs or RFLPs. All these markers (as well as classical markers, like blood groups) do not necessarily represent a random set of polymorphisms on the genome. These polymorphisms are thus not very representative of random mutations on the genome, because the most frequent polymorphisms (the mutations with the highest frequencies in the population) have a larger probability to be observed and thus to be selected in human variability studies. Rare mutations are thus certainly under-represented in SNP, RFLP and classical marker data sets (Nielsen 2000), which is one kind of ascertainment bias. Corrections for ascertainment bias have been developed for some analyses such as estimation of allele frequency distributions (Nielsen *et al.* 2004), but it is still not clear how this bias can affect other studies. Thus, we have investigated the influence of ascertainment bias in our results by carrying out the same analysis as above, but keeping only mutations for which the minor allele (the less frequent) is at least 5 or 10 per cent on average along our European transect.

Mismatch distribution

Mismatch distributions computed from biased SNPs with a global frequency greater than 10 per cent are ragged and multimodal, independently of the simulated demographic scenario (Fig. 4.5D). Ascertainment bias is thus erasing the signature of an expansion in the mismatch distributions. It seems therefore difficult to infer the role played by demography in shaping the ragged and multimodal mismatch distribution inferred from European Y chromosome SNP diversity (Pereira *et al.* 2001) until the influence of ascertainment bias can be efficiently taken into account.

Allele frequency clines

We examined the effect of ascertainment bias on AFCs over Europe. The fraction of SNPs showing AFCs increases extremely (> 47 per cent, see Table 4.1) when considering SNPs with minor allele frequency larger than 5 per cent over the studied transect. This proportion is again independent of the Palaeolithic/Mesolithic contribution represented by γ (Table 4.1). It shows that the frequency of AFCs increases by an order of magnitude in cases of ascertainment bias. This results can explain why AFCs have been so commonly detected in Europe using SNPs (Rosser *et al.* 2000) and classical markers (Menozzi *et al.* 1978; Sokal *et al.* 1991), but hardly with DNA sequences (Richards *et al.* 1996; 1998; 2002). AFCs may result from demographic expansions, but the probability of observing these clines is clearly related to the age of the SNPs. Indeed, the age of a given mutation is related to its frequency (Kimura & Ohta 1973; Slatkin & Rannala 2000 and see also Fig. 4.5A). If the mutation is old enough to have been present at the time of the demographic expansion, it could have been spread by the expansion wave (Edmonds *et al.* 2004) and then been clinally distributed along the axis of the expansion (Klopfstein *et al.* 2006). One thus expects to see a correlation between the proportion of old mutations in a sample and the probability of observing AFCs.

Concluding remarks

We have presented a new simulation framework to study the effect of more realistic scenarios of range expansion on genetic variability. This approach has allowed us to exclude with much greater confidence than previous attempts, that Neanderthals ever significantly contributed to the modern human gene pool. We also showed that the amount of local Palaeolithic/Mesolithic contribution after the European Neolithic transition could potentially be estimated, since the shape of the coalescent tree depends on this contribution (Fig. 4.5). The more Near Eastern lineages that have been brought into Europe during the Neolithic diffusion, the larger the proportion of bimodal mismatch distributions that can be expected. On the contrary, mismatch distributions should be unimodal and smooth if the local Palaeolithic/Mesolithic contribution was important. We have also shown that the observation of AFCs cannot be considered as support for demic diffusion, as AFCs can be created either by a Palaeolithic expansion of early modern humans or by a Neolithic expansion of farmers. The only way to determine if European AFCs were created during the Neolithic or during the Palaeolithic would be to date them. This seems difficult without a very precise

molecular clock, and because many loci would be required to obtain meaningful estimates.

Our results also underline the importance of taking into account in the analyses the type of genetic markers used. Indeed, the choice of the markers plays a role in the information about demography revealed by the genetic data. For example, we have shown that ascertainment bias (when rare mutations are underrepresented in the sample) promotes the observation of a cline of allelic frequency after a spatial expansion but erases its signature in mismatch distributions. The use of markers with ascertainment bias is thus useful for finding traces of expansion in populations when looking at clines, but not for mismatch distributions.

We did not consider here the last glacial maximum era (LGM) that could have potentially affected human distribution in Europe by potentially inducing large migrations towards southern refuges, but also extinctions in the north (Bocquet-Appel & Demars 2000; Housley *et al.* 1997). The modelling of range contraction is far more difficult than that of range expansions. However, even if humans had actually retreated to southern refuges during the LGM, pre-existing clines could have been potentially re-established during the post-glacial re-expansions from these refuges. Integrating LGM influences within our demographic models would require us to take spatial and temporal environmental heterogeneity into account (Ray *et al.* this volume), which is a challenging task.

In conclusion, realistic simulations are a very powerful tool to predict the expected pattern of molecular diversity resulting from the interactions of a series of populations. The integration of these simulations into an Approximate Bayesian Computation framework (Beaumont *et al.* 2002) should allow us to estimate the parameters of range expansion and the degree of assimilation between Palaeolithic/Mesolithic and Neolithic populations.

Acknowledgements

We are grateful to Shuichi Matsumura and to an anonymous reviewer for their helpful comments on this manuscript. This work was made possible thanks to Swiss NFS grants no 31-054059-98 and 3100A0-100800 to LE.

References

Ammerman, A. & L.L. Cavalli-Sforza, 1984. *The Neolithic Transition and the Genetics of Populations in Europe*. Princeton (NJ): Princeton University Press.

Arias, P., 1999. The origins of the Neolithic along the Atlantic coast of continental Europe. *Journal of World Prehistory* 13, 403–64.

Austerlitz, F., S. Mariette, N. Machon, P.H. Gouyon & B. Godelle, 2000. Effects of colonization processes on genetic diversity: differences between annual plants and tree species. *Genetics* 154, 1309–21.

Barbujani, G. & G. Bertorelle, 2001. Genetics and the population history of Europe. *Proceedings of the National Academy of Sciences of the USA* 98, 22–5.

Barbujani, G. & I. Dupanloup, 2002. DNA variation in Europe: estimating the demographic impact of Neolithic dispersals, in *Examining the Farming/Language Dispersal Hypothesis*, eds. P. Bellwood & C. Renfrew. (McDonald Institute Monographs.) Cambridge: McDonald Institute for Archaeological Research, 421–31.

Barbujani, G. & A. Pilastro, 1993. Genetic evidence on origin and dispersal of human populations speaking languages of the Nostratic macrofamily. *Proceedings of the National Academy of Sciences of the USA* 90, 4670–73.

Barbujani, G., R.R. Sokal & N.L. Oden, 1995. Indo-European origins: a computer-simulation test of five hypotheses. *American Journal of Physical Anthropology* 96, 109–32.

Beaumont, M.A., W. Zhang & D.J. Balding, 2002. Approximate Bayesian computation in population genetics. *Genetics* 162, 2025–35.

Bocquet-Appel, J.-P. & P.Y. Demars, 2000. Population kinetics in the Upper Palaeolithic in western Europe. *Journal of Archaeological Science* 27, 551–70.

Chikhi, L., 2002. Admixture and the demic diffusion model in Europe, in *Examining the Farming/Language Dispersal Hypothesis*, eds. P. Bellwood & C. Renfrew. (McDonald Institute Monographs.) Cambridge: McDonald Institute for Archaeological Research, 435–47.

Chikhi, L., G. Destro-Bisol, G. Bertorelle, V. Pascali & G. Barbujani, 1998. Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proceedings of the National Academy of Sciences of the USA* 95, 9053–8.

Currat, M. & L. Excoffier, 2004. Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biology* 2, 2264–74.

Currat, M. & L. Excoffier, 2005. The effect of the Neolithic expansion on European molecular diversity. *Proceedings of the Royal Society of London B Biological Sciences* 272, 679–88.

Currat, M., N. Ray & L. Excoffier, 2004. SPLATCHE: a program to simulate genetic diversity taking into account environmental heterogeneity. *Molecular Ecology Notes* 4, 139–42.

Dupanloup, I., G. Bertorelle, L. Chikhi & G. Barbujani, 2004. Estimating the impact of prehistoric admixture on the genome of europeans. *Molecular Biology and Evolution* 21, 1361–72.

Edmonds, C.A., A.S. Lillie & L.L. Cavalli-Sforza, 2004. Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the USA* 101, 975–9.

Excoffier, L., 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* 13, 853–64.

Gronenborg, D., 1999. A variation on a basic theme: the tran-

sition to farming in southern central Europe. *Journal of World Prehistory* 13, 123–210.

Hagelberg, E., 2003. Recombination or mutation rate heterogeneity? Implications for Mitochondrial Eve. *Trends in Genetics* 19, 84–90.

Handt, O., S. Meyer & A. von Haeseler, 1998. Compilation of human mtDNA control region sequences. *Nucleic Acids Research* 26, 126–12.

Housley, R.A., C.S. Gamble, M. Street & P. Pettit, 1997. Radiocarbon evidence for the lateglacial human recolonisation of northern Europe. *Proceedings of the Prehistoric Society* 63, 25–54.

Hublin, J.J., F. Spoor, M. Braun, F. Zonneveld & S. Condemi, 1996. A late Neanderthal associated with Upper Palaeolithic artefacts. *Nature* 381, 224–6.

Hudson, R.R., 1990. *Gene Genealogies and the Coalescent Process*. Oxford: Oxford University Press.

Kimura, M. & T. Ohta, 1973. Eukaryotes-prokaryotes divergence estimated by 5S ribosomal RNA sequences. *Nature New Biology* 243, 199–200.

King, J.P., M. Kimmel & R. Chakraborty, 2000. A power analysis of microsatellite-based statistics for inferring past population growth. *Molecular Biology and Evolution* 17, 1859–68.

Kingman, J.F.C., 1982. The coalescent. *Stochastic Processes and their Applications* 13, 235–48.

Klopfstein, S., M. Currat & L. Excoffier, 2006. The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution* 23(3), 482–90.

Krings, M., H. Geisert, R.W. Schmitz, H. Krainitzki & S. Pääbo, 1999. DNA sequence of the mitochondrial hypervariable region II from the Neandertal type specimen. *Proceedings of the National Academy of Sciences of the USA* 96, 5581–5.

Krings, M., A. Stone, R.W. Schmitz, H. Krainitzki, M. Stoneking & S. Pääbo, 1997. Neandertal DNA sequences and the origin of modern humans. *Cell* 90, 19–30.

Lotka, A.J., 1932. The growth of mixed populations: two species competing for a common food supply. *Journal of the Washington Academy of Sciences* 22, 461–9.

Mazurié de Keroualin, K., 2003. *Genèse et diffusion de l'agriculture en Europe: agriculteurs, chasseurs, pasteurs.* Paris: Errance.

Mellars, P., 2004. Neanderthals and the modern human colonization of Europe. *Nature* 432, 461–5.

Menozzi, P., A. Piazza & L. Cavalli-Sforza, 1978. Synthetic maps of human gene frequencies in Europeans. *Science* 201, 786–92.

Nielsen, R., 2000. Estimation of population parameters and recombination rates from single nucleotide polymorphisms. *Genetics* 154, 931–42.

Nielsen, R., M.J. Hubisz & A.G. Clark, 2004. Reconstituting the frequency spectrum of ascertained single-nucleotide polymorphism data. *Genetics* 168, 2373–82.

Nordborg, M., 1998. On the probability of Neanderthal ancestry. *American Journal of Human Genetics* 63, 1237–40.

Nordborg, M., 2001. Coalescent theory, in *Handbook of Statistical Genetics*, eds. D. Balding, C. Bishop & C. Cannings. New York (NY): John Wiley & Sons, 179–212.

Ovchinnikov, I.V., A. Götherström, G.P. Romanova, V.M. Kharitonov, K. Liden & W. Goodwin, 2000. Molecular analysis of Neanderthal DNA from the nothern Caucasus. *Nature* 404, 490–93.

Pereira, L., I. Dupanloup, Z.H. Rosser, M.A. Jobling & G. Barbujani, 2001. Y-chromosome mismatch distributions in Europe. *Molecular Biology and Evolution* 18, 1259–71.

Ray, N., M. Currat & L. Excoffier, 2003. Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology and Evolution* 20, 76–86.

Ray, N., M. Currat, P. Berthier & L. Excoffier, 2005. Recovering the geographic origin of early modern humans by realistic and spatially explicit simulations. *Genome Research* 15, 1161–7.

Relethford, J.H., 2001. Absence of regional affinities of Neandertal DNA with living humans does not reject multiregional evolution. *American Journal of Physical Anthropology* 115, 95–8.

Richards, M., 2003. The Neolithic invasion of Europe. *Annual Review of Anthropology* 32, 135–62.

Richards, M., H. Corte-Real, P. Forster, *et al.*, 1996. Palaeolithic and Neolithic lineages in the European mitochondrial gene pool. *American Journal of Human Genetics* 59, 185–203.

Richards, M.B., V.A. Macaulay, H.-J. Bandelt & B.C. Sykes, 1998. Phylogeography of mitochondrial DNA in western Europe. *Annual of Human Genetics* 62(3), 241–60.

Richards, M., V. Macaulay, E. Hickey, *et al.*, 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *American Journal of Human Genetics* 67, 1251–76.

Richards, M., V. Macaulay, A. Torroni & H.-J. Bandelt, 2002. In search of geographical patterns in European mitochondrial DNA. *American Journal of Human Genetics* 71, 1168–74.

Rogers, A.R. & H. Harpending, 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9, 552–69.

Rosser, Z.H., T. Zerjal, M.E. Hurles, *et al.*, 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *American Journal of Human Genetics* 67, 1526–43.

Schmitz, R.W., D. Serre, G. Bonani, *et al.*, 2002. The Neandertal type site revisited: interdisciplinary investigations of skeletal remains from the Neander Valley, Germany. *Proceedings of the National Academy of Sciences of the USA* 99, 13,342–7.

Semino, O., G. Passarino, P.J. Oefner, *et al.*, 2000. The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290, 1155–9.

Serre, D., A. Langaney, M. Chech, *et al.*, 2004. No evidence of Neandertal mtDNA contribution to early modern humans. *PLoS Biology* 2, E57.

Slatkin, M. & R.R. Hudson, 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* 129, 555–62.

Slatkin, M. & B. Rannala, 2000. Estimating allele age. *Annual*

Review of Genomics and Human Genetics 1, 225–49.
Sokal, R.R., N.L. Oden & C. Wilson, 1991. Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature* 351, 143–5.
Tajima, F., 1989. The effect of change in population size on DNA polymorphism. *Genetics* 123, 597–601.
Volterra, V., 1926. Variations and fluctuations of the numbers of individuals in animal species living together, in *Animal Ecology*, ed. R.N. Chapman. Reprinted in 1931. New York (NY): McGraw Hill, 409–48.