

This article has been published in *Molecular Biology and Evolution* by Oxford University Press

Accepted manuscript.

João Pimenta, Alexandra M Lopes, David Comas, António Amorim, Miguel Arenas, Evaluating the Neolithic Expansion at Both Shores of the Mediterranean Sea, *Molecular Biology and Evolution*, Volume 34, Issue 12, December 2017, Pages 3232–3242, <https://doi.org/10.1093/molbev/msx256>

General rights:

© The Author 2017. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.

Research article

Evaluating the Neolithic Expansion at Both Shores of the Mediterranean Sea

João Pimenta^{1,2,3,4}, Alexandra M. Lopes^{1,2}, David Comas³, António Amorim^{1,2,4} and Miguel Arenas^{1,2,5}

¹*Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal.*

²*Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal.*

³*Institute of Evolutionary Biology (CSIC-UPF). Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona, Spain.*

⁴*Faculty of Sciences, University of Porto, Porto, Portugal.*

⁵*Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain.*

Email addresses:

JP: jpimenta@ipatimup.pt

AL: alopes@ipatimup.pt

DC: david.comas@upf.edu

AM: aamorim@ipatimup.pt

MA: marenas@uvigo.es

Corresponding authors:

João Pimenta

Population Genetics & Evolution

Instituto de Investigação e Inovação em Saúde

University of Porto

Rua Alfredo Allen, 208

4200-135 Porto, Portugal

Email address: jpimenta@ipatimup.pt

Phone: +351 220 408 800

Miguel Arenas

Department of Biochemistry, Genetics and Immunology

University of Vigo

36310 Vigo, Spain

Email address: marenas@uvigo.es

Phone: (+34) 986 130050

Running head: Neolithic Expansion across the Mediterranean Sea

Keywords: Neolithic expansion, X chromosome, approximate Bayesian computation, Mediterranean Sea, Strait of Gibraltar

Abstract

During the Neolithic, human populations underwent cultural and technological developments that led to an agricultural revolution. Although the population genetics and evolution of European Neolithic populations have been extensively studied, little is known regarding the Neolithic expansion in North Africa with respect to Europe. One could expect that the different environmental and geological conditions at both shores of the Mediterranean Sea could have led to contrasting expansions. In order to test this hypothesis, we compared the Neolithic expansion in Europe and North Africa accounting for possible migration between them through the Strait of Gibraltar. We analysed the entire X chromosome of 580 individuals from 20 populations spatially distributed along the North of Africa and Europe. Next, we applied approximate Bayesian computation based on extensive spatially explicit computer simulations to select among alternative scenarios of migration through the Strait of Gibraltar and to estimate population genetics parameters in both expansions. Our results suggest that, despite being more technologically advanced, Neolithic populations did not expand faster than Palaeolithic populations, which could be interpreted as a consequence of a more sedentary lifestyle. We detected reciprocal Neolithic migration between the Iberian Peninsula and North Africa through the Strait of Gibraltar. Counterintuitively, we found that the studied Neolithic expansions presented similar levels of carrying capacity and migration, and occurred at comparable speeds, suggesting a similar demic process of substitution of hunter-gatherer populations. Altogether, the Neolithic expansion through both Mediterranean shores was not so different, perhaps because these populations shared similar technical abilities and lifestyle patterns.

Introduction

The Neolithic transition in Europe and North Africa started in the Near East around 10,000 years ago (ya) and generated a revolution in the living-style with the emergence of farmers after previous hunter-gatherer Palaeolithic populations (Zeder and Hesse 2000; Skoglund et al. 2012; Broushaki et al. 2016; Gallego-Llorente et al. 2016; Lazaridis et al. 2016; Meiklejohn et al. 2017). Two contrasting hypotheses were postulated for the expansion of Agriculture, the *demic diffusion* model (Ammerman and Cavalli-Sforza 1984; Sokal et al. 1991) and the *cultural diffusion* model (Fort 2012). The first one supports a migratory flux of Neolithic populations across Europe, which partially replaced the indigenous hunter-gatherer populations and implies gene flow between hunter-gatherers and farmers. The second model predicts the cultural adoption of the Neolithic lifestyle by neighbour populations with limited gene flow from Near Eastern populations (Sampietro et al. 2007; Fort 2012; Fu et al. 2012). Agriculture spread rapidly along the Southeast of Europe (Zeder 2008), leading to a pronounced cultural revolution, and left profound genetic marks in European populations (Currat and Excoffier 2005; Bramanti et al. 2009). In contrast to Europe, the *tempo and mode* of the Neolithic expansion throughout the North of Africa is still obscure, mostly because of a lack of archaeological and biological data (Barich 2014). Most evidence of the presence of domesticated species in North Africa dates around 8,000 ya and older (Linseele et al. 2014; Madella et al. 2014), suggesting an early Neolithic presence in the region, which is in agreement with the Mediterranean expansion route (Zeder 2008; Isern et al. 2017). Beforehand, one could predict that the transition to a farmer lifestyle in North Africa could have been slower than in Europe due to the arid conditions of the Sahara desert. It is also noteworthy that the region suffered frequent and drastic climatic changes during the Palaeolithic, with wet and arid periods (Adkins et al. 2006; McGee et al. 2013), which could have caused major demographic and cultural shifts (Clarke et al. 2016). However, North Africa experienced a long African Humid Period (AHP) between 12,000 and 5,000 ya (Adkins et al. 2006; Dunne et al. 2012; McGee et al. 2013), coincident with the Neolithic expansion period.

Importantly, the Mediterranean Sea was a geographical barrier to migration between Neolithic populations of North Africa and Europe, generating great cultural differences between both regions, exemplified by the expansion of Indo-European languages in Europe (Haak et al. 2015) and the expansion of Afro-Asiatic languages in North Africa (Ruhlen 1991). However, the geographical proximity at the westernmost region of both continents (Strait of Gibraltar) could have provided a migration corridor between the two shores (Bosch et al. 2001; González-Pérez et al. 2003; Rhouda et al. 2009; Botigué et al. 2013), although still little

is known about the specific direction and amount of migration through this Strait during the Neolithic.

Here we aim to evaluate and compare the Neolithic expansion on both shores of the Mediterranean Sea accounting for the role of the Strait of Gibraltar as a potential migration corridor between them. Using an approximate Bayesian computation (ABC) method based on extensive spatially explicit computer simulations, we investigated the population genetics and evolution of the Neolithic expansion in Europe and North Africa. First, we selected the evolutionary scenario of migration through the Strait of Gibraltar that fitted best with the real data and second, under the previously selected evolutionary scenario, we estimated population genetics parameters for the Neolithic expansion at each shore of the Mediterranean Sea. For this purpose, we built an X-chromosome dataset that combines data from several SNP dataset panels and includes samples from Europe, North Africa, Middle East and Sub-Saharan Africa. The use of the X-chromosome SNPs allowed us to perform an ABC analysis in a computational cost-benefit manner, and avoids errors of haplotype reconstruction since phasing is directly inferred in male individuals.

Material and Methods

Data and Genotyping

A dataset composed by 20 populations with 846 individuals (580 males and 266 females) was built with published data (Figure 1 and Table S1; Supplementary Material). The dataset included 11 European populations, Finland (FIN), Northern and Western Europeans currently in Utah (CEU), Great-Britain (GBR), Tuscany (TSI), East Spain (ESP) and Central Spain (CSP) populations retrieved from 1,000 Genomes Phase 3 (The 1000 Genomes Project Consortium 2012); Basque (BAS), Andalusia (AND) and Galicia (GAL) populations retrieved from Henn et al. (2012) and 1,000 Genomes Phase 3; and Porto (POR) and Lisbon (LIS) populations retrieved from Lopes et al. (2013). It also included 7 North African populations retrieved from Henn et al. (2012) and a Sub-Saharan [Yoruba (YRI)] populations from 1,000 Genomes Phase 3. Finally, a Syrian population retrieved from Arauna et al. (2017).

A quality control filter was applied to the X chromosome data using PLINK 1.07 (Purcell et al. 2007) and SNPs with missing rate higher than 10%, SNPs that failed the Hardy-Weinberg equilibrium under a threshold of 0.05 and SNPs with Minor Allele Frequency (MAF) lower than 0.05 were excluded to avoid biases. In addition, individuals with a missing rate higher than 10% and those with an identity by state, for the autosomes, larger than 85%, were also excluded. Moreover, SNPs within both pseudoautosomal regions (PAR1 and PAR2) and those

that were heterozygous in the X specific region in male individuals were also excluded. A total of 16,939 SNPs were maintained to proceed with the phasing step.

Phasing

As a sanity check, the dataset was phased using the SHAPEIT2 software (Delaneau et al. 2013), the HapMap PhaseII b37 (The International Hapmap Consortium 2003) with the genetic map and the 1000 genomes dataset as the reference panel (The 1000 Genomes Project Consortium 2012). The dataset was phased after an alignment with the reference panel and the removal of SNPs that did not align. Next, all genotypes were phased simultaneously with a total of 100 iterations, half of them for burn-in and pruning stages. A total of 16,917 SNPs were kept for subsequent analyses. For the ABC analysis, only male samples were considered, since X chromosome haplotypes can be directly (without biases) identified.

Spatially explicit computer simulations

Palaeolithic and Neolithic expansions were simulated with the evolutionary framework SPLATCHE2 (Ray et al. 2010). SPLATCHE2 is a well-established spatially explicit computer simulator that has been widely used to investigate human evolution (Benguigui and Arenas 2014). This program is based on three major steps. The first one consists of a forward-in-time simulation of the whole population under demographic parameters such as population size at the onset of the simulation, population growth rate and number of generations. This demographic simulation is performed over a two-dimensional (2D) landscape (grid of demes) that considers a carrying capacity per deme (number of individuals that can be sustained by the resources of the deme) and where individuals can move between demes according to a migration rate and under a 2D stepping-stone migration model (Kimura and Weiss 1964). The landscape used for this study was obtained by a geographic information system (GIS) and includes 4,814 demes with deme size 150×150 km following Alves et al. (2016) (Figure 1). A land bridge was added to connect Great Britain with mainland Europe, allowing the colonization of these islands. The second step is a backward-in-time (coalescent) simulation that generates the evolutionary history of a user-specified sample accounting for the history of the whole population stored in the previous step (further details in Currat et al. 2004). In the third step, the program simulates the genetic data of the sample. To perform this task, a randomly selected sequence is assigned to the root node of the previously simulated coalescent history and is evolved going forward-in-time over such coalescent history, according to a substitution model of evolution, to generate genetic sequences for the sampled nodes (Yang 2006; Arenas 2012). Altogether, the simulated genetic data is

influenced by the user-specific population genetics parameters (i.e., large population size can lead to high genetic diversity) (Ray et al. 2010).

We performed computer simulations following settings described in recent studies (François et al. 2010; Arenas et al. 2013; Alves et al. 2016). The Palaeolithic expansion was assumed to start from the Ethiopian region (Pritchard et al. 2000) at T_{START} generations ago with an ancestral population size of N_{ANC} haploid individuals. At each generation, individuals migrated to the adjacent demes at a migration rate M_{PALEO} and the population density for each deme was determined by the population growth rate G_{PALEO} and the carrying capacity K_{PALEO} (Currat et al. 2004). A range contraction was simulated to mimic the *last glacial maximum* (LGM) period in Africa and Eurasia described by Alves et al. (2016). Briefly, this range contraction consisted of 30 consecutive events that progressively reduced the habitable regions (i.e., at the North of Europe or increasing deserts size; Figure S1, Supplementary Material) of human settlements (Arenas et al. 2012). The range contraction started at 25,000 ya and took 3,000 years to reach its maximum. Next, individuals were only living at the refugia (South of Europe, North Africa and South of Africa; see Figure S1) during 4,000 years and finally re-expanded at 18,000 ya (Alves et al. 2016).

In addition, we simulated a Neolithic expansion that started from the Middle East (Zeder 2008) at T_{NEO} generations ago and with an ancestral size at the onset of the expansion of N_{ANCNEO} haploid individuals and a population growth rate G_{NEO} . Here, the map was divided into three different geographic regions: Europe, North Africa and peripheral regions (i.e., Arabian Peninsula and Sub-Saharan Africa) (Figure 2A). Each of these regions had a specific migration rate (M_{NEOEUR} , M_{NEONA} and M_{NEOPER}) and carrying capacity (K_{NEOEUR} , K_{NEONA} and K_{NEOPER}), actually these are the only parameters implemented in SPLATCHE2 with spatio-temporal variation. The contribution of the Palaeolithic to the final genetic pool was considered with the parameter interbreeding rate $Irate$ (Barbujani et al. 1995; Currat and Excoffier 2004; Currat and Excoffier 2005; Currat et al. 2008; François et al. 2010; Currat and Excoffier 2011; Arenas et al. 2013), which is assumed to be spatio-temporal invariable in the simulator.

All the parameters were drawn from Uniform prior distributions (Table S2; Supplementary Material) based on previous studies (Arenas et al. 2013; Alves et al. 2016) and were estimated with approximate Bayesian computation (ABC; details below).

Next, based on the previously simulated demographic history, we simulated genetic data for the sample of 580 male individuals of the 20 populations. Each individual included an X chromosome sequence of 16,917 SNPs divided in 14 independent blocks (Table S3; Supplementary Material). These blocks were determined by the presence of particular regions without recombination based on published data (Tillmar et al. 2008) and based on estimated recombination hotspots for the CEU, POR and YRI populations (chosen arbitrarily)

with the program LDhat 2.2 (McVean et al. 2004). LDhat was run following the recommendations provided in the software documentation. We applied a sliding window of 2,000 SNPs with an overlap of 500 SNPs and performed 6 independent runs (2 per population) with 10 million iterations each (Figure S2A; Supplementary Material). We specified 2,000 iterations per sampling and a block penalty of 20 for changes in recombination rate ρ ($\rho=4N_e r$). Following the documentation, 10% of iterations were discarded as burn-in. Convergence between the independent runs was assessed considering a recombination rate threshold of 20, however most SNPs presented a difference lower than 1 indicating a very high convergence (Figure S2BCD). To perform the computer simulations, the MAF value was set to be not lower than 0.05 for every block, in line with the SNP filter applied to the real dataset, as described above.

Alternative Scenarios

In order to properly study the Neolithic expansion over Europe and over North Africa, we explored genetic signatures of migration through the Strait of Gibraltar during the Neolithic, assuming absence or little migration between these regions before that period (e.g., Timmermann and Friedrich 2016). Although some studies suggested Palaeolithic migration through this Strait (Currat et al. 2010), other studies described or assumed lack of this migration (Ferembach 1985; Bocquet-Appel and Demars 2000; Liu et al. 2006; Timmermann and Friedrich 2016). We believe that if Palaeolithic migration through the Strait took place, the corresponding genetic signatures could be insignificant due to the small number of migrants and evolutionary processes like genetic drift. Next, Neolithic migration through the Strait of Gibraltar could affect Neolithic populations at both shores of the Mediterranean Sea by exchanging genetic material between them. Therefore, first, we investigated the fitting of alternative evolutionary scenarios to the real data with ABC (details below). One of these scenarios considered migration in both directions between Iberian Peninsula and North Africa (*fullMIG*, Figure 2Ba), with migration from North to South and from South to North through the Strait of Gibraltar with rates M_{nsSG} and M_{snSG} , respectively. The other scenario considered lack of migration (*noMIG*, Figure 2Bb).

Next, since *fullMIG* better fitted with the real data (see Results), we designed two additional scenarios to compare both directions of migration through the Strait of Gibraltar. Both scenarios allowed migration but under only one direction (anisotropic migration): a scenario of migration from the Iberian Peninsula to North Africa (*nsMIG*, Figure 2Bc) at M_{nsSG} rate and, other scenario of migration from North Africa to the Iberian Peninsula (*snMIG*, Figure 2Bd) with rate M_{snSG} .

Finally, we investigated model selection under some additional combinations of evolutionary scenarios: (i) the four scenarios together (*fullMIG* vs *noMIG* vs *nsMIG* vs *snMIG*) and (ii) the scenario of bidirectional migration against each scenario of unidirectional migration (*fullMIG* vs *nsMIG* and, *fullMIG* vs *snMIG*). However here note that *fullMIG* includes a different prior distribution for the migration rate towards each direction, whose sampling can sometimes lead to situations (large anisotropic migration) near to *nsMIG* or *snMIG* and this must be considered in the analysis of these scenarios.

Approximate Bayesian Computation for Selection of Evolutionary Scenarios and Parameter Estimation

The real data was analysed with ABC (Beaumont et al. 2002; Beaumont 2010), a statistical approach to obtain posterior probabilities in the selection among alternative scenarios and the estimation of parameters. Next, we describe the different steps of our ABC method.

Computer Simulations. For each evolutionary scenario (of the 4 above described), we performed a total of 100,000 spatially explicit computer simulations according to the prior distributions described in Table S2 (Supplementary Material).

Summary Statistics. Summary statistics (SS) from real and simulated datasets were obtained with ARLEQUIN ver.3.5 (Excoffier and Lischer 2010). To perform this task, the 20 populations were classified into eight geographical groups (described in Figure 1) that allowed the estimation of diverse group-based SS. This grouping is important because the ABC performance reduces quickly as the number of statistics grows (Beaumont 2010). The computed SS included: number of alleles (k), heterozygosity (H), number of segregating sites (s) and, F_{ST} between groups and populations. Next, we selected a small set (36) of the most informative SS for the estimations (Table S4; Supplementary Material). This was done by testing the contribution of the different SS to the estimations from 100 simulated datasets (hereafter *test datasets*) that are independent of the 100,000 simulated datasets used to generate the ABC method. We did not apply partial least square (PLS) components to reduce the number of SS (Wegmann et al. 2009) because our selected SS led to accurate estimations and because PLS components can generate SS without biological meaning and thus being difficult to interpret, and also can generate biases when comparing components that were separately computed (Aeschbacher et al. 2012).

Selection of Evolutionary Scenarios. As noted above, we generated 100,000 simulated datasets for each of the 4 migration scenarios. In addition, we simulated 100 *test datasets* for each

scenario to evaluate the accuracy of the method. We applied 4 ABC methods to perform the selection of the best-fitting evolutionary scenario to the real data: the *rejection* method developed by Pritchard et al. (1999); and the methods *rejection*, *multinomial logistic regression* and *neural network*, implemented in the *abc* package of R (Csilléry et al. 2012). Following the authors recommendation we retained 1% of simulations that were closer to the observed data (Csilléry et al. 2012). Before analysing the real data, the accuracy of the ABC methods was assessed with the *test datasets* (considering them as *true* data) and by cross-validation based on 100 permutations and a tolerance of 1% (following the recommendations described in the package *abc* (Csilléry et al. 2012)). Although traditionally the goodness of fit analysis is only applied to the preferred scenario, we performed goodness of fit analyses (based on principal component analysis (PCA) and histogram of null distributions from the selected SS (Csilléry et al. 2012)) to the four investigated evolutionary scenarios.

Parameters Estimation. We performed the parameters estimation under the evolutionary scenario (*fullMIG*), which is the scenario that fitted best with the real data. We applied the *rejection* method implemented in the *abc* package (which generated the most accurate inferences for the studied dataset) by retaining 1,000 simulated datasets of the closest simulations. The robustness of the parameter estimation was assessed with the 100 *test datasets* simulated under the *fullMIG* scenario, considering them as *true* data. Then we computed the distance between the true value and the estimated value (median, mean and mode) from each of the 100 *test datasets*.

Results

We first describe the selection between alternative evolutionary scenarios of migration through the Gibraltar strait. After that, we present the parameter estimation under the best fitting evolutionary scenario, with a focus on those parameters that are specific for the Neolithic expansion over each shore of the Mediterranean Sea.

Migration through the Strait of Gibraltar occurred in Both Directions

The selection of the evolutionary scenario (*fullMIG* vs *noMIG*) presented very high posterior probabilities to identify the correct scenario. We found that all the applied ABC methods were able to identify the correct scenario with probabilities that ranged from 0.78 (*rejection* method implemented in the *abc* package) to 0.98 (*multinomial logistic regression* and *neural network* methods) (Table S5; Supplementary Material).

Next, real data clearly fitted better with the *fullMIG* scenario (posterior probabilities around 0.99 under all the ABC methods excepting the rejection method of *abc* package (0.85) (Figure S3A). Goodness of fit analyses also indicated this finding (Figure S4; Supplementary Material). Altogether the results suggest reciprocal migration through the Strait of Gibraltar.

Given that the real data fitted better with the presence of migration, we next investigated the direction of the migration through the Strait of Gibraltar. The validation of the selection of the evolutionary scenario (*nsMIG* vs *snMIG*) also presented very high posterior probabilities to identify the correct scenario. We found that all the ABC methods performed acceptably to recognize the correct scenario with probabilities that ranged from 0.85 (*rejection* method implemented in the *abc* package) to 1.00 (*Pritchard's rejection* method) (Table S6; Supplementary Material).

Next, the fitting of real data with both evolutionary scenarios, *nsMIG* and *snMIG*, suggested migration in both directions, although with a higher probability for the *nsMIG* scenario (Figure S3B). The results varied among ABC methods but led to the same overall conclusion.

Next, we describe the results (evaluation of model selection and analysis of real data) from the two additional comparisons of evolutionary scenarios, (i) the four evolutionary scenarios together (*fullMIG* vs *noMIG* vs *nsMIG* vs *snMIG*) and (ii) the evolutionary scenario of bidirectional migration against each evolutionary scenario of unidirectional migration (*fullMIG* vs *nsMIG* and, *fullMIG* vs *snMIG*).

Concerning the evaluation of the model selection, when the four scenarios are evaluated together (Table S7; Supplementary Material), the scenario *noMIG* was recognized with the highest probability (reaching 0.85) when applying the *multinomial logistic regression* and the *neural network* methods, while the three other scenarios presented lower probabilities (which is expected since *fullMIG* can sometimes be near to *nsMIG* or *snMIG*, see Materials and methods) but showed the highest probability with the correct scenario. Concerning comparisons between the scenario of bidirectional migration with each scenario of unidirectional migration, we found that all the ABC methods performed acceptably to recognize the correct scenario with probabilities between 0.65 and 0.85 (Tables S8 and S9; Supplementary Material). Again despite of *fullMIG* can sometimes be near to *nsMIG* or *snMIG* (Materials and methods).

Concerning the fitting with real data, the highest probability corresponded to the *fullMIG* scenario, as expected followed by the scenarios with unidirectional migration since these scenarios are very near to the *fullMIG* scenario (Figure 3, see also Figure S4). We found

similar results when comparing the scenario of bidirectional migration with each scenario of unidirectional migration (Figures S5 and S6; Supplementary Material).

Similar Neolithic Expansion at Both Shores of the Mediterranean Sea

The estimation of the population genetics parameters was validated with the *tests datasets* of the best-fitting evolutionary scenario (*fullMIG*). These estimations presented a performance with overall small estimation errors that fell within the 50% HPDI with respect to the true value (Table S10; supplementary material), although for some parameters related with very ancestral populations (i.e., ancestral Palaeolithic population size) the precision was low, as expected and also found in other studies (Wegmann and Excoffier 2010; Alves et al. 2016). The estimated population genetics parameters for the real data are shown in Table 1 and in Figure S7 (Supplementary Material). Again, some parameters (especially those related with Palaeolithic populations and ancestral population sizes) presented wide posterior distributions, which is expected due to loss of very ancestral genetic signatures (Fagundes et al. 2007; Wegmann and Excoffier 2010; Alves et al. 2016). However we preferred to account for this variation instead of fixing parameters' values since previous studies disagree on specific parameters values (i.e., Benguigui and Arenas 2014). Regarding the Palaeolithic population, our estimated date for the onset of the expansion of Anatomically Modern Humans (AMH) in Africa is similar (≈ 100 kya) to that obtained by Alves et al. (2016), Voight et al. (2005) and Fagundes et al. (2007), and like in those studies we also obtained a wide posterior distribution. The estimated ancestral size ($\approx 15,000$) was slightly larger but still in agreement with previous analyses based on other genetic markers and also with wide posterior distributions (Fagundes et al. 2007; Alves et al. 2016). The estimation of the onset of the out-of-Africa expansion (≈ 70 kya but with a wide posterior distribution) is very similar to that obtained by Alves et al. (2017) and within the intervals obtained by Fagundes et al. (2007). The carrying capacity estimated for the Palaeolithic population ($\approx 1,500$) was slightly larger than in other studies (Hamilton et al. 2007; Deshpande et al. 2009) but falls within the estimations by Alves et al. (2016). The population growth rate per generation (≈ 0.70) fell within the range estimated in other studies for hunter-gatherer populations (between 0.2 and 0.9) (Alves et al. 2016). The migration rate for Palaeolithic populations (≈ 0.23) also fell within the range estimated by Alves et al. (2016). The estimated interbreeding rate was ≈ 0.011 , which corresponds to approximately 71% of admixture between Palaeolithic and Neolithic populations (obtained with a polynomial regression based on Arenas et al. (2013)) and falls within the range described in other studies (Bramanti et al. 2009; Lazaridis et al. 2014; Haak et al. 2015; Lazaridis et al. 2016).

Regarding the Neolithic population, the estimated time for the onset of the Neolithic expansion (around 10 kya) was in agreement with previous findings (Goring-Morris and Belfer-Cohen 2011). As expected, the estimated Neolithic ancestral size (around 23,000) and carrying capacity ($\approx 6,200$) were higher than for the Palaeolithic population (Table 1). However, the population growth rate and migration rate for Neolithic and Palaeolithic populations were similar. We did not find significant differences (at the 95% HPDI) between the migration rate from the Iberian Peninsula to the North of Africa and the migration rate in the opposite direction. Indeed and as expected, the migration rate through the Strait (≈ 0.15) was much lower than the migration rate in other regions, which is indicative of the barrier to gene flow generated by the Strait. Importantly, the Neolithic migration rate through Europe and through the North of Africa presented similar levels (≈ 0.23). Indeed, the Neolithic carrying capacity for European and North African populations also exhibited similar levels (Table 1).

Discussion

The change from the hunter-gatherer lifestyle to the more sedentary farmers of the Neolithic was a major transition in human history (Bocquet-Appel 2011). The Near Eastern Neolithic revolution expanded to several regions, including Europe and North of Africa. However, little is known about the Neolithic expansion in North of Africa and its comparison with the expansion in Europe. Here, we evaluate the Neolithic progression in Europe, a region where the expansion from Anatolia was fast and agriculture was adopted in $\approx 3,000$ years (Bocquet-Appel et al. 2012), and in North Africa, a region still poorly studied. To address this aim we analysed the entire X chromosome (16,917 SNPs) from 580 male individuals that belong to 20 populations from Europe, Africa and Near East.

Migration through the Strait of Gibraltar during the Neolithic is abundantly verified by archaeological and genetic data (Hervella et al. 2016), even in studies based on uniparental markers (Bosch et al. 2001; Gonzalez et al. 2003; González-Pérez et al. 2003; Cruciani et al. 2004; Ennafaa et al. 2009; Rhouda et al. 2009) and genome-wide data (Botigué et al. 2013; Arauna et al. 2017). Although we also detected migration through this strait, we found that the Strait of Gibraltar was a barrier to gene flow since the estimated migration rate through this strait was lower than the estimated migration rates in Europe or in the North of Africa (Table 1). Our results support a reciprocal gene flow through the Strait of Gibraltar, which suggests European influence on the genetic diversity of North African populations (and *vice versa*), increasing their heterogeneity. Concerning North Africa, this is in agreement with the recent study by Arauna et al. (2017), which showed that North African populations present,

in general, high levels of genetic heterogeneity and low geographical structure. Note that North Africa has a peculiar geographic landscape that contributed for its complex genetic diversity, limited by the Sahara desert and the Mediterranean Sea and connected to the Middle East through the Sinai and Arabian Peninsulas. Both the Sahara desert and the Mediterranean Sea operated as permeable barriers, not completely blocking the contact between North Africa and the rest of the African continent and Europe (Henn et al. 2012; Botigué et al. 2013; Arauna et al. 2017). Those migrations, together with the particular above described migration through the Strait of Gibraltar, played an important role on current genetic diversity of North African populations (Henn et al. 2012; Arauna et al. 2017). By contrast, Southern Europe might have presented a lower number of different migration routes, leading to a higher population structure (Hofmanová et al. 2016; Baker et al. 2017).

The parameter estimation showed that Neolithic migration and carrying capacity in Europe and North Africa were very similar, even presenting comparable speeds of expansion on both regions. This finding fits well with the existence of: *i*) maritime migrations that could potentiate a regular contact between both shores (Fernández et al. 2014; Paschou et al. 2014) and *ii*) favourable climatic conditions in both Southern Europe and North Africa that facilitated the adoption of the farmer lifestyle (Dunne et al. 2012; Manning and Timpson 2014; Olalde et al. 2015). Archaeological data shows that the introduction of the Neolithic in the Iberian Peninsula involved a pioneer expansion of the first farmers by sea (Zilhão 2001) and that this rapid expansion most likely occurred through long-distance dispersal (LDD) events along the coast (Isern et al. 2017). Furthermore, the archaeological evidence in North Africa and the Iberian Peninsula suggests a simultaneous arrival of the Neolithic by seafarers based on similar pottery and manufacturer techniques in both regions (Cortés Sánchez et al. 2012; Linstadter et al. 2012). Also, these results suggest a regular contact between western Mediterranean populations, which could be favoured by their geographical proximity (through the Strait of Gibraltar) and are in agreement with our selected evolutionary scenario of reciprocal migration between both shores.

Regarding the speed of the Neolithic range expansion, we performed a set of computer simulations based on the estimated posterior distributions (within the 95% HPDI) to estimate the dates of Neolithic arrival to the Iberian Peninsula and to perform comparisons with other studies (Isern et al. 2017). We found that the timing based on empirical data (Cortés Sánchez et al. 2012; Linstadter et al. 2012; Bernabeu Aubán et al. 2015; Isern et al. 2017) fell within our estimated time intervals, which indeed may provide an additional validation of our models. Next, our models generated a similar time of arrival of Agriculture

at the Iberian Peninsula and Western North Africa, which is in agreement with radiocarbon dates of Neolithic plants found in both regions (Morales et al. 2013; Isern et al. 2017).

The transition from hunter-gatherer to farmer lifestyle is still shrouded in controversy with two main hypotheses proposed for the way it took place: cultural or demic diffusion. Archaeological and classical genetic analyses of the European Neolithic expansion suggested a more influential role of the demic diffusion at the continental level (Cavalli-Sforza et al. 1993; Cavalli-Sforza et al. 1995), while the cultural diffusion had a secondary role, despite possible regional variations (Fort 2012; Rasteiro and Chikhi 2013; Fort et al. 2015). Recent whole-genome data (Lazaridis et al. 2014; Lazaridis et al. 2016) reinforces the idea of the Neolithic demic diffusion from the Near East and differential admixture with hunter-gatherers in Europe. The spread of the Neolithic in Europe was estimated to be around 1.0 km per year (Pinhasi et al. 2005), which is in agreement with predictions for the demic diffusion model. Our results suggest a similar scenario for the Neolithic expansion in North Africa considering that this region presented a similar migration rate to the corresponding one in Europe (Table 1). However, despite our estimates of a similar general velocity of the Neolithic front-wave in Europe and in North Africa, the Neolithic transition of Northern African populations could have been more heterogeneous. Archaeological evidence denotes a longer period of coexistence of hunter-gatherers and farmers in North Africa than in the Iberian Peninsula, which could be attributed to harsher environmental conditions for crops and cattle production and could lead to a slower Neolithization process by cultural assimilation (Linstadter et al. 2012). Indeed, genetic data shows some discontinuity between both shores of the Mediterranean (Botigué et al. 2013) despite the regular contact between the two regions, at least since the emergence of the Neolithic, as indicated by our results and archaeological studies (Linstadter et al. 2012; Mulazzani et al. 2016). Our results follow those based on archaeological data (in terms of timing), which might suggest that the Neolithic cultural component was more important for the contact between both shores than a genetic exchange.

Interestingly, we did not find significant differences between migration rates of Palaeolithic and Neolithic populations despite the larger carrying capacities and ancestral sizes presented in the Neolithic. These results, which are in agreement with previous studies (e.g., Pinhasi et al. 2005), could be caused by a space-competition between Mesolithic and Neolithic populations leading to a slowdown of the front wave transition of the Neolithic (Fort and Méndez 1999; Isern and Fort 2010; Isern et al. 2012).

Although the X chromosome is usually ignored in population genetics studies, it encloses some useful features (in comparison to genetic markers on the autosomes, Y chromosome and mtDNA) that we took advantage while performing this study. The X chromosome has a lower recombination rate than the autosomes due to its particular inheritance pattern, with recombination in the X chromosome non-pseudoautosomal region only occurring in females, resulting in the presence of long chromosomal segments that share a common history. Furthermore, haplotypes can be directly inferred in male individuals. In contrast to the Y chromosome and mtDNA, genetic drift has a lower effect on X chromosome markers while having a larger number of loci with potential different histories (Schaffner 2004). Additionally, male haplotypes of a given generation are the result of meiotic recombination in females in the previous generation and half of the X chromosomes in females in the following generation, meaning that for studies involving a large number of generations, problems concerning sex-bias are mostly avoided indicating that results would not be substantially different using autosomal markers. Indeed, Goldberg and Rosenberg (2015) demonstrated that, under a model of constant migration rate with no sex-bias, the admixture fractions for both autosomes and X chromosome are very similar when considering a large number of generations. Interestingly, a recent study based on three X chromosome regions with little or no recombination where European populations were analysed (Delsler et al. 2017), showed evidence of pre-Neolithic expansions at 15-18 kya and diversity in one of the studied regions supported the existence of a Neolithic expansion around 10 kya. This is in line with our model for a post LGM recolonization of Europe and a Near East Neolithic expansion, respectively. In conclusion, our study and other recent studies (Goldberg and Rosenberg 2015; Delsler et al. 2017) showed that the X chromosome per se can provide a large amount of useful information for population genetics analyses aimed to dissect the complex genetic history of modern humans, not merely as a complement for other genetic markers.

Supplementary Material

Tables S1-S10 and figures S1-S7 are available in the Supplementary Material.

Acknowledgments

We would like to thank Lara R. Arauna (Universitat Pompeu Fabra) for technical advice with the datasets used. IPATIMUP integrates the i3S Research Unit, which is partially supported by FCT in the framework of the project “Institute for Research and Innovation in Health Sciences” (POCI-01-0145-FEDER-007274). JP, AML and MA and are funded by the Portuguese Government through the FCT fellowship SFRH/BD/97200/2013 and research contracts IF/01262/2014 and IF/00955/2014, respectively. MA was also supported by the “Ramón y

Cajal” grant RYC-2015-18241 from the Spanish Government. DC was supported by the Spanish MINECO grant CGL2013-44351-P.

Literature Cited

- Adkins J, deMenocal P, Eshel G. 2006. The “African humid period” and the record of marine upwelling from excess ^{230}Th in Ocean Drilling Program Hole 658C. *Paleoceanography* 21.
- Aeschbacher S, Beaumont MA, Futschik A. 2012. A novel approach for choosing summary statistics in approximate Bayesian computation. *Genetics* 192:1027–1047.
- Alves I, Arenas M, Currat M, Hanulova AS, Sousa VC, Ray N, Excoffier L. 2016. Long-distance dispersal shaped patterns of human genetic diversity in Eurasia. *Mol. Biol. Evol.* 33:946–958.
- Ammerman AJ, Cavalli-Sforza L. 1984. *The Neolithic Transition and the Genetics of Populations in Europe*. (Press PU, editor.). Princeton University Press
- Arauna LR, Mendoza-Revilla J, Mas-Sandoval A, Izaabel H, Bekada A, Benhamamouch S, Fadhlouli-Zid K, Zalloua P, Hellenthal G, Comas D, et al. 2017. Recent historical migrations have shaped the gene pool of Arabs and Berbers in North Africa. *Mol. Biol. Evol.* 34:318–329.
- Arenas M. 2012. Simulation of molecular data under diverse evolutionary scenarios. *PLoS Comput. Biol.* 8.
- Arenas M, François O, Currat M, Ray N, Excoffier L. 2013. Influence of admixture and paleolithic range contractions on current European diversity gradients. *Mol. Biol. Evol.* 30:57–61.
- Arenas M, Ray N, Currat M, Excoffier L. 2012. Consequences of range contractions and range shifts on molecular diversity. *Mol. Biol. Evol.* 29:207–218.
- Baker KH, Gray HWI, Ramovs V, Mertzaniidou D, Akin Peksen Ç, Bilgin CC, Sykes N, Hoelzel AR. 2017. Strong population structure in a species manipulated by humans since the Neolithic: the European fallow deer (*Dama dama dama*). *Heredity (Edinb)*. 119:16–26.
- Barbujani G, Sokal RR, Oden NL. 1995. Indo-European origins - A computer simulation test of five hypotheses. *Am. J. Phys. Anthropol.* 96:109–132.
- Barich BE. 2014. Northwest Libya from the early to late Holocene: New data on environment and subsistence from the Jebel Gharbi. *Quat. Int.* 320:15–27.
- Beaumont MA. 2010. *Approximate Bayesian Computation in Evolution and Ecology*. Annu.

- Rev. Ecol. Evol. Syst. 41:379–406.
- Beaumont MA, Zhang W, Balding DJ. 2002. Approximate Bayesian computation in population genetics. *Genetics* 162:2025–2035.
- Benguigui M, Arenas M. 2014. Spatial and temporal simulation of human evolution. Methods, frameworks and applications. *Curr. Genomics* 15:245–255.
- Bernabeu Aubán J, Michael Barton C, Pardo Gordó S, Bergin SM. 2015. Modeling initial Neolithic dispersal. The first agricultural groups in West Mediterranean. *Ecol. Modell.* 307:22–31.
- Bocquet-Appel J-P. 2011. When the world's population took off: the springboard of the Neolithic Demographic Transition. *Science* 333:560–561.
- Bocquet-Appel J-P, Demars PY. 2000. Neanderthal contraction and modern human colonization of Europe. *Antiquity* 74:544–552.
- Bocquet-Appel JP, Naji S, Vander Linden M, Kozłowski J. 2012. Understanding the rates of expansion of the farming system in Europe. *J. Archaeol. Sci.* 39:531–546.
- Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J. 2001. High-Resolution Analysis of Human Y-Chromosome Variation Shows a Sharp Discontinuity and Limited Gene Flow between Northwestern Africa and the Iberian Peninsula. *Am. J. Hum. Genet* 68:1019–1029.
- Botigué LR, Henn BM, Gravel S, Maples BK, Gignoux CR, Corona E, Atzmon G, Burns E, Ostrer H, Flores C, et al. 2013. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc. Natl. Acad. Sci. U. S. A.* 110:11791–11796.
- Bramanti B, Thomas MG, Haak W, Unterlaender M, Jores P, Tambets K, Antanaitis-Jacobs I, Haidle MN, Jankauskas R, Kind C-J, et al. 2009. Genetic Discontinuity Between Local Hunter-Gatherers and Central Europe's First Farmers. *Science* (80-.). 326:137–140.
- Broushaki F, Thomas MG, Link V, López S, Dorp L Van, Kirsanow K, Hofmanová Z, Diekmann Y, Cassidy LM, Díez-del-molino D, et al. 2016. Early Neolithic genomes from the eastern Fertile Crescent. *Science* (80-.). 353:499–503.
- Cavalli-Sforza L, Menozzi P, Piazza A. 1995. *The History and Geography of Human Gene*. (Press PU, editor.).
- Cavalli-Sforza LL, Menozzi P, Piazza A. 1993. Demic expansions and human evolution. *Science* (80-.). 259:639–646.
- Clarke J, Brooks N, Banning EB, Bar-Matthews M, Campbell S, Clare L, Cremaschi M, di Lernia S, Drake N, Gallinaro M, et al. 2016. Climatic changes and social transformations in the Near East and North Africa during the “long” 4th millennium BC: A comparative study of

- environmental and archaeological evidence. *Quat. Sci. Rev.* 136.
- Cortés Sánchez M, Jiménez Espejo FJ, Simón Vallejo MD, Gibaja Bao JF, Carvalho AF, Martínez-Ruiz F, Gamiz MR, Flores JA, Paytan A, López Sáez JA, et al. 2012. The Mesolithic-Neolithic transition in southern Iberia. *Quat. Res.* 77:221–234.
- Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B, et al. 2004. Phylogeographic Analysis of Haplogroup E3b (E-M215) Y Chromosomes Reveals Multiple Migratory Events Within and Out Of Africa. *Am. J. Hum. Genet.* 74:1014–1022.
- Csilléry K, François O, Blum MGB. 2012. *ABC*: An R package for approximate Bayesian computation (ABC). *Methods Ecol. Evol.* 3:475–479.
- Currat M, Excoffier L. 2004. Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol.* 2.
- Currat M, Excoffier L. 2005. The effect of the Neolithic expansion on European molecular diversity. *Proc. R. Soc. B Biol. Sci.* 272:679–688.
- Currat M, Excoffier L. 2011. Strong reproductive isolation between humans and Neanderthals inferred from observed patterns of introgression. *Proc. Natl. Acad. Sci. U. S. A.* 108:15129–15134.
- Currat M, Poloni ES, Sanchez-Mazas A. 2010. Human genetic differentiation across the Strait of Gibraltar. *BMC Evol. Biol.* 10:237–255.
- Currat M, Ray N, Excoffier L. 2004. *SPLATCHE*: A program to simulate genetic diversity taking into account environmental heterogeneity. *Mol. Ecol. Notes* 4:139–142.
- Currat M, Ruedi M, Petit RJ, Excoffier L. 2008. The hidden side of invasions: Massive introgression by local genes. *Evolution (N. Y.)*. 62:1908–1920.
- Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. 2013. Haplotype estimation using sequencing reads. *Am. J. Hum. Genet.* 93:687–696.
- Delsler PM, Neumann R, Ballereau S, Hallast P, Batini C, Zadik D, Jobling MA. 2017. Signatures of human European Palaeolithic expansion shown by resequencing of non recombining X-chromosome segments. *Eur. J. Hum. Genet.* 25:485–492.
- Deshpande O, Batzoglou S, Feldman MW, Cavalli-Sforza LL. 2009. A serial founder effect model for human settlement out of Africa. *Proc. R. Soc. B Biol. Sci.* 276:291–300.
- Dunne J, Evershed RP, Salque MM, Cramp L, Bruni S, Ryan K, Biagetti S, di Lernia S. 2012. First dairying in green Saharan Africa in the fifth millennium BC. *Nature* 486:390–394.
- Ennafaa H, M Cabrera V, K Abu-Amero K, M González A, B Amor M, Bouhaha R, Dzimir N, Ben Ammar-Elgaaïed A, M Larruga J. 2009. Mitochondrial DNA haplogroup H structure in

- North Africa. *BMC Genet.* 10:8.
- Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564–567.
- Fagundes NJ, Ray N, Beaumont M, Neuenschwander S, Salzano FM, Bonatto SL, Excoffier L. 2007. Statistical evaluation of alternative models of human evolution. *Proc Natl Acad Sci U S A* 104:17614–17619.
- Ferembach D. 1985. On the origin of the iberomaurusians (Upper palaeolithic: North Africa). A new hypothesis. *J. Hum. Evol.* 14:393–397.
- Fernández E, Pérez-Pérez A, Gamba C, Prats E, Cuesta P, Anfruns J, Molist M, Arroyo-Pardo E, Turbón D. 2014. Ancient DNA Analysis of 8000 B.C. Near Eastern Farmers Supports an Early Neolithic Pioneer Maritime Colonization of Mainland Europe through Cyprus and the Aegean Islands. *PLoS Genet.* 10:e1004401.
- Fort J. 2012. Synthesis between demic and cultural diffusion in the Neolithic transition in Europe. *Proc. Natl. Acad. Sci. U. S. A.* 109:18669–18673.
- Fort J, Crema ER, Madella M. 2015. Modeling Demic and Cultural Diffusion: An Introduction. *Hum. Biol.* 87:141–149.
- Fort J, Méndez V. 1999. Time-Delayed Theory of the Neolithic Transition in Europe. *Phys. Rev. Lett.* 82:867–870.
- François O, Currat M, Ray N, Han E, Excoffier L, Novembre J. 2010. Principal component analysis under population genetic models of range expansion and admixture. *Mol Biol Evol* 27:1257–1268.
- Fu Q, Rudan P, Pääbo S, Krause J. 2012. Complete mitochondrial genomes reveal neolithic expansion into Europe. *PLoS One* 7.
- Gallego-Llorente M, Connell S, Jones ER, Merrett DC, Jeon Y, Eriksson A, Siska V, Gamba C, Meiklejohn C, Beyer R, et al. 2016. The genetics of an early Neolithic pastoralist from the Zagros, Iran. *Sci. Rep.* 6:31326.
- Goldberg A, Rosenberg NA. 2015. Beyond 2/3 and 1/3: The complex signatures of sex-biased admixture on the X chromosome. *Genetics* 201:263–279.
- González-Pérez E, Via M, Esteban E, López-Alomar A, Mazieres S, Harich N, Kandil M, Dugoujon J-M, Moral P. 2003. Alu insertions in the Iberian Peninsula and north west Africa--genetic boundaries or melting pot? *Coll Antropol* 27:491–500.
- Gonzalez AM, Brehm A, Pérez JA, Maca-Meyer N, Flores C, Cabrera VM. 2003. Mitochondrial DNA affinities at the atlantic fringe of Europe. *Am. J. Phys. Anthropol.* 120:391–404.
- Goring-Morris AN, Belfer-Cohen A. 2011. Neolithization Processes in the Levant: The Outer

- Envelope. *Curr. Anthropol.* 52:S195–S208.
- Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, Brandt G, Nordenfelt S, Harney E, Stewardson K, et al. 2015. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* 522:207–211.
- Hamilton MJ, Milne BT, Walker RS, Burger O, Brown JH. 2007. The complex structure of hunter-gatherer social networks. *Proc. Biol. Sci.* 274:2195–2202.
- Henn BM, Botigué LR, Gravel S, Wang W, Brisbin A, Byrnes JK, Fadhlaoui-Zid K, Zalloua PA, Moreno-Estrada A, Bertranpetit J, et al. 2012. Genomic ancestry of North Africans supports back-to-Africa migrations. *PLoS Genet.* 8:e1002397.
- Hervella M, Svensson EM, Alberdi A, Günther T, Izagirre N, Munters AR, Alonso S, Ioana M, Ridiche F, Soficaru A, et al. 2016. The mitogenome of a 35,000-year-old *Homo sapiens* from Europe supports a Palaeolithic back-migration to Africa. *Sci. Rep.* 6:25501.
- Hofmanová Z, Kreutzer S, Hellenthal G, Sell C, Diekmann Y, Díez-del-Molino D, van Dorp L, López S, Kousathanas A, Link V, et al. 2016. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl. Acad. Sci.* 113:6886–6891.
- Isern N, Fort J. 2010. Anisotropic dispersion, space competition and the slowdown of the Neolithic transition. *New J. Phys.* 12.
- Isern N, Fort J, Vander Linden M. 2012. Space Competition and Time Delays in Human Range Expansions. Application to the Neolithic Transition. *PLoS One* 7.
- Isern N, Zilhão J, Fort J, Ammerman AJ. 2017. Modeling the role of voyaging in the coastal spread of the Early Neolithic in the West Mediterranean. *Proc. Natl. Acad. Sci.* 114:897–902.
- Kimura M, Weiss GH. 1964. The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. *Genetics* 49:561–576.
- Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamarra B, Sirak K, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. *Nature* 536.
- Lazaridis I, Patterson N, Mittnik A, Renaud G, Mallick S, Sudmant PH, Schraiber JG, Castellano S, Kirsanow K, Economou C, et al. 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* 513:409–413.
- Linseele V, Van Neer W, Thys S, Phillipps R, Cappers R, Wendrich W, Holdaway S. 2014. New archaeozoological data from the fayum “neolithic” with a critical assessment of the evidence for early stock keeping in egypt. *PLoS One* 9.
- Linstadter J, Medved I, Solich M, Weniger GC. 2012. Neolithisation process within the Alboran

- territory: Models and possible African impact. *Quat. Int.* 274:219–232.
- Liu H, Prugnolle F, Manica A, Balloux F. 2006. A geographically explicit genetic model of worldwide human-settlement history. *Am. J. Hum. Genet.* 79:230–237.
- Lopes AM, Aston KI, Thompson E, Carvalho F, Gonçalves J, Huang N, Matthiesen R, Noordam MJ, Quintela I, Ramu A, et al. 2013. Human Spermatogenic Failure Purges Deleterious Mutation Load from the Autosomes and Both Sex Chromosomes, including the Gene DMRT1. *PLoS Genet.* 9.
- Madella M, García-Granero JJ, Out WA, Ryan P, Usai D. 2014. Microbotanical evidence of domestic cereals in Africa 7000 years ago. *PLoS One* 9.
- Manning K, Timpson A. 2014. The demographic response to holocene climate change in the Sahara. *Quat. Sci. Rev.* 101:28–35.
- McGee D, deMenocal PB, Winckler G, Stuut JWB, Bradtmiller LI. 2013. The magnitude, timing and abruptness of changes in North African dust deposition over the last 20,000 yr. *Earth Planet. Sci. Lett.* 371–372:163–176.
- McVean G a T, Myers SR, Hunt S, Deloukas P, Bentley DR, Donnelly P. 2004. The fine-scale structure of recombination rate variation in the human genome. *Science* 304:581–584.
- Meiklejohn C, Merrett DC, Reich D, Pinhasi R. 2017. Direct dating of human skeletal material from Ganj Dareh, Early Neolithic of the Iranian Zagros. *J. Archaeol. Sci. Reports* 12:165–172.
- Morales J, Pérez-Jordà G, Peña-Chocarro L, Zapata L, Ruíz-Alonso M, López-Sáez JA, Linstädter J. 2013. The origins of agriculture in North-West Africa: Macro-botanical remains from Epipalaeolithic and Early Neolithic levels of Ifri Oudadane (Morocco). *J. Archaeol. Sci.* 40:2659–2669.
- Mulazzani S, Belhouchet L, Salanova L, Aouadi N, Dridi Y, Eddargach W, Morales J, Tombret O, Zazzo A, Zoughlami J. 2016. The emergence of the Neolithic in North Africa: A new model for the Eastern Maghreb. *Quat. Int.* 410:123–143.
- Olalde I, Schroeder H, Sandoval-Velasco M, Vinner L, Lobon I, Ramirez O, Civit S, Borja PG, Salazar-Garcia DC, Talamo S, et al. 2015. A common genetic origin for early farmers from mediterranean cardial and central european LBK cultures. *Mol. Biol. Evol.* 32:3132–3142.
- Paschou P, Drineas P, Yannaki E, Razou A, Kanaki K, Tsetsos F, Padmanabhuni SS, Michalodimitrakakis M, Renda MC, Pavlovic S, et al. 2014. Maritime route of colonization of Europe. *Proc. Natl. Acad. Sci. U. S. A.* 111:9211–9216.
- Pinhasi R, Fort J, Ammerman AJ. 2005. Tracing the origin and spread of agriculture in Europe. *PLoS Biol.* 3:1–9.

- Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol Biol Evol* 16:1791–1798.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81:559–575.
- 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.
- Ray N, Currat M, Foll M, Excoffier L. 2010. SPLATCHE2: A spatially explicit simulation framework for complex demography, genetic admixture and recombination. *Bioinformatics* 26:2993–2994.
- Rhouda T, Martínez-Redondo D, Gómez-Durán A, Elmtili N, Idaomar M, Díez-Sánchez C, Montoya J, López-Pérez MJ, Ruiz-Pesini E. 2009. Moroccan mitochondrial genetic background suggests prehistoric human migrations across the Gibraltar Strait. *Mitochondrion* 9:402–407.
- Ruhlen M. 1991. *A Guide to the World's Languages*. (Press SU, editor.).
- Sampietro ML, Lao O, Caramelli D, Lari M, Pou R, Marti M, Bertranpetit J, Lalueza-Fox C. 2007. Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. *Proc Biol Sci* 274:2161–2167.
- Schaffner SF. 2004. The X chromosome in population genetics. *Nat. Rev. Genet.* 5:43–51.
- Skoglund P, Malmström H, Raghavan M, Storå J, Hall P, Willerslev E, Gilbert MTP, Götherström A, Jakobsson M, Menozzi P, et al. 2012. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* 336:466–469.
- Sokal RR, Oden NL, Wilson C. 1991. Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature* 351:143–145.
- The 1000 Genomes Project Consortium. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56–65.
- The International Hapmap Consortium. 2003. The International HapMap Project. *Nature* 426:789–796.
- Tillmar AO, Mostad P, Egeland T, Lindblom B, Holmlund G, Montelius K. 2008. Analysis of linkage and linkage disequilibrium for eight X-STR markers. *Forensic Sci. Int. Genet.* 3:37–41.
- Timmermann A, Friedrich T. 2016. Late Pleistocene climate drivers of early human migration.

Nature 538:92–95.

- Voight BF, Adams AM, Frisse L a, Qian Y, Hudson RR, Di Rienzo A. 2005. Interrogating multiple aspects of variation in a full resequencing data set to infer human population size changes. *Proc. Natl. Acad. Sci. U. S. A.* 102:18508–18513.
- Wegmann D, Excoffier L. 2010. Bayesian inference of the demographic history of chimpanzees. *Mol. Biol. Evol.* 27:1425–1435.
- Wegmann D, Leuenberger C, Excoffier L. 2009. Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics* 182:1207–1218.
- Yang Z. 2006. *Computational Molecular Evolution*. (Press OU, editor.).
- Zeder MA. 2008. Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact. *Proc. Natl. Acad. Sci. U. S. A.* 105:11597–11604.
- Zeder MA, Hesse B. 2000. The initial domestication of goats (*Capra hircus* L.) in the Zagros mountains 10,000 years ago. *Science* (80-.). 287:2254–2257.
- Zilhão J. 2001. Radiocarbon evidence for maritime pioneer colonization at the origins of farming in west Mediterranean Europe. *Proc. Natl. Acad. Sci. U. S. A.* 98:14180–14185.

Figure captions

Figure 1. Location and sample size (580 male individuals) of the 20 populations that composed the studied dataset. Populations were divided in 8 geographic groups abbreviated as EUR (European populations), BAS (Basque Country population), IBE (Iberian populations), WNA (western North Africa populations), CNA (central North Africa populations), ENA (eastern North Africa populations), SYR (Syrian population) and YRI (Yoruba population).

Figure 2. Landscape features for the spatially explicit computer simulations of the Neolithic expansion. (A) 2D map of the three different regions considered for the Neolithic expansion: Europe, North Africa and peripheral regions. Uninhabited regions (i.e., deserts) are shown in light brown. The origin of the simulated Neolithic expansion is shown in blue. (B) Alternative evolutionary scenarios of migration through the Strait of Gibraltar: (a) reciprocal migration between both regions (*fullMIG*), (b) lack of migration between both regions (*noMIG*), (c) migration allowed only from the Iberia Peninsula to North Africa (*nsMIG*), (d) migration allowed only from North Africa to the Iberia Peninsula (*snMIG*).

Figure 3. Fitting of the four alternative evolutionary scenarios of migration through the Gibraltar Strait (*fullMIG*, *noMIG*, *nsMIG* and *snMIG*) to the real data. Posterior probabilities of evolutionary scenarios were estimated with four different ABC methods (the *rejection* method by Pritchard and the *rejection*, *multinomial logistic regression* and *neural network* methods implemented in the *abc* package (Csilléry et al. 2012)).

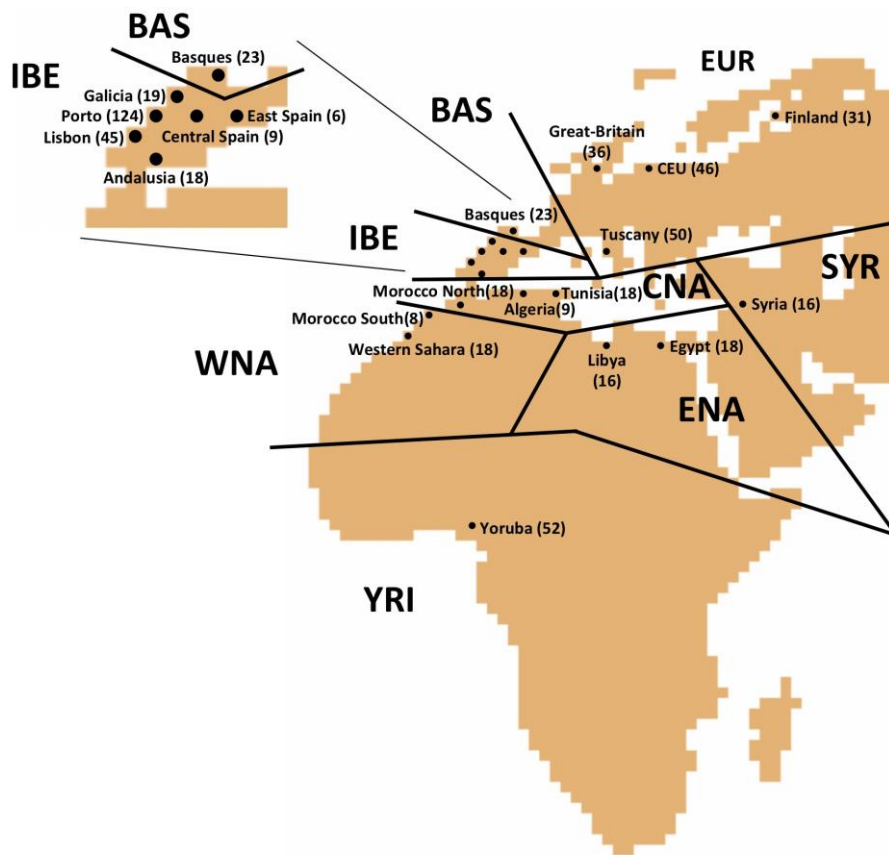


Figure 1.

A



B

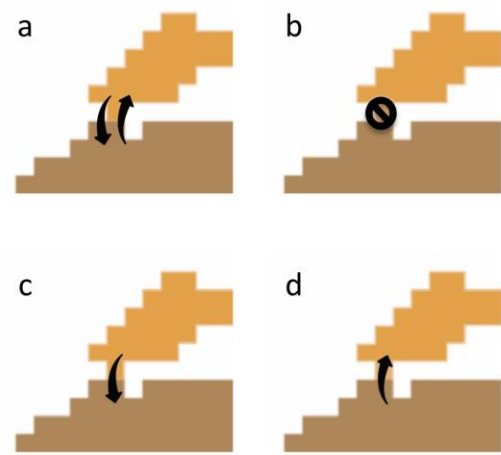


Figure 2.

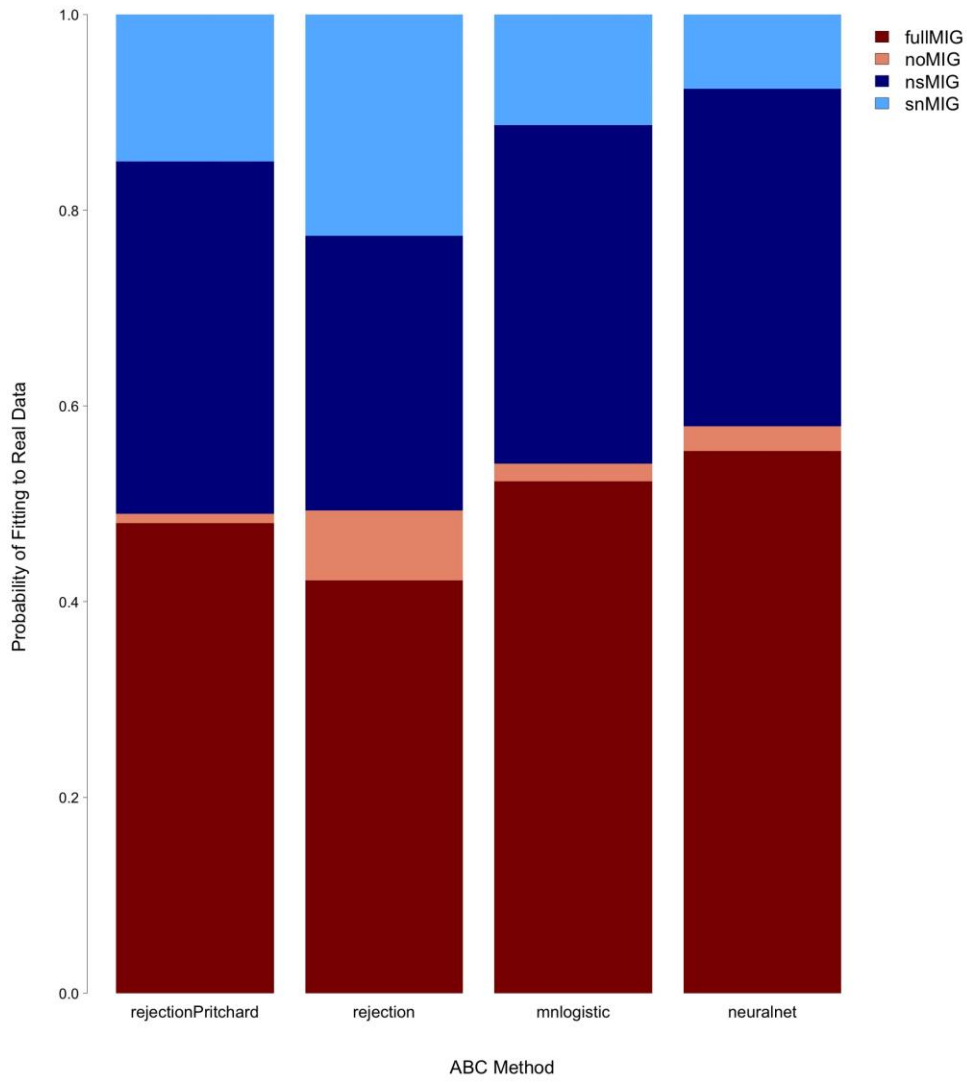


Figure 3.

Table 1. Population genetics parameters estimated with ABC. For each parameter, we present the median, mean and mode of the estimated posterior distribution. Parameters in bold correspond to comparisons between European and North African Neolithic populations. In italic we indicate the most accurate measure (median, mean or mode) according to Table S10. 95% HPDI indicates the 95% highest posterior probability.

Process	Population genetics parameter	Median	Mean	Mode	95% HPDI
Time	Initial Expansion of modern humans (T_{START}) ^a	100,313	<i>100,182</i>	88,041	81,048–119,050
	Out of Africa (T_{OOA}) ^a	66,313	65,864	<i>75,516</i>	50,899–79,300
	Initial Neolithic Expansion (T_{NEO}) ^a	10,425	<i>10,279</i>	11,321	8,174–11,900
Population size and demographics	Ancestral size Palaeolithic (N_{ANC})	13,966	<i>13,428</i>	16,687	5,726–19,625
	Ancestral size Neolithic (N_{ANCNEO})	26,917	27,354	<i>21,314</i>	6,272–49,117
	Population growth rate Palaeolithic (G_{PALEO})	0.671	<i>0.662</i>	0.730	0.414–0.890
	Population growth rate Neolithic (G_{NEO})	0.671	0.665	<i>0.836</i>	0.416–0.888
Migration	Migration rate Palaeolithic (M_{PALEO})	0.205	0.191	<i>0.233</i>	0.075–0.248
	Migration rate Neolithic in Europe (M_{NEOEUR})	0.210	0.199	0.232	0.092–0.248
	Migration rate Neolithic in North Africa (M_{NEONA})	0.192	0.183	0.231	0.074–0.248
	Migration rate Neolithic in Rest of the World (M_{NEOPER})	0.167	0.163	<i>0.223</i>	0.056–0.247
	Migration rate at Strait of Gibraltar - North to South route (M_{nsSG})	0.149	0.151	<i>0.138</i>	0.056–0.245
	Migration rate at Strait of Gibraltar - South to North route (M_{snSG})	0.153	0.152	<i>0.122</i>	0.055–0.245
Carrying capacity	Carrying capacity Palaeolithic (K_{PALEO})	1,477	1,414	<i>1,845</i>	571–1,970
	Carrying capacity Neolithic in Europe (K_{NEOEUR})	5,173	4,996	6,315	2,285–6,928
	Carrying capacity Neolithic in North Africa (K_{NEONA})	5,034	4,841	6,155	2,280–6,920
	Carrying capacity Neolithic in Rest of the World (K_{NEOPER})	4,681	4,655	6,384	2,224–6,863
Admixture	Interbreeding rate (I_{rate})	0.015	0.017	<i>0.011</i>	0.004–0.042

^aTime is shown in years, assuming a 25 year generation time.