



Rasteiro, R., & Chikhi, L. (2013). Female and Male Perspectives on the Neolithic Transition in Europe: Clues from Ancient and Modern Genetic Data. *PLoS ONE*, 8(4), [e60944].
<https://doi.org/10.1371/journal.pone.0060944>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1371/journal.pone.0060944](https://doi.org/10.1371/journal.pone.0060944)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Public Library of Science at <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0060944> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/pure/about/ebr-terms>

Female and Male Perspectives on the Neolithic Transition in Europe: Clues from Ancient and Modern Genetic Data

Rita Rasteiro^{1*‡}, Lounès Chikhi^{1,2,3*}

1 Instituto Gulbenkian de Ciência, Oeiras, Portugal, **2** CNRS, Université Paul Sabatier, ENFA, UMR 5174 EDB (Laboratoire Evolution et Diversité Biologique), Toulouse, France, **3** Université de Toulouse; UPS; EDB (Laboratoire Evolution et Diversité Biologique), Toulouse, France

Abstract

The arrival of agriculture into Europe during the Neolithic transition brought a significant shift in human lifestyle and subsistence. However, the conditions under which the spread of the new culture and technologies occurred are still debated. Similarly, the roles played by women and men during the Neolithic transition are not well understood, probably due to the fact that mitochondrial DNA (mtDNA) and Y chromosome (NRY) data are usually studied independently rather than within the same statistical framework. Here, we applied an integrative approach, using different model-based inferential techniques, to analyse published datasets from contemporary and ancient European populations. By integrating mtDNA and NRY data into the same admixture approach, we show that both males and females underwent the same admixture history and both support the demic diffusion model of Ammerman and Cavalli-Sforza. Similarly, the patterns of genetic diversity found in extant and ancient populations demonstrate that both modern and ancient mtDNA support the demic diffusion model. They also show that population structure and differential growth between farmers and hunter-gatherers are necessary to explain both types of data. However, we also found some differences between male and female markers, suggesting that the female effective population size was larger than that of the males, probably due to different demographic histories. We argue that these differences are most probably related to the various shifts in cultural practices and lifestyles that followed the Neolithic Transition, such as sedentism, the shift from polygyny to monogamy or the increase of patrilocality.

Citation: Rasteiro R, Chikhi L (2013) Female and Male Perspectives on the Neolithic Transition in Europe: Clues from Ancient and Modern Genetic Data. PLoS ONE 8(4): e60944. doi:10.1371/journal.pone.0060944

Editor: Carles Lalueza-Fox, Institut de Biologia Evolutiva - Universitat Pompeu Fabra, Spain

Received: March 21, 2012; **Accepted:** March 5, 2013; **Published:** April 17, 2013

Copyright: © 2013 Rasteiro, Chikhi. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: RR was funded by a Fundação para a Ciência e Tecnologia (FCT, Portugal, <http://www.fct.pt>) grant (ref. SFRH/BD/30821/2006). LC was partly funded by the CNRS (Centre National de la Recherche Scientifique, France), the "Laboratoire d'Excellence (LABEX, http://www6.inra.fr/labex-tulip_eng/layout/set/print)" entitled TULIP (ANR-10-LABX-41) and the FCT grant PTDC/BIA-BDE/71299/2008. HPC resources from CALMIP (www.calmip.cict.fr), Toulouse (Grant 2010-P1038 and Grant 2012-P1244 to LC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: rr147@le.ac.uk (RR); chikhi@igc.gulbenkian.pt (LC)

‡ Current address: Department of Genetics/School of Historical Studies, University of Leicester, Leicester, United Kingdom

Introduction

Major progress has been made in the use of genetic data to reconstruct the demographic history of human populations and compare alternative models of human origins [1], [2], [3]. Despite these advances, one of the most important cultural, economic and demographic revolutions in human prehistory, the Neolithic transition [4], remains the subject of continuing and hotly debated controversies [3], [5], [6], [7], [8], [9], [10], [11]. Even for Europe, where most genetic studies have been carried out, there is a major disagreement among archaeologists and anthropologists [9], [12], [13], [14], [15] and among geneticists [5], [8], [16], [17]. Some favour the hypothesis that this process resulted from an active migratory process starting in the Near East, where the domestication of Old World animals and plants began [9], whereas others believe that it was merely due to cultural contact between hunter-gathering and farming societies. These two extreme alternatives are usually encapsulated in two widely used models assuming either demic diffusion (DDM) [18] or cultural diffusion (CDM) [19]. The CDM predicts that there should be no

or very little contribution in Europe from the Near Eastern populations. The genetic consequences of the DDM are much less straightforward and depend on the details of the spatial processes that took place during the expansion, including the importance of intermarriage (admixture) events between farmers and hunter-gatherers (*HG*) [1], [8], [20]. For instance, Chikhi *et al.* [8] showed that even assuming that farmers represented 90% of all the newly formed farming societies (and with only 10% of *HG*) as they expanded into Europe, the average contribution of Near Eastern genes in Europe could be as low as a few per cent, due to a dilution effect along the expansion axis, and close to zero on the western borders of Europe. They stressed a fundamental asymmetry between the two models in terms of genetic patterns and the need to use model-based approaches explicitly accounting for drift and admixture. These points were also stressed by Currat and Excoffier [1], who used more complex and sophisticated models.

Until now, one of the major limitations in the studies published is the fact that they either use mtDNA or NRY (non-recombinant region of the Y-chromosome) data, which are sometimes claimed to favour opposite models [21], even though they have never been

used jointly. For instance, mtDNA data are often claimed to support CDM [5], [6], [7] whereas NRY data would support the DDM [8], [22], [23]. It is indeed very tempting to imagine that, during the Neolithic expansion in Europe, male farmers eliminated *HG* males whereas they integrated *HG* females in the newly founded farming societies, hence generating an asymmetry between male and female lineages similar to that described between Bantu speakers and African *HG* societies [24] or during the colonization of the Americas by Europeans [25].

In addition, recent technological advances have allowed the use of ancient DNA (aDNA) from early *HG* and farmer societies, hence raising new hopes that the long-lasting controversy between the CDM and DDM can be resolved. However, the recent attempts to model the colonization of Europe using ancient and modern DNA jointly [26], [27], [28], [29], have assumed very simple models that fail to incorporate crucial aspects of the demographic history of early Europeans including Neolithic farmers. They have also, in most cases, failed to use some recent advances in population genetics modelling and statistical inference. This has led to contradictory and inconsistent conclusions as we shall discuss here.

In a recent work [30], we have carried out one of the first studies where mtDNA and NRY data were analysed jointly to model ancient demographic events. Here, we continue along that road and use a simple admixture model (Figure S1) to study the spread of agriculture in Europe, by expanding the modern NRY dataset [23] and by adding modern mtDNA data [5] (see SI Material and Methods). We also take an Approximate Bayesian Computation (ABC) approach [31], [32] using one of the largest aDNA dataset available [27], to identify the demographic scenarios that could explain both modern and ancient DNA data.

We show for the first time that (i) there are no major contradictions between NRY and mtDNA data, (ii) both exhibit a clear decrease of the Neolithic contribution with geographic distance from the Near East, (iii) both favour a DDM. But there are also differences between the two markers. We show that (iv) the female effective population size was larger than that of the males, suggesting that the demographic history of males and females was significantly different before and during the Neolithic transition, probably due to differences in the migration patterns and mating systems prior to and after the arrival of agriculture. By combining evidence from both modern and ancient mtDNA we also demonstrate that (v) genetic drift and population structure were extremely important in both *HG* and farming societies, explaining why aDNA data can produce many alleles with frequencies that are significantly different from present-day frequencies and (vi) that aDNA also supports the DDM. Altogether, we propose a synthetic model of colonization that accounts for both modern and ancient mtDNA and NRY data.

Results

Admixture Analyses: The Neolithic Contribution Decreases with Distance from the Near East, for both NRY and mtDNA Data

Figures 1A (mtDNA) and S2 (NRY) show the posterior distributions for p_1 , the Palaeolithic contribution to the European populations analysed. As expected from simulations [33], [34], the distributions are rather wide and each single population estimate has a large standard error, confirming that population genetic parameters estimated using single *locus* data are rarely very accurate. Nevertheless, when all populations are considered jointly, a clear geographic pattern is seen in both the new NRY and mtDNA (Figure 1B) datasets. This pattern shows that the

proportion of Neolithic genes ($1-p_1$) decreases from modal values of around 100% in Greece and Cyprus, to 75% in Romania, 30% in France and 20% in Spain (Figure 1B). This confirms previous results that used another independent NRY data set [8]. This trend is detected for the first time in mtDNA data, which have repeatedly been claimed to exhibit no SE-NW spatial pattern [5], [6]. Figure 1B shows that the three (two NRY and one mtDNA) datasets produce the same general trend, hence supporting a parallel decrease of female and male lineages from Neolithic farmers in the genome of modern Europeans, as we move away from the Near-East. The two NRY datasets exhibit differences, due to the fact that different populations were sampled, different numbers of SNPs were genotyped, and sample sizes were also different between the two. However, one of the NRY datasets exhibits a cline that is near identical to the cline detected for mtDNA. This strongly suggests that the difference between mtDNA and one of the NRY datasets is not greater than expected under stochasticity.

The Neolithic Transition in the Caucasus and European islands: NRY Admixture Analyses

Another set of new results is found with the NRY samples from the Caucasus (Armenia, Georgia and Ossetia). First, the admixture level of these populations is exactly at the level expected if they had been on a SE-NW expansion axis (*i.e.* along the general direction of farmers expansion towards Europe during the Neolithic), even though they are geographically located NE of the Fertile Crescent and not NW (Figure S3A). Second, when the Caucasus data are analysed independently from the rest of the data, we find a significant geographical trend, as expected if agriculture has expanded demically from the Near East outwards in several directions, *i.e.* not just towards Europe (Figure S4A), as predicted by Renfrew's theory linking the expansion of Indo-European with the expansion of agriculture [35], [36]. Third, the same analysis performed using populations that are unlikely to have played a major role during the Neolithic transition, due to their geographic location (*i.e.* negative controls, see SI Material and Methods) exhibit no such trend despite their much larger sample sizes (Figure S3B). Fourth, contrary to the negative controls used, several European islands population samples (East Anglia, Ireland, Cyprus and Sardinia and British Isles populations) appear to also fit within the general decrease in admixture across Europe (Figure S4B). Thus, we find clines in the Caucasus and European Islands but not in populations from the Eastern/Northern Europe.

Drift in Paternal and Maternal Lineages: NRY and mtDNA Data Support the DDM but not the Same Demographic Histories

Genetic drift is represented by parameter t_i that represents the ratio of T , the time since the admixture event, and N_i the effective size of population i (see Figure S1). Thus, genetic drift in the different parental populations is represented by the parameters t_1 and t_2 for the Palaeolithic and Neolithic populations, respectively. Each of the t_1 and t_2 posterior distributions is obtained independently by the analysis of one European population (Figures S5A–B, S6A–B). First, we find that the t_1 posterior values are always higher than the t_2 values suggesting that genetic drift has been more important in the “Palaeolithic” than in the “Neolithic” parental population, in agreement with a later population size increase related with the arrival of agriculture. Second, for all the European populations analysed the t_1 (and t_2) posterior distribu-

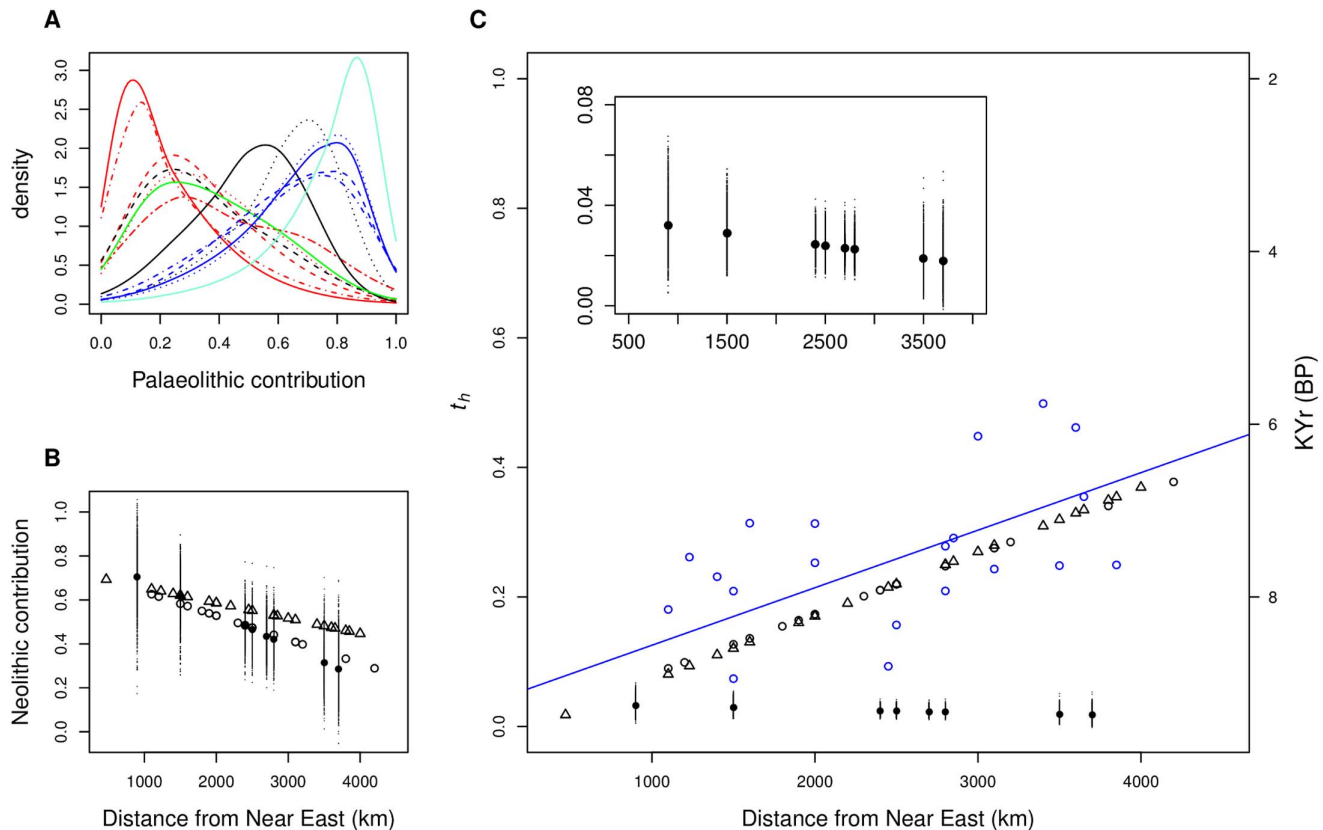


Figure 1. Spatial variation of admixture and drift across Europe. In (A) are represented the posterior distributions of the Palaeolithic contribution (HG contribution to modern European), for each of the populations analysed, using mtDNA data. Each curve corresponds to the analysis of a specific admixed population (Armenia – red, Caucasus – dashed red, Azeri – dotted red, Egypt – dotdash red, Iran – twodash red, Central Mediterranean – black, East Mediterranean – dashed black, West Mediterranean – dotted black, Southeast Europe – green, North and Central Europe – blue, Northeast Europe – dashed blue, Northwest Europe – dotted blue, Alps, dotdash blue and Scandinavia – aquamarine). (B) Linear regression of Neolithic contribution, against geographical distance from Near East. For each of the samples, one $1-p_1$ value was randomly sampled from the corresponding posterior distribution. A linear regression was then calculated between this set of values and geographic distance. This process was repeated 1,000 times to obtain the empirical distribution of regression curves. The fitted values using mtDNA data are plotted for each of the 1,000 replicates. As fitted values are plotted, they can occur outside the range (0–1). Mean values for each population are represented by solid circles (mtDNA data) and open triangles and circles (for two different NRY datasets, Rosser et al. [23] and Semino et al. [17], respectively). In (C) a similar approach was used to represent the linear regression of t_h (drift in the admixed populations) against geographic distance from the Near East. Mean values for each population for mtDNA and NRY datasets are plotted, with symbol codes as in (B). The close-up inset shows the mtDNA regression on a different scale for the Y-axis. Mean values for each population are represented for the sake of clarity. Calibrated radiocarbon dates of Neolithic archaeological sites [13] (see table S4) are also plotted against the distance from the Near East (blue open circles), with the linear regression represented by the blue line.
doi:10.1371/journal.pone.0060944.g001

tions are tightly clustered, rather than spread out, even though each analysis is performed independently. Third, the different t_1 posterior values are more diverse (*i.e.* less clustered) than the t_2 distributions, which is expected if the early HG populations were differentiated, due to their smaller effective sizes. Fourth, the t_1 and t_2 posteriors obtained for the mtDNA datasets support much lower values than the corresponding NRY t_1 and t_2 posteriors, suggesting a much larger female (N_f) than male (N_m) population effective size and/or higher female gene flow.

Fifth, Figure 1C shows the results for the parameter t_h which represents drift in the different European populations since the admixture event. We find that for NRY data, t_h is positively correlated with distance from the Near East and with the earliest date of arrival of agriculture in the different locations based on archaeological artefacts (*i.e.* drift increases for European populations that had a HG lifestyle for a longer period and admixed later). In other words, the male global effective size will always be larger in the Near East (see also Figure S7). For the mtDNA data, the

geographical trend is very different. Low t_h values are observed in the Near East, but instead of increasing with distance they exhibit (almost) no trend (see inset in Figure 1C showing a decrease). It thus appears that the mtDNA and NRY t_h results require different explanations for the demographic history of males and females, while favouring both the DDM. Sixth, differences between males and females are also observed when measures of genetic diversity (H_s) and differentiation (F_{ST}) are regressed against geographic distance from the Near East. For mtDNA, genetic differentiation between Europeans and Near Easterners increases much less with increasing geographical distance than for NRY data (Figure S8A). In agreement with this trend, differences in diversity levels are also less important in mtDNA than in NRY data (Figure S8B). Both support a higher N_f and/or higher female migration rates.

Ancient DNA, Coalescent Simulations and Model Identification Using ABC

Figure 2 represents the three demographic scenarios tested together with their posterior probabilities, using two ABC model choice algorithms on aDNA data [27]. Whether we use the multinomial logistic regression (MLR) method of Beaumont [37] or the non-linear heteroscedastic neural network (NCH) approach of Blum and François [32], the support for the Total Panmixia (TP) model is *nil*, whereas the best supported model, with a posterior probability >0.957 , is the Split with Differential Growth (SDG) model which assumes a differential growth between Neolithic and Palaeolithic farmers (see also Figure S9). These results suggest that structure is required between *HG* and farmers to explain the observed data (SDG and S [Split] *vs.* TP) and that differential growth is also required (SDG *vs.* S). Furthermore, the parameters estimated for the SDG suggest that the growth rate in the *HG* populations, during the Palaeolithic, was very low or null (see table S1).

The same kind of results, but using another approach, is shown in Figure 3. This figure represents the estimated probability of obtaining F_{ST} values that are equal or higher

than those observed in the real data ($P_{S>O}$), for the three scenarios. A two-tailed test was also applied and the results were (qualitatively) identical i.e. the result of the statistical test did not change (not shown). The data simulated under the TP model (Figure 3 A–C) show results identical to those obtained by Bramanti and colleagues [27], hence validating our simulation approach and the exaggerated simplicity of the model used by these authors. For this model, the parameter space explaining the observed data is extremely limited. However, as soon as structure is incorporated in the models (S and SDG), the number of parameter combinations (N_{UP} and N_N) for which large F_{ST} values are observed becomes very large hence allowing for many realistic scenarios to explain the observed data. This is true for the S model (Figure 3 D–F) and even more when we introduce differential growth in the model (Figure 3 G–I). For instance, the $P_{S>O}$ values in the SDG model panels can be as high as 0.99 for the *HG vs.* farmer comparisons or as high as one for the *HG vs.* modern European comparison, showing that simple structured models produce high F_{ST} values for reasonable parameter values. Conversely, the simulations for the TP model have maximum $P_{S>O}$ values

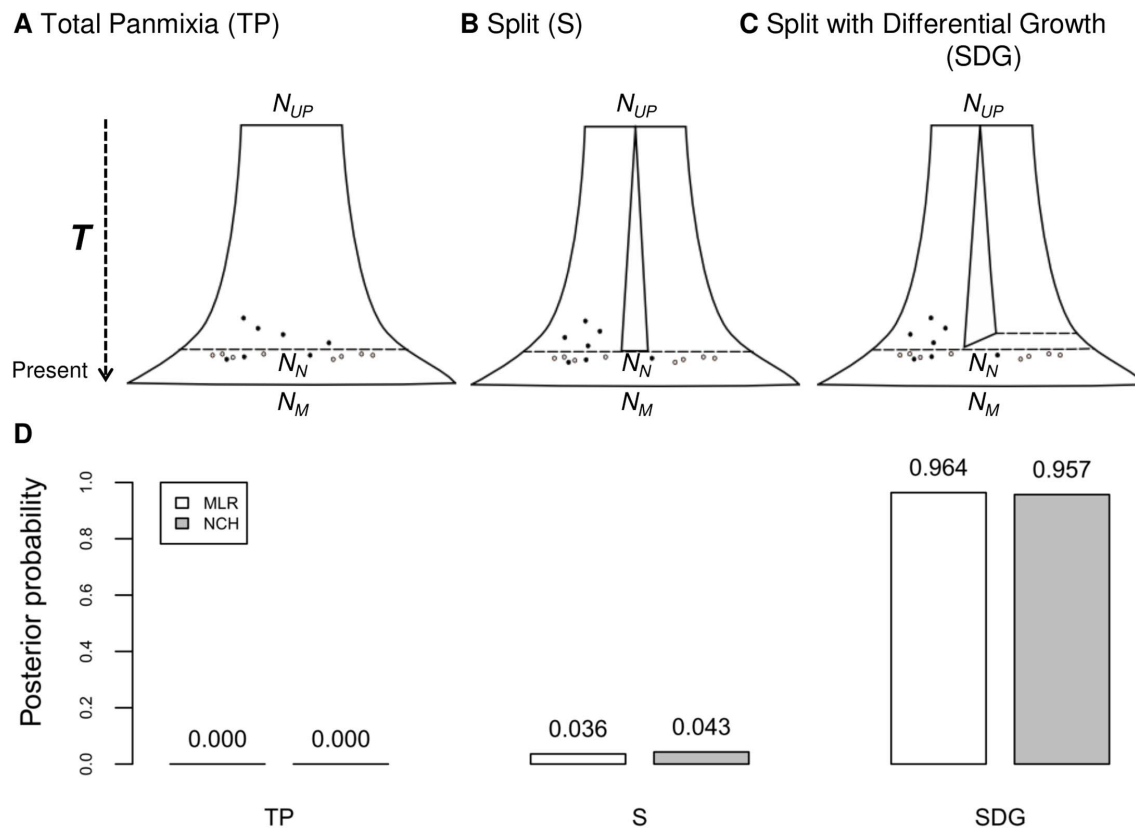


Figure 2. Demographic models used in the aDNA analysis and their posterior probabilities. Three different demographic models were tested using ancient and modern mtDNA data. The Total Panmixia (TP) model (A) follows the assumptions of Bramanti et al. [27], where *HG* and farmers were part of the same panmictic population over Central Europe and were never separated in different populations or communities. This model was used assuming a single modern female effective population size N_M and two periods of exponential growth: i) the first starting with an Upper Palaeolithic (UP) population of effective size N_{UP} , sampled from an ancestral African female population of constant size, corresponding to the initial colonization of Central Europe 45,000 years ago and ii) the second following the Neolithic Transition 7,500 years ago, from a population of effective size N_N . Both N_{UP} and N_N population sizes were allowed to vary using the same priors as in [27]. In the Split Model (S) (B), the UP population was structured in two sub-populations of equal size, 45,000 years ago. These sub-populations were assumed to grow independently (no gene flow), until they joined together at the beginning of the Neolithic, in Central Europe. The Split with Differential Growth (SDG) model (C) is similar to the S model but has a more complex splitting, in which one of the two sub-populations was allowed to have a higher growth rate between 10,000 and 7,500 years ago. In (D) are represented the posterior probabilities under each model, calculated using the ABC framework, for two different types of post-rejection adjustments: MLR (white bars) and NCH (grey bars). doi:10.1371/journal.pone.0060944.g002

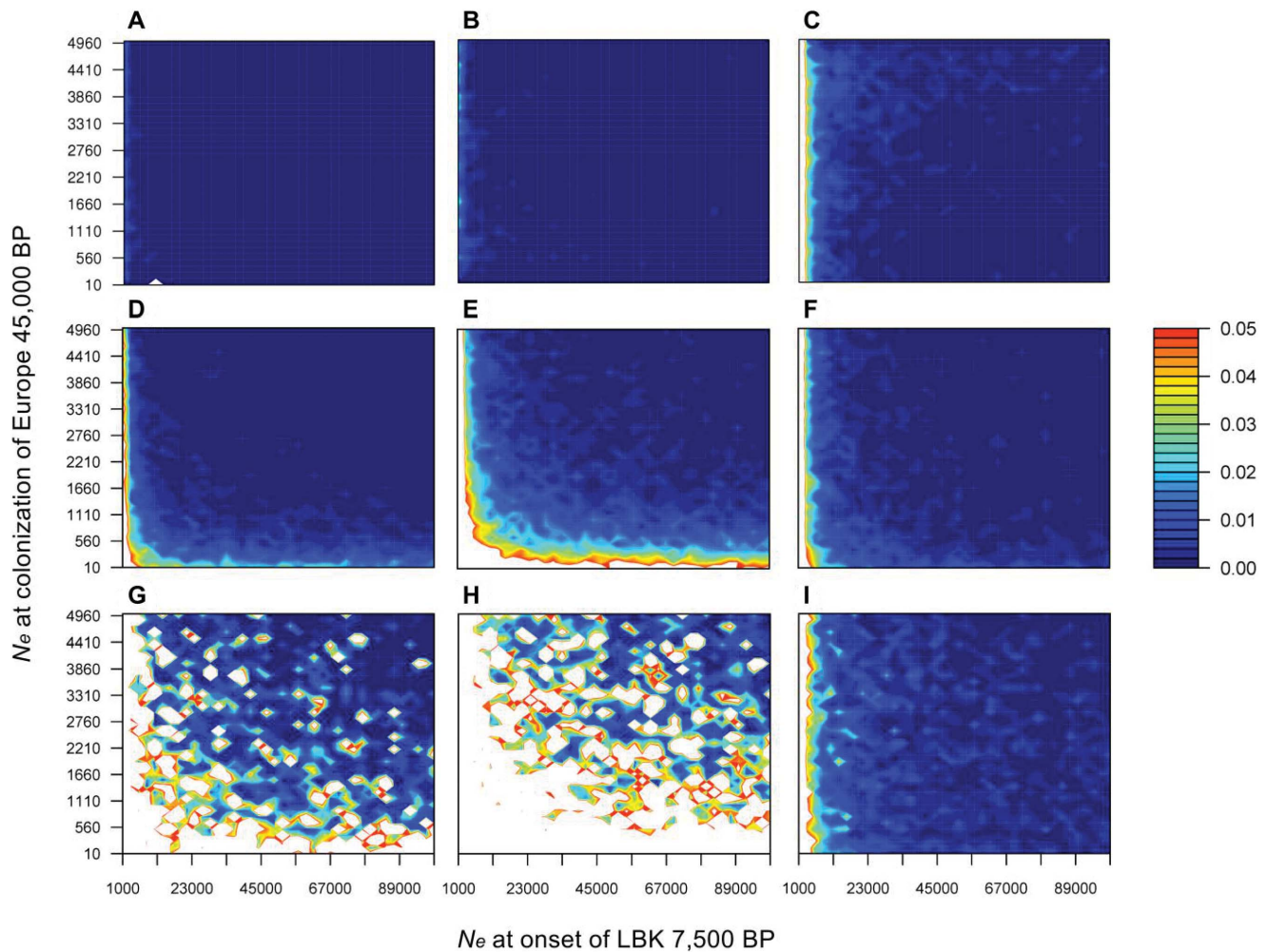


Figure 3. Probability of obtaining genetic differentiation (F_{ST}) values larger than those observed in the real data. The panels in each row correspond to data simulated under the TP model (A, B, C), the S model (D, E, F) and the SDG model (G, H, I) (see Figure 2, for models definitions). Each column corresponds to a specific pairwise F_{ST} comparison, namely between HG and early farmers (A, D, G), HG and modern Europeans (B, E, H), and early farmers and modern Europeans (C, F, I). The x- and y-axis represent the values used for the female effective size N_N (at the onset of the Central European Neolithic 7,500 years ago) and N_{UP} (45,000 years ago), respectively. The colour key gives the probability of obtaining a F_{ST} value equal or greater than that observed. The white shaded area corresponds to parameter combinations for which this probability is greater than 0.05.

doi:10.1371/journal.pone.0060944.g003

of 0.018 for the first comparison and 0.032 for the latter, in agreement with the values found by Bramanti and colleagues [27] (see table S2).

Discussion

Both Contemporary NRY and mtDNA Data Support DDM, but Tell Different Demographic Histories

Our analyses, using contemporary data, suggest that there is a parallel decrease in the NRY and mtDNA Neolithic contributions to the European populations with increasing distance from the Near East. This is not compatible with a model of cultural diffusion and requires demic movement of both male and female farmers, from the Near East, as agriculture spread into Europe, in agreement with archaeological data [12], [13], [15], [38]. This parallel decrease also suggests that both males and females admixed with the local Palaeolithic populations that inhabited Europe at the time, resulting in a progressive dilution of the Near Eastern genes. We also found that the demic diffusion process was

centrifugal, with samples from the Caucasus fitting in the general trend, as was already suggested by Renfrew [35] and others [39] and in agreement with linguistic data too [36]. Moreover, the European islands appear also to fit within this trend. This suggests that the sea did not represent a major barrier to the Neolithic expansion and that the peopling of these islands was not subjected to major drift effects or radically different admixture histories compared to neighbouring continental populations [12].

It therefore appears that, when we use one coherent statistical framework, both datasets from male [8], [23] and female [5] markers, support the DDM. These results are at odds with the original conclusions drawn by Richards *et al.* [5] (*i.e.* using only mtDNA), who advocated that mtDNA data favoured the CDM. However, they are in agreement with the clines described by Rosser *et al.* [23] (*i.e.* only with NRY data). It is worth noting that the methods used by the two studies are not comparable. Richards *et al.* [5] used the age of mtDNA mutations and haplogroups to date major demographic events. This kind of approach has been criticised as it can lead to misinterpretation of the data [3], [10],

[40]. Rosser *et al.* [23] used spatial autocorrelation methods instead, to identify statistically significant clines. This method has been similarly criticised, as a cline in itself does not indicate the time at which it was established. Model-based approaches, like those applied here, explicitly state the assumptions used to make inference and are probably the most suitable to infer demographic parameters [1], [8], [41], such as the Neolithic contribution to European populations.

The fact that extant NRY and mtDNA both support the DDM does not imply that other details of the male and female demography were identical, particularly in relation with the amount of drift experienced by each sex [42]. Indeed, our results point to a higher N_f over N_m , in agreement with the larger coalescence times for mtDNA [43], [44]. But before addressing this issue and proposing a model accounting for these results we turn to the aDNA results.

aDNA Supports Demic Diffusion

The first aDNA study using model-based approaches, on samples identified as Linear Pottery Culture (LBK), argued in favour of CDM [26]. Later, the same LBK data was compared to samples from Palaeolithic/Mesolithic archaeological sites and modern data from the same region, by Bramanti *et al.* [27]. They interpreted the genetic differentiation observed in the real data as being too high to “*be explained by population continuity alone*” [27], hence arguing for a Neolithic immigration in Central Europe. These two studies [26], [27] had in common that all DNA samples, ancient and modern alike, were assumed to belong to the same panmictic population (see Figure 2A). While this may seem surprising, the model assumed in these two studies is the one that we call Total Panmixia. This model assumes that there was no population structure and that *HG* and farmers were allowed to mate freely, making the distinction between *HG* and farmers unclear, to say the least.

What our new aDNA simulation framework suggests is that it is actually possible to explain the large genetic differentiation between samples if we explicitly model both population structure and different population growth rates between Neolithic and Palaeolithic populations before they admixed. In a recent work, Haak and colleagues [29] also allowed for some population structure, namely between populations of Central Europe and the Near East. Their results suggested an affinity between the first LBK farmers and modern Near Easterners, but they still could not explain the high population differentiation encountered between the LBK farmers and present-day Central European populations. On the contrary, our SDG model, could explain the high F_{ST} values encountered between *HG* and farmers and between farmers (or *HG*) and modern-day Central Europeans. We believe that the main difference with the Haak *et al.* study [29] is that they did not allow variable population growth rates in their simulations. However, by varying the growth rates between *HG* and farmers, as between the onset of farming and the following period, we could explain these high F_{ST} values.

Differential growth between farmers and *HG* is supported by anthropological and archaeological data [45], [46]. Indeed, at the onset of the Neolithic expansion in the Near East and in the front of the wave of expansion, it has been shown that a very high growth rate is expected from the colonizing populations until their size reaches the new carrying capacity ceilings [45]. Interestingly, our estimates suggest that the female growth rate remained quasi-constant during the Palaeolithic, and that there was an expansion with the advent of farming, which is also in agreement with archaeological data [47], [48]. Such an increase in N_f could also be explained by an increase in gene flow

following the arrival of farming, for instance if it was accompanied by a change in post-marital residence patterns in females. This is in agreement with a simulation study by Rasteiro *et al.* [30]. These authors simulated genetic data (mtDNA and Y-chromosome) for 45 scenarios by varying the amount of admixture between *HG* and farmers populations and the patterns of post-marital residence behaviour, hence allowing for a shift after the arrival of agriculture. This is also in agreement with strontium data recently published demonstrating a sudden increase in female gene flow after the arrival of agriculture in the Balkans [38] or in the LBK [49].

Towards an Integrated Model of Neolithic Transition

Altogether, the work presented here allows us to draw a coherent integrated model for the Neolithic transition in Europe which accounts for both the congruent admixture results between mtDNA and NRY data, their difference in terms of diversity and differentiation (drift), and the constraints imposed by the aDNA data. On that basis, we propose (i) an establishment of farming communities in Europe by a demic diffusion process, with an origin in the Near East, in agreement with archaeological [12], [13], [15], [46], [50] and anthropological studies [14], [51], [52], along with a process of admixture with the local *HG* [51]; (ii) a spread in different directions from the Near East, with the Caucasus and European Islands being part of this gradual expansion, in agreement with Renfrew’s theory of Indo-European languages [35], [36]. Furthermore, we propose that (iii) both male and female farmers were involved in this demic movement in agreement with strontium data [38], [49], and that (iv) the demographic histories of the two sexes were probably different during and perhaps before the Neolithic transition. In particular, we propose that the difference in the amount of drift experienced by males and females can be explained by a change in the patterns of gene flow and by a shift in human mating systems, from polygyny to monogamy during to the Neolithic transition. Below we go through the rationale and data that corroborate this scenario.

As noted above, one of our main results is that $N_f > N_m$ and/or that migration rates were higher in females compared with males (Figure 1C). Anthropological, linguistic and archaeological evidence suggest that the transition from hunting-gathering to farming or herding communities usually leads to an increase in patrilocality (*i.e.* when the marital residence is the groom’s birthplace) due to the fact that males tend to control and inherit wealth (*i.e.* the land or the herds), hence leading to higher female migration rates [49], [53], [54], [55], [56], [57], [58], [59]. Given that forager communities do not accumulate wealth, migration patterns are more likely to be symmetrical, and this is indeed what has been observed. In other words, sedentism that accompanied the Neolithic transition [60] is expected to have led to a decrease in male gene flow, whereas female gene flow would either have remained constant or would have increased to compensate the decrease in male gene flow. This would explain two of our results, namely the higher mtDNA diversity, the higher NRY differentiation, and the higher difficulty found by several authors to identify clines in mtDNA data, compared to NRY. Interestingly, this would also be in agreement with the larger coalescent times described for mtDNA compared to NRY [43], [44] and would partly explain the results and interpretation of Richard *et al.* [5].

Another cultural change that is thought to have taken place in Europe during the Neolithic transition is a shift from polygyny to monogamy [61], [62]. In fact, several Neolithic burials [55], [58] show evidence of nuclear families, which may

reflect a monogamous marriage system. A shift from polygyny to monogamy would have the effect of decreasing male variance in reproductive success, since more males would now be able to mate, and consequently would increase N_m . This could result in a signal of population growth in NRY data that would be more recent compared to that observed in mtDNA and is exactly what Dupanloup and colleagues [63] have argued and found. Our results are in good agreement with theirs. Indeed, we found that t_h increased in males but not in females as we moved away from the Near East (Figure 1C), with t_h being the ratio of T , the time since the admixture event, and N_h , the effective size of the admixed population. Given that T necessarily decreases as we move away from the Near East, an increase of this ratio suggests that the decrease of T was compensated by a rapid increase in N_h . In other words, the admixture process between HG and farmers led to a very rapid increase in the effective size of the male population whereas this increase was more limited in females. Indeed, a shift from polygyny to monogamy would have less influence on N_f , which would anyway be higher than that of males, due to their lower variance in reproductive success. Altogether, a model in which human societies began to adopt farming as a means of subsistence, with the correlated patrilocality and monogamy as a mating system, would be in agreement with all the results presented here, including the aDNA (for instance it was rather impressive to find that the most probable scenarios, independently inferred no significant growth in Palaeolithic females). It also allows us to put in a single picture, results from several genetic and anthropological studies.

While we claim that a more coherent picture emerges from our results, we cannot claim that other scenarios could not also explain the results. Many layers of complexities could be added. For instance, female hypergamy (*i.e.* the fact that lower social status women are more likely to mate with males from a higher status than the opposite) has been described in several human migration and colonization events [24], [25], [64], and it is believed that it probably happened during the Neolithic transition in Europe [51], with HG females marrying into farmer communities [65]. Qualitatively, female hypergamy would increase female mobility and lead to low levels of mtDNA genetic differentiation between populations. Thus, one should expect lower mtDNA gradients and (almost) no geographic trend in drift, which is exactly what we see. The exclusion of HG males would lead to an increase of NRY genetic differentiation, explaining the clear geographic trend found in genetic drift. However, we must add that this scenario, which may indeed have taken place, would not as easily fit with the admixture patterns that we find and which are similar in males and females. Also, it does not fit with the recent strontium isotope data [38], [49]. Thus, at this stage, we would be cautious before arguing for or against female hypergamy. We also insist on the fact that the patterns identified here correspond to global patterns, and are not in contradiction with regional studies arguing against the demic diffusion. Several processes are likely to have taken place during the millennia corresponding to the arrival of farming communities in Europe. Similarly, it is increasingly clear that different routes (coastal or continental) were followed by different groups of humans. Still, the genetic data point to a major input from Near Eastern populations. This cannot be explained by cultural diffusion at a European scale and, as we have argued repeatedly, the general approach using the age of haplogroups or haplotypes to reconstruct human prehistory still awaits formal validation [3], [8], [11], [66], despite the large literature that uses it [10], [11], [40].

Our study represents the first attempt to integrate contemporary mtDNA and NRY data, together with aDNA. This has allowed us to draw a coherent picture of the Neolithic Transition in Europe, which not only provides an explanation for the patterns of genetic diversity found today and in our past, but also for the apparent contradiction between phylogeographic and model-based studies. The aDNA modelling approach described here could be applied to other aDNA datasets and we have applied it to data from an Iberian Neolithic population [67]. The results from these independent data appear to validate the suggestion that structured models with varying growth rates explain better the genetic distances observed between ancient and modern DNA than simpler models. The Neolithic transition in Europe is one of the most studied periods of human prehistory and the source of much debate. It is our hope that the work presented here may help provide a consistent framework to address certain aspects of this long-standing controversy.

Materials and Methods

Estimating Admixture/Interbreeding between Palaeolithic HG and Neolithic Farmers Using Extant Genetic Data

We applied a Bayesian full-likelihood method based on a simple admixture model that assumes that in a given moment in the past, an “admixed” population H (representing the European populations), is formed by members of two independent parental populations, P_1 and P_2 (representing HG and the farmers, respectively), whose contributions to H are p_1 and p_2 ($p_2 = 1 - p_1$), respectively (see Figure S1). After the admixture event, the three populations are isolated and assumed to evolve independently under pure genetic drift, represented by parameter $t_i = T/N_i$ (t_1, t_2 and t_h for populations P_1, P_2 and H , respectively). This method, already applied to the Neolithic Transition in previous works [8], [68], [69], is described in Chikhi *et al.* [33] and implemented in LEA [70] and ParLEA [71]. It has been shown that both the cultural and demic diffusion models can be seen as extreme cases of an admixture model, whereby two or more parental populations mixed in the past to produce the hybrid ancestors of present-day populations [1], [8]. Thus, in extreme cases of admixture, with no genetic contribution of one of the parental populations, we would expect that the gene pool of present-day populations is similar to the Mesolithic HGs, in the case of CDM, or to the Neolithic farmers, in the case of DDM.

aDNA and Coalescent Analysis

We used Bayesian Serial SimCoal software [72], [73] to simulate data, by tracing the ancestry of the female modern samples and incorporating aDNA samples of both HG and farmers, for each of the three models described in Figure 2. We explored 2,500 parameter combinations using fifty equally spaced values, sampled from the priors for both N_{UP} (ranging from 10 to 5,000) and N_N (between 1,000 and 100,000), that is using the same range as in [27].

The selection of the best demographic model was carried out under an ABC framework [32], [37]. The same approach was applied to estimate parameters for the selected model (SDG) [31], [32]. The validation of this procedure is fully described in SI Material Methods (see also table S3).

Further details regarding the admixture analysis in modern and ancient data and the data sets used in this study may be found in *Text S1*.

Supporting Information

Figure S1 Admixture model used by the Chikhi *et al.* [1] method. See SI Methods for more details and reference information.
(TIF)

Figure S2 Palaeolithic contribution to modern European (p_I) posterior distributions, for each of the European populations analysed, using NRY data [2]. Each curve corresponds to the analysis of a specific hybrid (admixed) population. In (A) are represented all the populations used in this study and in (B) are the populations used as negative control. See Text S1 for more details and reference information.
(TIF)

Figure S3 Linear regression of Neolithic contribution ($1-p_I$), against geographical distance from the Near East, using NRY data [2]. In (A) are represented all the populations used in this study and in (B) are the populations used as negative control. Mean values for each population are represented by red circles. See SI Methods for more details and reference information.
(TIF)

Figure S4 Linear regression of Neolithic contribution ($1-p_I$) against geographical distance from the Near East, using NRY data [2]. In (A) are represented the Caucasus populations (note the different scale on the x-axis) and in (B) are the European Islands population samples (Cyprus, Sardinian, UK and Ireland) used in this study. Mean values for each population are represented by red circles. See Text S1 for more details and reference information.
(TIF)

Figure S5 Distributions of the t_i 's for all populations, using NRY [2]. (A) Posterior distributions of t_1 . The different curves represent the amount of genetic drift, since the admixture event, between the present sample of Basques and the ancestral populations of HG that interbred with the incoming farmers. (B) Posterior distributions of t_2 . As in (A), but for the drift between the Near East and the first farmer populations. The colour codes are as in Figure S2A. See Text S1 for more details and reference information.
(TIF)

Figure S6 Distributions of the t_i 's for all populations, using mtDNA [3] (A) Posterior distributions of t_1 . (B) Posterior distributions of t_2 (see Figure S5 for a more detailed explanation). Note that the panel B has a different scale on the x-axis compared to panel A and Figure S5. See Text S1 for more details and reference information.
(TIF)

Figure S7 Estimated effective population sizes for the admixed populations (N_h) and their distance from the Near East. The N_h values were calculated for the Rosser *et al.* [2] dataset using archaeological dates from table S4 and a generation time of 25 years. See Text S1 for more details and reference information.
(TIF)

Figure S8 Genetic diversity and differentiation, across Europe. In (A), the H_e values for each European population analysed are regressed against the geographic distance from the Near East, both for NRY (solid circles) and mtDNA (open circles). The linear regressions calculated from these points are represented by the solid (NRY) and dashed (mtDNA) lines. In (B), each point

represents pairwise F_{ST} values, between European populations and the Near East, regressed against distance from the latter. The symbol and line codes are as in (A).
(TIF)

Figure S9 Split with differential growth model (SDG), with name of the demes.
(TIF)

Table S1 Demographic parameters estimated under the Split with Differential Growth (SDG) model. Weighted (ω) median, 5% and 95% percentiles values are represented for the N_e at the Neolithic and Upper Palaeolithic. Deme 1 and 2 correspond to the demes without and with differential growth, respectively (see Figure S9).
(PDF)

Table S2 Maximum probability of obtaining genetic differentiation (F_{ST}) values larger than those observed in the real data. Maximum probability values of obtaining a simulated F_{ST} value higher than that observed ($P_{s>o}$), for each of the models (TP - Total Panmixia, S - Split, SDG - Split with Differential Growth) and pairwise comparisons analysed (see Figure 3). See Text S1 for more details and reference information.
(PDF)

Table S3 Validation of the ABC model selection procedure. Each row corresponds to the percentage of times that a model (TP - Total Panmixia, S - Split, SDG - Split with Differential Growth) was assigned to each of the models, by a higher posterior probability. When data are simulated under the S model our results show that a significant proportion of the data sets are identified as being generated under another model (and as many as 44.7% are assigned to the TP model). This is less the case for the data generated under the TP model (but still they represent as much as 25% altogether) and even less under the SDG model. Thus despite non negligible error rates, these simulations suggest that there is a bias favouring the TP model, and much less the S and SDG models. One reason for this is that the ABC algorithm used here followed the procedure of Bramanti and colleagues [5], and was only based on three statistics, which were available. However, the results also show that the SDG model is the model which is most easily identified with nearly 88% of positive results. Given that the results obtained from the real data provide no support for the TPM, and less than 5% for the S model, we are confident that the inference of the model is unlikely to be incorrect hence demonstrating the importance of differential growth. This explains why Haak *et al.* [6] were unable to explain the observed F_{ST} values with their split model. See Text S1 for more details and reference information.
(PDF)

Table S4 Calibrated radiocarbon dates of Neolithic archaeological sites (from Pinhasi *et al.* [4]). Location and type of Neolithic culture (EN- Early Neolithic, LBK- Linear Pottery Culture) are also represented in this table. See Text S1 for reference information.
(PDF)

Text S1 Support Materials and Methods and references to SI Tables and SI Figures.
(PDF)

Acknowledgments

The authors are grateful to C. van Schaik, A. Coutinho, G. Gomes, V. Sousa, J. Salmons, J. Alves, S. Davis and two anonymous reviewers for useful comments on earlier versions of this manuscript. Simulations were

partly carried out using HPC resources from CALMIP (Toulouse, France) and from Instituto Gulbenkian de Ciência (IGC, Oeiras, Portugal). We thus thank G. Gomes and P. Fernandes, respectively, for allowing us to use the HULK simulation server and HERMES HPC Centre (FCT grant H200741/re-equip/2005) at IGC. E. Danchin, C. Thébaud and B. Crouau-Roy, A. Borelli are also thanked for their support.

References

1. Currat M, Excoffier L (2005) The effect of the Neolithic expansion on European molecular diversity. *Proc R Soc B* 272: 679–688.
2. Fagundes NJR, Ray N, Beaumont M, Neuenschwander S, Salzano FM et al. (2007) Statistical evaluation of alternative models of human evolution. *Proc Natl Acad Sci USA* 104: 17614–17619.
3. Goldstein DB, Chikhi L (2002) Human migrations and population structure: what we know and why it matters. *Annu Rev Genomics Hum Genet* 3: 129–152.
4. Mithen S (2007) Did farming arise from a misapplication of social intelligence?. *Philos Trans R Soc B* 362: 705–718.
5. Richards M, Macaulay V, Hickey E, Vega E, Sykes B et al. (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* 67: 1251–1276.
6. Richards M, Macaulay V, Torroni A, Bandelt H (2002) In search of geographical patterns in European mitochondrial DNA. *Am J Hum Genet* 71: 1168–1174.
7. Richards M (2003) The Neolithic invasion of Europe. *Annu. Rev. Anthropol.* 32: 135–162.
8. Chikhi L, Nichols RA, Barbujani G, Beaumont MA (2002) Y genetic data support the Neolithic demic diffusion model. *Proc. Natl. Acad. Sci. U. S. A.* 99: 11008–11013.
9. Bellwood P (2004) *First farmers: the origins of agricultural societies*. Oxford: Blackwell Publishing, 360 p.
10. Barbujani G, Chikhi L (2006) Population genetics: DNAs from the European Neolithic. *Heredity* 97: 84–85.
11. Chikhi L (2009) Update to Chikhi et al.'s "Clinal variation in the nuclear DNA of Europeans" (1998): genetic data and storytelling—from archaeogenetics to astrogenetics?. *Hum Biol* 81: 639–643.
12. Bocquet-Appel JP, Naji S, Linden MV, Kozłowski JK (2009) Detection of diffusion and contact zones of early farming in Europe from the space-time distribution of ¹⁴C dates. *J Archaeol Sci* 36: 807–820.
13. Pinhasi R, Fort J, Ammerman AJ (2005) Tracing the origin and spread of agriculture in Europe. *PLoS Biol* 3: e410.
14. Pinhasi R, von Cramon-Taubadel N (2009) Craniometric data supports demic diffusion model for the spread of agriculture into Europe. *PLoS One* 4: e6747.
15. Gkiasta M, Russell T, Shennan S, Steele J (2003) Neolithic transition in Europe: the radiocarbon record revisited. *Antiquity* 77: 45–62.
16. Dupanloup I, Bertorelle G, Chikhi L, Barbujani G (2004) Estimating the impact of prehistoric admixture on the genome of Europeans. *Mol Biol Evol* 21: 1361–1372.
17. Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S et al. (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 290: 1155–1159.
18. Ammerman AJ, Cavalli-Sforza LL (1984) *The Neolithic transition and the genetics of populations in Europe*. Princeton: Princeton University Press, 200 p.
19. Zvelebil M, Zvelebil KV (1988) Agricultural transition and Indo-European dispersals. *Antiquity* 62: 574–583.
20. Chikhi L, Destro-Bisol G, Bertorelle G, Pascali V, Barbujani G (1998) Clines of nuclear DNA markers suggest a largely Neolithic ancestry of the European gene pool. *Proc Natl Acad Sci U S A* 95: 9053–9058.
21. Balter M (2009) Archaeology. Ancient DNA says Europe's first farmers came from afar. *Science* 325: 1189.
22. Balaresque P, Bowden GR, Adams SM, Leung H, King TE et al. (2010) A predominantly Neolithic origin for European paternal lineages. *PLoS Biol* 8: e1000285.
23. Rosser ZH, Zerjal T, Hurler ME, Adojaan M, Alavantic D et al. (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67: 1526–1543.
24. Quintana-Murci L, Quach H, Harmant C, Luca F, Massonnet B et al. (2008) Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. *Proc Natl Acad Sci USA* 105: 1596–1601.
25. Salzano FM (2004) Interethnic variability and admixture in Latin America—social implications. *Rev Biol Trop* 52: 405–415.
26. Haak W, Forster P, Bramanti B, Matsumura S, Brandt G et al. (2005) Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science* 310: 1016–1018.
27. Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P et al. (2009) Genetic discontinuity between local hunter-gatherers and Central Europe's first farmers. *Science* 326: 137–140.
28. Malmström H, Gilbert MTP, Thomas MG, Brandström M, Storå J et al. (2009) Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. *Curr Biol* 19: 1758–1762.
29. Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V et al. (2010) Ancient DNA from European early Neolithic farmers reveals their Near Eastern affinities. *PLoS Biol* 8: e1000536.
30. Rasteiro R, Bouttier P, Sousa VC, Chikhi L (2012) Investigating sex-biased migration during the Neolithic transition in Europe, using an explicit spatial simulation framework. *Proc Biol Sci* 279: 2409–2416.
31. Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* 162: 2025–2035.
32. Blum MGB, François O (2009) Non-linear regression models for Approximate Bayesian Computation. *Stat Comput* 20: 63–73.
33. Chikhi L, Bruford MW, Beaumont MA (2001) Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. *Genetics* 158: 1347–1362.
34. Sousa VC, Fritz M, Beaumont MA, Chikhi L (2009) Approximate Bayesian Computation without summary statistics: the case of admixture. *Genetics* 181: 1507–1519.
35. Renfrew C (1991) Before Babel: Speculations on the Origins of Linguistic Diversity. *Camb Archaeol J* 1: 3–23.
36. Gray RD, Atkinson QD (2003) Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426: 435–439.
37. Beaumont M (2008) Joint determination of topology, divergence time, and immigration in population trees. In: Matsumura S, Forster P, Renfrew C, editors. *Simulation, genetics, and human prehistory*. Cambridge: McDonald Institute for Archaeological Research, 135–154.
38. Boric D, Price TD (2013) Strontium isotopes document greater human mobility at the start of the Balkan Neolithic. *Proc Natl Acad Sci U S A* 110: 3298–3303.
39. Balanovsky O, Dibirova K, Dybo A, Mudrak O, Frolova S et al. (2011) Parallel evolution of genes and languages in the Caucasus region. *Mol Biol Evol* 28: 2905–2920.
40. Barbujani G, Bertorelle G, Chikhi L (1998) Evidence for Paleolithic and Neolithic gene flow in Europe. *Am J Hum Genet* 62: 488–492.
41. Chikhi L, Beaumont MA (2005) Modelling human genetic history. In: Dunn MJ, Jorde LB, Little PFR, Subramaniam S, editors. *The encyclopedia of genetics, genomics, proteomics and bioinformatics*. New York: John Wiley & Sons.
42. Wilkins JF (2006) Unraveling male and female histories from human genetic data. *Curr Opin Genet Dev* 16: 611–617.
43. Tang H, Siegmund DO, Shen P, Oefner PJ, Feldman MW (2002) Frequentist estimation of coalescence times from nuclearotide sequence data using a tree-based partition. *Genetics* 161: 447–459.
44. Wilder JA, Mobasher Z, Hammer MF (2004) Genetic evidence for unequal effective population sizes of human females and males. *Mol Biol Evol* 21: 2047–2057.
45. Shennan S (2009) Evolutionary demography and the population history of the European early Neolithic. *Hum Biol* 81: 339–355.
46. Galeta P, Bruzek J (2009) Demographic model of the Neolithic transition in Central Europe. *Documenta Praehistorica* 36 (Neolithic Studies 16): 139–150.
47. Bocquet-Appel J, Demars P, Noiret L, Dobrowsky D (2005) Estimates of Upper Palaeolithic meta-population size in Europe from archaeological data. *J Archaeol Sci* 32: 1656–1668.
48. Gignoux CR, Henn BM, Mountain JL (2011) Rapid, global demographic expansions after the origins of agriculture. *Proc Natl Acad Sci U S A* 108: 6044–6049.
49. Bentley RA, Bickle P, Fibiger L, Nowell GM, Dale CW et al. (2012) Community differentiation and kinship among Europe's first farmers. *Proc Natl Acad Sci U S A* 109: 9326–9330.
50. Price TD, Bentley RA, Luning J, Gronenborn D, Wahl J (2001) Prehistoric human migration in the Linearbandkeramik of Central Europe. *Antiquity* 75: 593–603.
51. Bentley RA, Chikhi L, Price TD (2003) The Neolithic transition in Europe: comparing broad scale genetic and local scale isotopic evidence. *Antiquity* 77: 63–66.
52. Bocquet-Appel J (2002) Paleoanthropological traces of a Neolithic Demographic Transition. *Curr Anthropol* 43: 637–650.
53. Baker M, Jacobsen J (2006) A Human capital-based theory of postmarital residence rules. *J Law Econ Organ* 23: 208–241.
54. Bentley RA, Price TD, Luning J, Gronenborn D, Wahl J et al. (2002) Human migration in early Neolithic Europe. *Curr Anthropol* 43: 799–804.
55. Bentley RA, Wahl J, Price TD, Atkinson TC (2008) Isotopic signatures and hereditary traits: snapshot of a Neolithic community in Germany. *Antiquity* 82: 290–304.
56. Cavalli-Sforza LL, Minch E (1997) Paleolithic and Neolithic lineages in the European mitochondrial gene pool. *Am J Hum Genet* 61: 247–254.

Author Contributions

Conceived and designed the experiments: RR LC. Performed the experiments: RR LC. Analyzed the data: RR LC. Wrote the paper: RR LC.

57. Fortunato L, Jordan F (2010) Your place or mine? A phylogenetic comparative analysis of marital residence in Indo-European and Austronesian societies. *Philos Trans R Soc B* 365: 3913–3922.
58. Haak W, Brandt G, de Jong HN, Meyer C, Ganslmeier R et al. (2008) Ancient DNA, Strontium isotopes, and osteological analyses shed light on social and kinship organization of the Later Stone Age. *Proc Natl Acad Sci U S A* 105: 18226–18231.
59. Langergraber KE, Siedel H, Mitani JC, Wrangham RW, Reynolds V et al. (2007) The genetic signature of sex-biased migration in patrilocal chimpanzees and humans. *PLoS One* 2: e973.
60. Bellwood P, Oxenham M (2008) The expansions of farming societies and the role of the Neolithic demographic transition. In: Bocquet-Appel J, Bar-Yosef O, editors. *The neolithic demographic transition and its consequences*. Dordrecht: Springer. 13–34.
61. Lagerlöf N (2010) Pacifying monogamy. *J Econ Growth* 15: 235–262.
62. Fortunato L (2011) Reconstructing the history of marriage strategies in Indo-European-speaking societies: monogamy and polygyny. *Hum Biol* 83: 87–105.
63. Dupanloup I, Pereira L, Bertorelle G, Calafell F, Prata MJ et al. (2003) A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. *J Mol Evol* 57: 85–97.
64. Thomas MG, Stumpf MPH, Härke H (2006) Evidence for an apartheid-like social structure in early Anglo-Saxon England. *Proc Biol Sci* 273: 2651–2657.
65. Bentley RA, Layton RH, Tehrani J (2009) Kinship, marriage, and the genetics of past human dispersals. *Hum Biol* 81: 159–179.
66. Beaumont MA, Nielsen R, Robert C, Hey J, Gaggiotti O et al. (2010) In defence of model-based inference in phylogeography. *Mol Ecol* 19: 436–446.
67. Gamba C, Fernández E, Tirado M, Deguilloux MF, Pemonge MH et al. (2012) Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. *Mol Ecol* 21: 45–56.
68. Rasteiro R, Chikhi L (2009) Revisiting the peopling of Japan: an admixture perspective. *J Hum Genet* 54: 349–354.
69. Belle EMS, Landry P, Barbujani G (2006) Origins and evolution of the Europeans' genome: evidence from multiple microsatellite loci. *Proc R Soc B* 273: 1595–1602.
70. Langella O, Chikhi L, Beaumont MA (2001) LEA (likelihood-based estimation of admixture): a program to estimate simultaneously admixture and time since the admixture event. *Mol Ecol Notes* 1: 357–358.
71. Giovannini A, Zanghirati G, Beaumont MA, Chikhi L, Barbujani G (2009) A novel parallel approach to the likelihood-based estimation of admixture in population genetics. *Bioinformatics* 25: 1440–1441.
72. Anderson CNK, Ramakrishnan U, Chan YL, Hadly EA (2005) Serial SimCoal: a population genetics model for data from multiple populations and points in time. *Bioinformatics* 21: 1733–1734.
73. Excoffier L, Novembre J, Schneider S (2000) SIMCOAL: a general coalescent program for the simulation of molecular data in interconnected populations with arbitrary demography. *J Hered* 91: 506–509.