The mixed genetic origin of the first farmers of Europe

Nina Marchi[1,2#], Laura Winkelbach[3#], Ilektra Schulz[2,4#], Maxime Brami[3#], Zuzana Hofmanová[2,4], Jens Blöcher[3], Carlos S. Reyna-Blanco[2,4], Yoan Diekmann[3], Alexandre Thiéry[1,2], Adamandia Kapopoulou[1,2], Vivian Link[2,4], Valérie Piuz[1], Susanne Kreutzer[3], Sylwia M. Figarska[3], Elissavet Ganiatsou[5], Albert Pukaj[3], Necmi Karul[6], Fokke Gerritsen[7], Joachim Pechtl[8], Joris Peters[9,10], Andrea Zeeb-Lanz[11], Eva Lenneis[12], Maria Teschler-Nicola[13,14], Sevasti Triantaphyllou[15], Sofija Stefanović[16], Christina Papageorgopoulou[5], Daniel Wegmann[2,4^*], Joachim Burger[3^*], Laurent Excoffier[1,2^*]

[#] These authors contributed equally. [^] These authors contributed equally.

* Corresponding authors (laurent.excoffier@iee.unibe.ch; jburger@uni-mainz.de; daniel.wegmann@unifr.ch)

Affiliations:

[1]CMPG, Institute for Ecology and Evolution, University of Berne, Berne, Switzerland.
[2]Swiss Institute of Bioinformatics, Lausanne, Switzerland.
[3]Palaeogenetics Group, Institute of Organismic and Molecular Evolution (iomE), Johannes Gutenberg University Mainz, Mainz, Germany.
[4]Department of Biology, University of Fribourg, Fribourg, Switzerland.
[5]Laboratory of Physical Anthropology, Department of History and Ethnology, Democritus University of Thrace, Komotini, Greece.
[6]Department of Prehistory, İstanbul University, Turkey.
[7]Netherlands Institute in Turkey, Istanbul, Turkey.
[8]Institute of Archaeology, Innsbruck University, Austria.
[9]Institute of Palaeoanatomy, Domestication Research and the History of Veterinary Medicine, LMU Munich, Munich, Germany.
[10]SNSB, State Collection of Anthropology and Palaeoanatomy Munich, Germany.
[11]Generaldirektion Kulturelles Erbe Rheinland-Pfalz, Speyer, Germany.
[12]Department of Prehistoric and Historical Archaeology, University of Vienna, Austria.
[13]Department of Anthropology, Natural History Museum of Vienna, Austria.
[14]Department of Evolutionary Anthropology, University of Vienna, Austria.
[15]Department of History and Archaeology, Aristotle University of Thessaloniki, Thessaloniki, Greece.
[16]Biosense Institute, University of Novi Sad/Department of Archaeology, Laboratory for Bioarchaeology, University of Belgrade, Serbia.

Alexandre Thiéry's present address: Cardio-CARE AG, Davos-Wolfgang, Switzerland.
Susanne Kreutzer's present address: Functional Genomics Center Zurich/GEML, ETH Zurich, Zurich, Switzerland & Department of Biology, ETH, Zurich, Switzerland.
Sylwia M. Figarska's present address: Department of Cardiology, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands.

Summary

While the Neolithic expansion in Europe is well described archaeologically, the genetic origins of European first farmers and their affinities with local hunter-gatherers (HGs) remain unclear. To infer the demographic history of these populations, the genomes of 15 ancient individuals located between Western Anatolia and Southern Germany were sequenced to high quality, allowing us to perform population genomics analyses formerly restricted to modern genomes. We find that all European and Anatolian early farmers descend from the merging of a European and a Near Eastern group of HGs, possibly in the Near East, shortly after the Last Glacial Maximum (LGM). Western and Southeastern European HG are shown to split during the LGM, and share signals of a very strong LGM bottleneck that drastically reduced their genetic diversity. Early Neolithic Central Anatolians seem only indirectly related to ancestors of European farmers, who probably originated in the Near East and dispersed later on from the Aegean along the Danubian corridor following a stepwise demic process with only limited (2-6%) but additive input from local HGs.

Our analyses provide a time frame and resolve the genetic origins of early European farmers. They highlight the impact of Late Pleistocene climatic fluctuations that caused the fragmentation, merging and reexpansion of human populations in SW Asia and Europe, and eventually led to the world's first agricultural populations.

Introduction

The origins and spread of agriculture in Southwest (SW) Asia, often described as the 'Neolithic transition', have been under research for well over a century[1]. While early sedentary communities emerged at the end of the Pleistocene[2], crop cultivation and ungulate management developed in the Fertile Crescent after the Younger Dryas cold spell (12.9-11.7 kya) with the onset of warmer conditions at the beginning of the Holocene[3–5]. Starting 10.6 kya, shifts towards small-scale agriculture with imported cultivars and livestock management are observed among sedentarizing communities of Central Anatolia[6,7] (**Fig. 1**). Around 8.7-8.6 kya, farming traditions expanded into the wider Aegean region, including the western half of the Anatolian Peninsula[8,9]. The new subsistence economy reached Crete and the Greek mainland shortly thereafter[10,11]. From the Aegean, farming spread into the Central Mediterranean Basin and the Danubian corridor, reaching the Central Balkans by about 8.2 kya, and Austria and Southern Germany by 7.5 kya[12].

Even though agriculture was first invented in SW Asia, the genetic origins of Europe's first farmers remain elusive. Recent palaeogenetic findings revealed that most European farmers are genetically closer to Central and Northwestern (NW) Anatolian farmers than to Pre-Pottery Neolithic (PPN) farmers of the Southern Levant or the Zagros region of Western Iran, who were genetically well differentiated[13–18]. However, these findings mostly rely on a set of ascertained genomic sites[19] that cannot easily be used for demographic reconstruction, and the temporal framework they provide depends on the dating of tested samples[16].

In order to characterize the demographic history and origin of European and Anatolian farmers, we generated high quality palaeogenomes from two Mesolithic hunter-gatherers (HGs) and 13 Early Neolithic farmers (mean depth between 10.55X and 15.21X, **Table 1**). These individuals were chosen on a regular spatial and temporal gradient along the main expansion axis of the Neolithic from the Near East into Central Europe (**Fig. 1**). We combined these data with nine ancient genomes of similar high quality available for this region and period[14,17,18,20–24] (**Fig. 1**, **S11a, Table S4**). We used these complete ancient genomes to perform model-based demographic inference based on the site frequency spectrum (SFS) at neutral sites[25]. We thus obtained a precise scenario of the colonization of Europe by early farmers and their interactions with local HGs, and estimated population size changes, interactions and split times with high accuracy.

Results

Patterns of genomic diversity along the Danubian corridor

The genetic structure and affinities of ancient individuals

Both multidimensional scaling (MDS) performed on the neutral portion of ancient and modern genomes (**Fig. 2a**) and an admixture analysis (**Fig. S21**) revealed three main clusters of ancient samples, which are found overall much more differentiated than modern individuals of SW Asia and Europe: i) a cluster of European HGs, ii) a cluster of Early Neolithic individuals from Iran, here represented by a single genome from Wezmeh Cave, WC1, and iii) a cluster with all other Holocene individuals. In keeping with previous analyses based on a restricted set of SNPs[26,27], this MDS analysis (**Fig. 2a**) suggests strongest affinities of European and NW Anatolian Neolithic samples with modern Sardinians, with the exception of the Early Neolithic NW Anatolian individual Bar8 found to be closer to modern individuals from Greece, Albania and other individuals from Southern Europe. However, a MDS analysis performed on the whole genome including sites potentially affected by selection (**Fig. S19**) rather suggests that early farmers are closer to Southern Europeans other than Sardinians. The Early Neolithic individual from Iran (WC1) shows strongest genetic affinities with modern Iranians (**Fig. 2a**), and to a lesser extent with individuals from the Northern Caucasus, suggesting some genetic continuity in Iran since Neolithic times. Finally, Palaeolithic and Mesolithic HGs are generally distinct from all modern SW Asians and Europeans, the closest of whom are Baltic Sea individuals, Russians and Scandinavians.

Early farmers are genetically more diverse and decline in stature over time

While genetic diversity as quantified by the heterozygosity at neutral sites was much reduced in HGs, most Early Neolithic farmers show diversity levels only slightly lower than those of modern humans (**Fig. 2b**), with genomes from NW Anatolia at the lower end of the distribution[17]. We note a slight reduction of diversity in modern humans with distance from Anatolia to the West of Boncuklu (Spearman's $\rho = -0.344$, p-value = 0.028), but not to the East (Spearman's $\rho = 0.019$, p-value = 0.929), while no such simple trend is observed among early European farmers.

Compared to the other samples, the HG genomes, and in particular Bichon and SF12, show a larger proportion of short (2-10Mb) Runs of Homozygosity (ROHs, **Fig. 2c**), in keeping with previous results[17,28,29]. This is indicative of higher levels of remote inbreeding within European

HGs, likely due to smaller population sizes as corroborated by *MSMC2* analyses (**Fig. S25**). Among early farmers, WC1, Bon002, AKT16, STAR1 and Stuttgart also show a high proportion of short ROHs and seem to be drawn from small populations, too. Furthermore, WC1, Stuttgart, LEPE52, Bichon and Loschbour, as well as several modern individuals from the Near East, carry some very long ROHs (>10Mb), indicative of recent inbreeding between close relatives (potentially second cousins or closer[30], **Fig. S20**).

We find that the vast majority of early farmers in our dataset had intermediate to light skin complexion, while HGs had a darker skin tone (**Supp. Table 3**). A dark (brown to black) hair color was inferred for all but two samples, LEPE52 and VC3-2, who likely had light brown hair. Eye color variation was similarly low, with all samples showing high probabilities for brown eyes, except for two individuals of the Starčevo culture (STAR1 and VC3-2) which were likely blue-eyed.

Based on polygenic scores, we show that early farmers are shorter than HGs (Student t-test, $t = -2.427$, p-value = 0.027), and their stature declined between 8,300 and 7,000 BP (Pearson's $r = 0.6537$, p-value < 0.008, **Fig. S24**), suggesting that selection for short stature occurred during the Neolithic expansion along the Danubian corridor.

The allele associated with lactase persistence was not found in any of the analyzed ancient samples, consistent with an increase in frequency of these alleles at a later stage[31]. However, early farmers already show allele frequencies similar to contemporary Europeans for 6 out of 7 SNPs of the *FADS1/2* gene complex, potentially selected in populations with plant-based diet[32,33] (see Suppl. Information - Section 5, **Table S7**).

Demographic inference

A step-wise expansion of Neolithic farmers into Central Europe

We contrasted eight scenarios of the spread of farmers into Europe, using four Early Neolithic populations from Northern Greece, Central Serbia, Lower Austria and Southern Germany, and one HG population from Serbia, each represented by at least two individuals (**Fig. S30**, **Table S9**). We find that strict stepwise scenarios are better supported than scenarios allowing for a long-distance migration from the Aegean directly to Lower Austria (**Fig. S30e**, **Supp. Table 4**). Importantly, scenarios without HG introgression into early farmer populations are clearly rejected. It implies that early farmer communities incorporated a few HG individuals (2-6%, **Fig. S31**) at all major stages of the dispersal along the Danubian corridor. However, the total amount of HG contribution into the farmer gene pool did not necessarily increase along the expansion as the input of HG genes was almost matched by the input from other farming

154 communities (2-5%). This complex pattern of gene flow might explain the apparent lack of
155 genetic structure among early farmer individuals shown in the MDS plot (**Fig. 2a**), as well as
156 an absence of increasing HG ancestry along the Danubian corridor in our admixture analysis
157 (**Fig. S21**).

158 **A mixed ancestry of all European and Anatolian farmers originating just after the LGM**
159 In our initial model, the population ancestral to all European farmers is surprisingly found to
160 be the product of a substantial post-LGM admixture between a HG population, potentially from
161 Anatolia or the Near East, and a HG population closely related to the genomes from Vlasac,
162 Central Serbia called hereafter *east* and *central* (**Fig. S31**). To shed more light on this
163 admixture, we progressively added individuals of other populations. We started with a
164 population from NW Anatolia represented by an individual from Barcın (Bar25), which we
165 found to have diverged from the other Aegean population (Northern Greece) very recently at
166 the beginning of the Holocene (~9.9 kya, 95% CI 10.8-9.1). We then added the Neolithic
167 genome from Aktopraklık in NW Anatolia, which we estimate to have split very recently from
168 Barcın (9.2 kya, 95% CI 9.5-9.1; **Fig. S36d**, **S37**, **Supp. Table 4**). However, this individual
169 received massive genetic contributions from both surrounding farmers (25%, 95% CI 28-18)
170 and surrounding HGs (16%, 95% CI 22-14) (**Fig. S37**, **Supp. Table 4**), in line with the
171 admixture analysis (**Fig. S21**), f-statistics (**Fig. S52**), and its affiliation to the 'coastal Fikirtepe
172 horizon' thought to have been influenced by both Epipalaeolithic and Neolithic traditions[8].
173 Importantly, these extended analyses confirm the old (~19.4 kya, 95% CI 23.3-10.4) and
174 massive (41% central HG contribution, 95% CI 38-50) admixture between the two HG
175 populations, which are found to have diverged during the LGM (23.4 kya; 95% CI 31.5-21.2)
176 (**Fig. S31**, **Supp. Table 4**).
177 To further study the spread of Neolithic people into Europe, we added two Early Neolithic
178 individuals from Lepenski Vir (**Fig. S40**), a site in the Danube Gorges with long pre-Neolithic
179 traditions of fishing, hunting and gathering and without ecological conditions for agriculture.
180 These two individuals previously shown to resemble Neolithic farmers from NW Anatolia[16,34]
181 are found to be tightly connected to the Northern Greek early farmers, and could thus be part
182 of an early migration of farmers into the Balkans.
183 Adding an early farmer from Boncuklu in the Konya plain of Central Anatolia (Bon002,
184 **Fig. S42a**) revealed that the Boncuklu population also shows a mixed ancestry, and that it
185 diverged ~13.4 kya (95% CI 14.6-11.5) from the branch leading to the Aegeans. In addition, it
186 would have received quite large (8%, 95% CI 1-17) and recent (~11.8 kya, 95% CI 13.0-10.3)

HG admixture, and relatively little input from surrounding farmers (2%, 95% CI 0-9) (**Fig. S43**, **Supp. Table 4**).

The genetically distinct (**Fig. 2a**) early farmer from Wezmeh Cave in the Iranian Zagros region is inferred to have diverged from the HG population ancestral to Aegeans and Central Anatolians (**Fig. S42b-c**) during the LGM (~20.1 kya, 95% CI 20.9-19.6, **Fig. S44**, **Supp. Table 4**), before it received the massive admixture from the *central* HGs observed in all other investigated early farmers. Its genetic proximity with the pre-admixed HG population (**Fig. S50, Fig. S51c**) suggests that the latter was located in the Near East.

Finally, we investigated the relationship between two western European HGs from Bichon and Loschbour and our newly-sequenced Mesolithic individuals from Serbia. We find that Bichon and Loschbour have a common ancestor branching off the *central* HG ancestral population 23.3 kya (95% CI 23.3-20.0) (**Fig. S47c, S48**, **Supp. Table 4**), and that they diverged from each other soon after this split. In contrast, the Danube Gorges Mesolithic population from Vlasac diverged from the *central* HG group relatively late about 10.2 kya (95% CI 9.0-21.3) and remained well isolated afterwards with very little later admixture (<1%) **(Fig. S31**, **Supp. Table 4**). Altogether, this suggests that the LGM led to a fragmentation of HG populations in SW Asia and Europe with at least four genetically distinct groups: one related to Loschbour and Bichon (called *west*, subdivided into *west1* and *west2*, based on the old divergence between Loschbour and Bichon branches), one related to the Danube Gorges Mesolithic samples (*central*), another one that later received the massive Central HG introgression (*east1* then *admixed*), and a last one potentially further East (*east2*) related to WC1 (**Fig. 3a, 4a, S34**, **S48**, **Supp. Table 4**).

Discussion

The LGM shaped Holocene genetic diversity in SW Asia and Europe

We find that Holocene human genetic structure in SW Asia and Europe emerged briefly before or during the LGM with the initial separation 32-21 kya of a western-central European and an eastern group of HGs. Right after this initial split, the western-central European branch experienced a very strong bottleneck (equivalent to a single human couple for one generation) that decreased the diversity of all descending populations. Then, these HGs further divided 23.3-20.0 kya, leaving us with three genetically distinct groups in western-central Europe that potentially differentiated in separate LGM refuge areas (**Fig. 4a**). The ancestors of Loschbour and Bichon could have resided in separate refugia in South Western Europe, and the ancestors of the Mesolithic Vlasac population could have lived in a geographically distinct *central* refugium likely located around the Balkans and the Aegean. Broadly speaking, these refugial populations coincide later on with what archaeologists have identified as the areas of distribution of Magdalenian and Epigravettian traditions in Europe[35,36]. In contrast, the *eastern* group of HGs, which does not show any sign of a strong bottleneck and was potentially genetically more diverse, diverged further into at least three groups of Near Eastern HGs during the LGM: one that later massively admixed with central HGs to become the ancestors of later Anatolian and Aegean farmers, one leading to the ancestors of Iranian Neolithic farmers, and one to Neolithic populations in the southern Levant (respectively *east1*, *east2* and *east3* on **Fig 4a**). After the LGM, these HG populations re-expanded from their southern refugia probably due to improving climatic conditions[37], allowing previously separated *central* and *east1* refugial populations to overlap and admix 19 kya (**Fig. 4b**), and then to become a distinct population from which Northwestern, Central Anatolian and European farmers would later descend.

Even though Central Anatolia has previously been proposed to have hosted admixture events[15,17], the exact geographic location of the massive post-LGM admixture event is difficult to pinpoint, and even though we modeled a single pulse of gene flow, admixture could also have consisted in extensive gene flow over several generations and over a relatively large area. We can nevertheless envision two alternative scenarios of admixture and later migrations. *1) A demic diffusion scenario:* the admixture took place mainly in the Fertile Crescent (**Fig. 4b**), implying separate migrations from the western Fertile Crescent to Central Anatolia and the broader Aegean region including NW Anatolia. Given the genetic proximity between Epipalaeolithic Central Anatolian foragers and Early Neolithic farmers[15] and our inferred early

split of the Boncuklu population, an initial migration into the Central Anatolian Plateau could have occurred already before the Younger Dryas and thus well before the Neolithic transition (**Fig. 4c**). In contrast, the migration to NW Anatolia would have occurred at the time of the fully developed, ceramic Neolithic (**Fig. 4c**), characterized by the establishment of widespread mixed farming across large parts of Anatolia[38]. Archaeological observations suggest two separate routes of neolithization towards the broader Aegean region. The first one would be a land-route across the Anatolian plateau, with Barcın showing clear cultural but only remote genetic affinities with Central Anatolia. The second one would be a maritime route connecting seafaring communities of the Eastern Mediterranean and the Aegean region[39,40] (**Fig. 4c**).
2) A cultural diffusion scenario: under this scenario, the admixture event at the origin of the ancestors of Anatolian and European Neolithic farmers occurred further west, i.e. closer to the inferred location of the Aegean refugial population. This scenario, which is plausible given technological interactions between HG communities in the Eastern Mediterranean and the Aegean[41], assumes a pre-Neolithic expansion of Near Eastern refugial populations into NW Anatolia. It would also explain the appearance of Near Eastern-like genetic signals in post-LGM European HGs, which has been postulated for the period 14-17 kya[16,42,43].

Despite some continuity in flake-based lithic industries across the Mesolithic-Neolithic transition in Greece, the abrupt appearance of fully developed Neolithic lifeways in that region involving dozens of innovations at hundreds of newly-founded sites[44,45] seems difficult to be explained by cultural diffusion alone and appears to be more compatible with demic diffusion from the Fertile Crescent. In contrast, based on our genetic data, adoption of agriculture by indigenous HG communities is more likely in Central Anatolia, where early aceramic sites like Boncuklu and Aşıklı show experiments in crop cultivation and caprine management, with increasing dependence on food-production, including a heavy emphasis on caprines after 9.7 kya[6,46,47].

Further support for the *demic diffusion* scenario comes from *f*-statistics showing Levantine populations to share more drift with Aegeans than with Central Anatolian Neolithic individuals (**Fig. S57**). This signal could either be due to some long distance gene flow between the Aegeans and the Levant, a higher level of *central* HG admixture observed in Boncuklu (**Fig. S56**), or a combination of i) an early migration of the Boncuklu HG ancestors from the Fertile Crescent to Central Anatolia before the Younger Dryas (**Fig. 3a, 4c**), ii) some gene flow between people from the Levant and the ancestors of Aegeans, who would have remained in the Fertile Crescent and only later migrated to the West. However, early farmers from the Aegean are rather heterogeneous in their levels of shared drift with several populations,

277 including Levantine HGs and early Iranian farmers (**Fig. S58**), suggesting that the peopling of
278 the Aegean was a complex process.

279 **A demic diffusion of Neolithic farmers along the Danubian corridor**

280 From an archaeological point of view, there have been a large number of proposed explanations
281 for the introduction of Neolithic lifeways in Europe[48]. Our explicit modelling supports the
282 simplest of all demic models, namely a gradual spread/progressive migration of early farmers
283 originating in the wider Aegean region (NW Anatolia or Greece) and extending to Serbia along
284 the Balkans and the so-called Danubian corridor, then to Hungary (usampled) and Austria, and
285 eventually up to the Rhine valley in southwestern Germany (**Fig. 1a, 4c**). While this study
286 focused on the Danubian or continental route of Neolithic expansion, we expect Impresso- and
287 Cardial-related farmers who spread along the Mediterranean shoreline to have shared a similar
288 genetic background in the Aegean[14].
289 Low levels of admixture with local Mesolithic populations (2-6%) seem to have occurred at
290 each of the four modelled migration steps, suggesting that early farming communities were not
291 completely genetically isolated[49]. The inferred rates of admixture are slightly lower than
292 previously reported (3-9%[14,49,50]). Even though we have modelled this hybridization process to
293 have occurred from the same Mesolithic metapopulation to which the Danube Gorges
294 Mesolithic individuals from Vlasac belong, we cannot exclude that admixture in Austria and
295 Southern Germany occurred with a Western European Mesolithic metapopulation, to which
296 Loschbour and Bichon are connected, as previous work has suggested that different Mesolithic
297 backgrounds could have introgressed early farmer gene pools in different regions[50].

298 **Advantages of demographic modeling**

299 Our sequencing of ancient genomes at >10X, which triples the number of whole genomes
300 available for the Early Holocene period in Europe, has allowed us to perform genetic analyses
301 on an unbiased set of markers minimally impacted by selection, thus ideally suited for
302 reconstructing the Neolithic settlement history of Western Eurasia. Our results fit into the larger
303 picture of an Anatolian origin of the first farmers and a settlement of Europe from the wider
304 Aegean region[14]. We also confirm the deep structure between early farming communities in
305 the eastern Fertile Crescent and NW Anatolia[15,18,51]. However, even though early farmers of
306 Central Anatolia are rather similar to those of NW Anatolia[17], we show that unlike the
307 populations in the Aegean, they are not part of the Neolithic migration chain to Europe.
308 Two unexpected findings were i) that the group formerly called WHG had already split into
309 two subgroups (*west* later substructured, and *central* HGs) approximately 23 kya, and ii) that

310 all Danubian early farmers can be traced back to a mixed population with substantial
311 contributions from HGs as they appear later in SE Europe. Even though early farmers were
312 recognized to be genetically intermediate between other Near Eastern groups[15], or considered
313 as a mixture of other ancient[21] or modern[17] populations, this initial admixture signal remained
314 hidden to previous approaches as it was eroded by later genetic drift (**Fig. S50-51** and see
315 below).

316 **Model validation**

317 These results are by no means definitive, and additional high-depth Mesolithic and
318 Epipalaeolithic samples from Greece, SE Anatolia and, importantly, from the Northern Levant
319 are needed to confirm our results. Nevertheless, we can show that the demographic model
320 reported in **Fig. 3a-b** can reproduce observed population affinities and patterns of admixture.
321 Indeed, genomic data simulated under our best demographic scenario leads to population
322 relationships very similar to those observed on an MDS (**Fig. 3c**), providing an *a posteriori*
323 validation of our model-based approach. In addition, we find that the ancestors of NW
324 Anatolian and European farmers simulated just after their defining admixture event (19 kya)
325 were exactly intermediate on a MDS between the two source populations (**Fig. S50b**), and that
326 their position on the MDS drifts away over time from this initial intermediate position towards
327 that of the early farmers. Similarly, the estimation of admixture on the same simulated data
328 shortly after the admixture event correctly recovers admixture proportions, but this signal
329 declines rapidly and disappears by the time ancient samples are sequenced (**Fig. S51**). Ten
330 thousand years of genetic drift have erased the initial admixture, thus explaining why the hybrid
331 nature of early farmers had been previously unnoticed. Finally, we used these simulations to
332 construct an admixture graph[19] matching the scenario depicted in **Fig. 3** (**Fig. S53**). The
333 estimated qpGraph is compatible with all *f*-statistics calculated on the real data (**Fig. S54-55**)
334 and it also recovers an ancient admixture event of similar proportions between *central* and
335 *eastern* HG groups.

336 In sum, our population modelling allowed us to extract novel, unforeseen, but complementary
337 and far more detailed information on population affinities and past history than what one can
338 conclude from summary statistics or multivariate analysis alone. In addition, we now have a
339 time frame for the differentiation of the major groups populating SW Asia and Europe from
340 the LGM until the introduction of agriculture, highlighting the crucial role of climatic changes
341 in promoting population fragmentation and secondary contacts[52].

342 # Tables

343

344 **Table 1 - Archaeological and genetic information on the newly-sequenced genomes.**

345

Individual	Period (culture)	Site	Country	Age (cal. BP)	Mean Depth (X)	Genetic sex	Haplogroups	
							mtDNA	Y
VLASA7	LM	Vlasac	Serbia	8764-8340	15.21	XY	U5a2a	I2
VLASA32	LM	Vlasac	Serbia	9741-9468	12.65	XY	U5a2a	R1b1
AKT16	EN	Aktopraklık	Turkey	8635-8460	12.25	XX	K1a3	–
Bar25	EN	Barcın	Turkey	8384-8205	12.65	XY	N1a1a1	G2a2b2a1
Nea3	EN	Nea Nikomedeia	Greece	8327-8040	11.57	XX	K1a2c	–
Nea2	EN	Nea Nikomedeia	Greece	8173-8023	12.51	XX	K1a	–
LEPE48	TEN	Lepenski Vir	Serbia	8012-7867	10.92	XY	K1a1	C1a2b
LEPE52	E-MN	Lepenski Vir	Serbia	7931-7693	12.37	XY	H3	G2a2b2a1a1c
STAR1	EN (Starčevo)	Grad-Starčevo	Serbia	7589-7476	10.55	XX	T2e2	–
VC3-2	EN (Starčevo)	Vinča-Belo Brdo	Serbia	7565-7426	11.22	XY	HV-16311	G2a2a1a3~
Asp6	EN (LBK)	Asparn-Schletz	Austria	7575-7474	12.11	XY	U5a1c1	G2a2b2a3
Klein7	EN (LBK)	Kleinhadersdorf	Austria	7244-7000	11.30	XX	W1-119	
Dil16	EN (LBK)	Dillingen-Steinheim	Germany	7235-6998	10.60	XY	J1c6	C1a2b
Ess7	EN (LBK)	Essenbach-Ammerbreite	Germany	7050-6900	12.34	XY	U5b2c1	G2a2b2a1a1
Herx	EN (LBK)	Herxheim	Germany	7164-6993	11.46	XX	K1a4a1i	–

LM, Late Mesolithic; EN, Early Neolithic; TEN, Transformational/Early Neolithic; E-MN, Early-Middle Neolithic; LBK, Linearbandkeramik

346

Figures

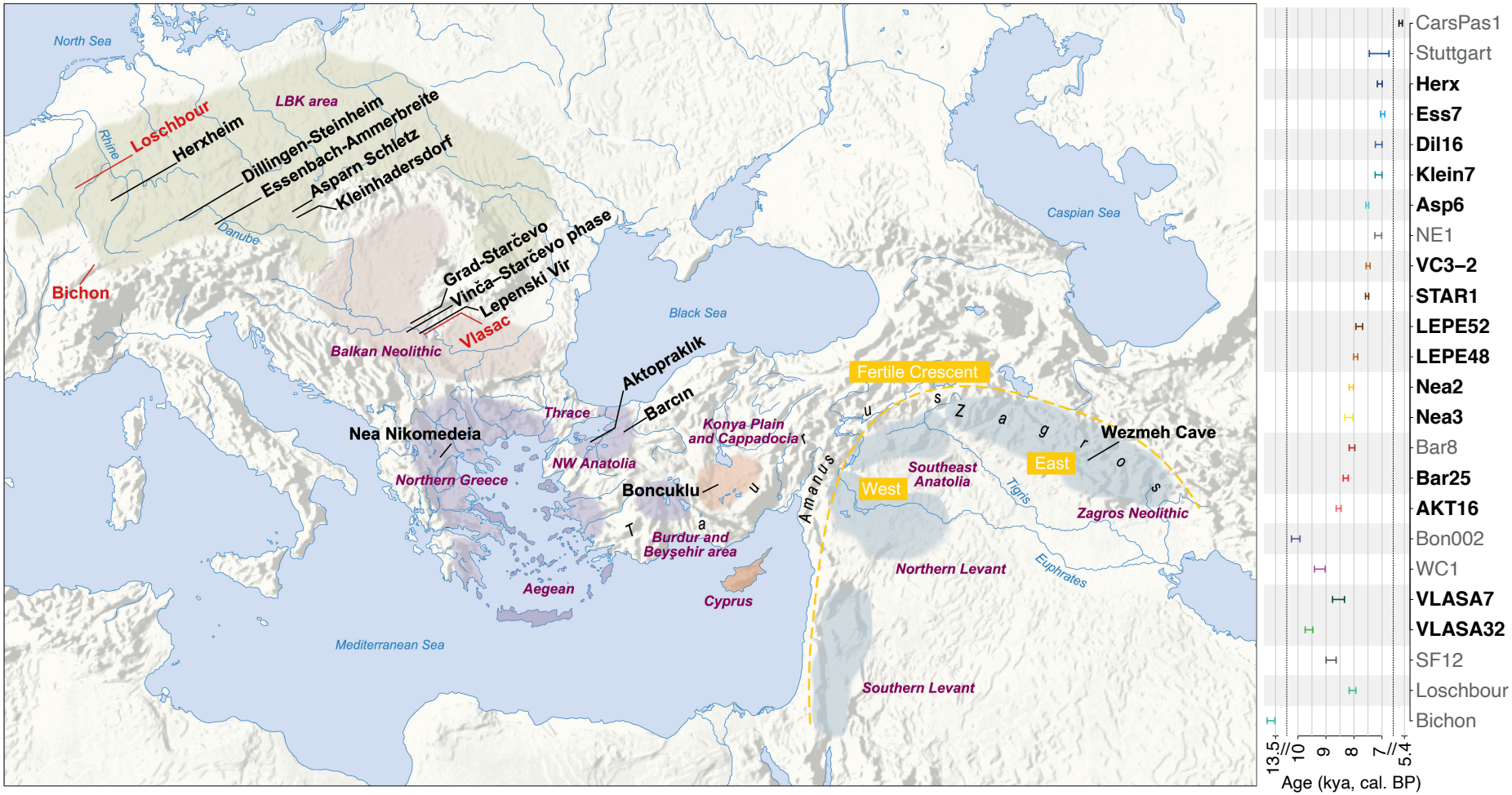

Figure 1 - Spatial and temporal distribution of the palaeogenomes analyzed in this study. Left: Location of the Neolithic (black), and Mesolithic or Palaeolithic (red) archaeological sites sampled for demographic modelling. Coloured areas reflect different chronological phases of agricultural expansion along the Eastern Mediterranean and Danubian routes of neolithization. **Right:** Chronological distribution of the 24 genomes analyzed in this study (see details in **Table 1** and **Table S4**), with the 15 newly-sequenced genomes shown in bold. The chronological interval at 2 sigma (95.4% probability) is shown for each directly-^{14}C dated sample, except for Stuttgart and Ess7, for which an approximate date is given based on the archaeological context.

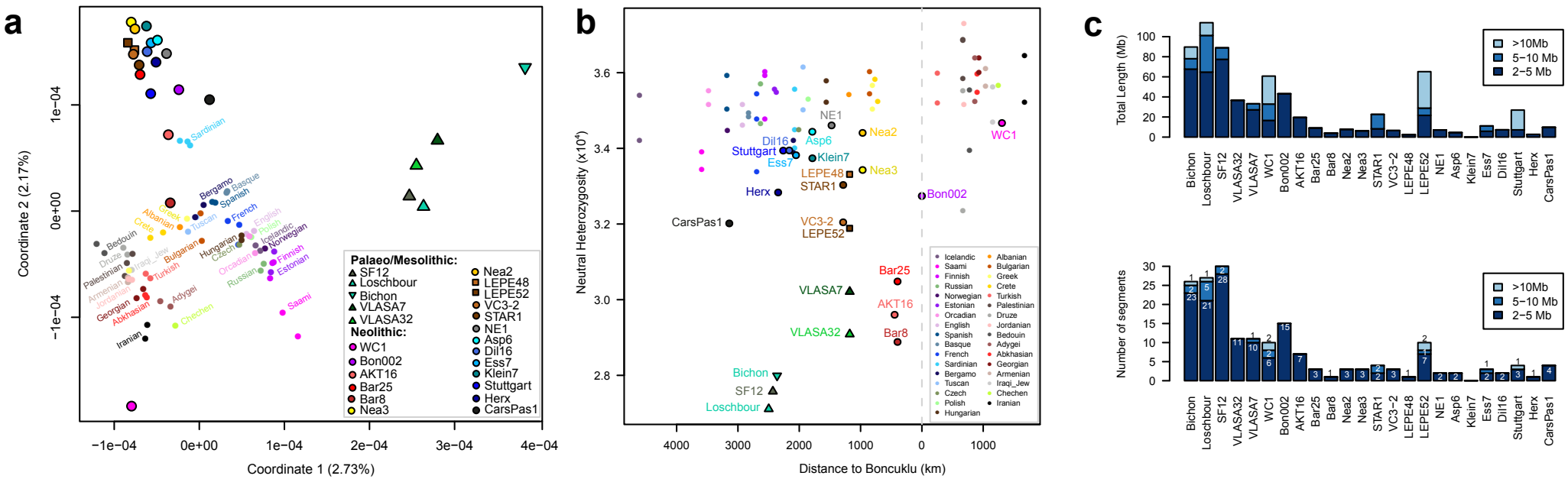

Figure 2 - Genetic relationships and diversity of high quality ancient genomes. a: Multidimensional Scaling Diversity (MDS) analysis performed on the neutrally evolving portion of ancient (n = 24) and modern (n = 65), shown as small circles) whole genomes from Europe and SW Asia (see Supplementary Information - Section 4). **b**: Heterozygosity computed at neutral sites in ancient and modern genomes plotted against air distance from Boncuklu in Central Anatolia. **c**: Runs of Homozygosity (ROHs) computed on imputed ancient whole genomes.

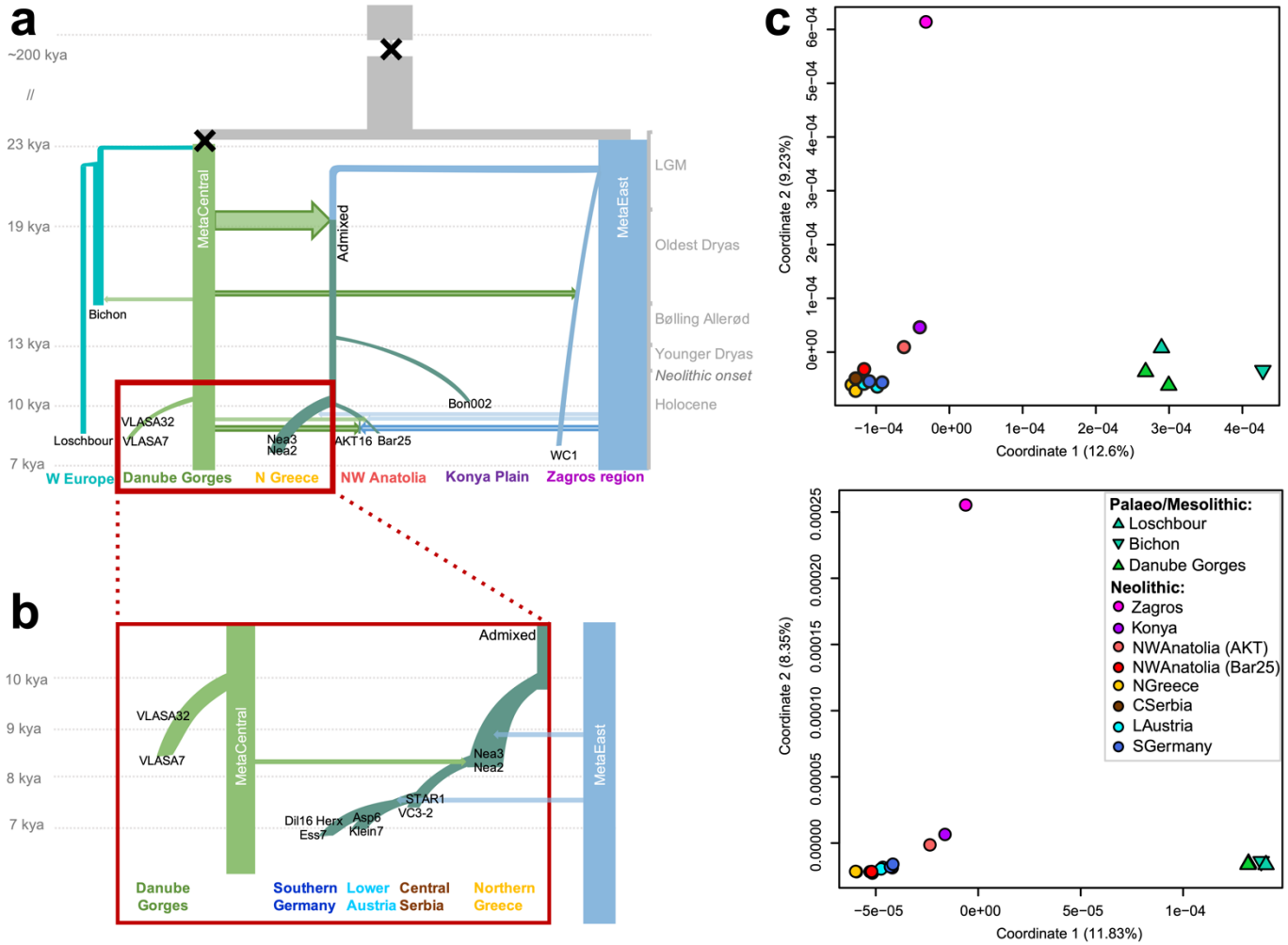

Figure 3 - Demographic scenario inferred from the sampled genomes and underlying genetic data. a: This demographic history was obtained by summarizing the best models of all tested scenarios. **b:** zoom-in on the red-square area in panel a. The X symbols indicate very strong bottlenecks that occurred on the HG ancestral branch before the divergence between Bichon-Loschbour and central European HGs and some 200 kya in the ancestral population. Only admixtures with point estimates ≥ 5% are represented with arrows (≥ 10% when arrows have a dark outline). **c:** MDS analyses performed on the neutrally evolving portion of the 17 ancient whole-genomes used in the demographic models (left) or on data simulated (right) according to the inferred ML parameters of the global scenario shown in panes a and b (See Supplementary Information - Section 7).

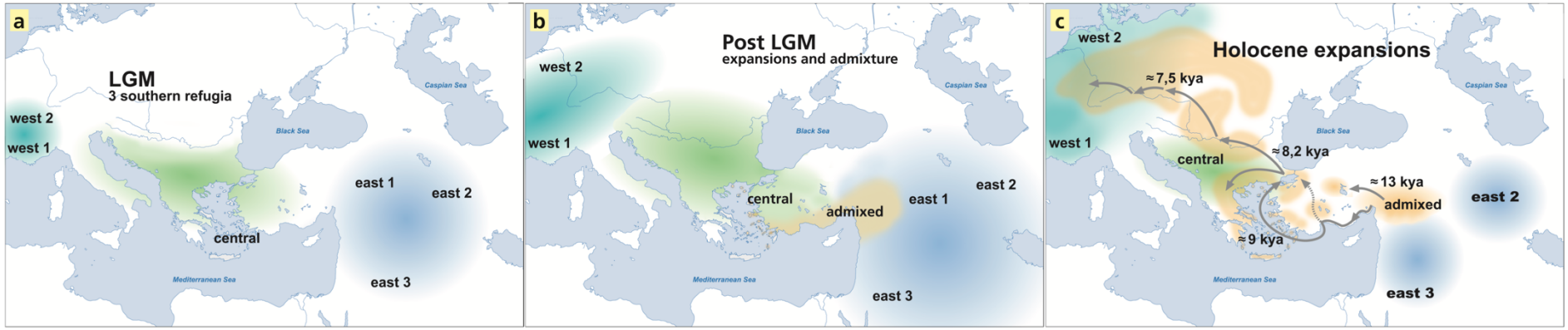

Figure 4 - Possible scenario of the population history of SW-Asia and Europe between the Last Glacial Maximum (LGM) and the Early Neolithic period, i.e. approximately 26,000 to 7,000 years ago. Note that the exact geographic distribution of the populations is very approximate. See main text for a detailed description.

Methods

Laboratory work

DNA extraction and sequencing: Palaeogenetic analyses were conducted in the dedicated ancient DNA facilities of the Palaeogenetics Group (Johannes Gutenberg University, Mainz), according to strict ancient DNA protocols. The ancient DNA samples were extracted from petrous bones (*Pars petrosa*) (Suppl. Information - Section 2). DNA extracts were treated with USER™ enzyme[53] and double-indexed libraries were prepared according to the protocol of Kircher *et al.*[54] with slight modifications. The libraries were sequenced on an Illumina HiSeq3000 (SE, 100 cycles or PE, 150 cycles) at the Next Generation Sequencing Platform at the University of Berne, Switzerland.

Bioinformatics pipeline

Genotype calling: We committed to process the 24 palaeogenomes as well as 77 modern SGDP genomes[55] with the same bioinformatic pipeline where possible. For the 15 newly sequenced genomes, we adapter-trimmed (TrimGalore! v.0.11.5), aligned with *bwa*, v.0.7.15[56] to the hs37d5 reference sequence[57], filtered for length ≥ 30 and mapping quality ≥ 30 (*SAMtools*, v.1.3, Li 2009) and marked PCR-duplicates (*Picard-tools*, v.2.9). Where available, we used the same pipeline for raw FASTQ-files of the 9 published palaeogenomes. In other cases, we transformed aligned BAM files to FASTQ files first. We also marked PCR-duplicates on the modern genomes. All samples underwent Local Realignment following *GATK* (v.3.7) guidelines but with a new approach for identifying indel sites[58] (Suppl. Information - Section 3). Reads containing soft-clipped positions were removed. In our snakemake-based *ATLAS-pipeline* (commit 6df90e7), we merged paired-end reads and called genotypes by taking potential post-mortem damages and base quality recalibration patterns of sequencing errors into account[59]. Additionally, we estimated genetic sex[60] and contamination[61,62]. All genotype calls were then filtered for read depth, genotype quality and allelic imbalance and polarised with the Chimpanzee reference genome.

Filtering data for demographic inference: In order to avoid biases due to background selection (BGS) and biased gene conversion (BGC) when estimating population diversity and relationships[63], we performed most of our genetic analyses on a "neutral" portion of the genome[64]. We thus extracted a restricted set of sites in regions with recombination rate >1.5 cM/Mb where BGS has little effect and with A↔T and G↔C mutations, which are not affected

409 by BGC or PMD. We also imputed and phased genotypes with *SHAPEIT4* v1.2[65] using default
410 parameters for sequence data, the HapMap phase II b37 genetic map, and the Haplotype
411 Reference Consortium[66] dataset as reference panel.

412 Population genetics

413 **Genetic relationships** among individuals were estimated from pairwise average nucleotide
414 divergence π_{XY}[67] computed on their neutral genomic portion and represented with a classical
415 multidimensional scaling (MDS) approach implemented in R (*cmdscale* function).
416 **Genomic heterozygosity** was computed as the amount of heterozygous sites found in the
417 neutrally evolving portion of each genome divided by the expected number of neutral sites
418 (Suppl. Information - Section 4).
419 **Admixture** clustering analyses were realized for two subsets of ancient individuals (Suppl.
420 Information - Sections 4&7), focusing on sites with no missing data, with R package *LEA*[68],
421 parameters K, alpha = 100, number of repetitions = 5, and run goodness-of-fit and admixture
422 coefficients being calculated in an unsupervised manner with function *snmf.*
423 **D-statistics** were computed on pseudo-haploid data from reference 1240K dataset v42.4 and
424 on majority calls generated with ATLAS at the ~1.2 mil. SNPs from the reference. D-statistics
425 was calculated in a form of *D(Individual1, Individual2; Individual3, Outgroup)* using
426 ADMIXTOOLS[19] and Mbuti as the outgroup. We used qpGraph to check the fit between f-
427 statistics.
428 **Runs of Homozygosity (ROHs)** were identified in genomes of modern and ancient Western
429 Eurasian individuals imputed using IBDSeq v. r1206[69] with default parameters but errormax =
430 0.005 and ibdlod = 2, and after artificially long tracts spanning assembly gaps or centromeres
431 were split into shorter tracks excluding the gap stretch.

432 Uniparental haplogroup determination

433 Mitochondrial haplogroups were determined from the BAM files for the 15 newly-sequenced
434 genomes using phy-mer[70] with K-mer minimal number of occurrences = 10; Y-chromosomal
435 haplogroups were determined using Yleaf[71] with minimal base-quality of 20 and base-majority
436 to determine an allele set to 90%.

437 # Phenotype predictions

438 Pigmentation phenotypes of hair, skin and eyes were predicted for each of the newly sequenced
439 samples with the HirisPlexS webtool[72,73] on genotypes or BAM files directly when genotypes
440 were missing (Suppl. Information - Section 5).
441 Genotypes for SNPs associated with additional phenotypes of interest were inspected manually
442 for each sample: rs4988235 variant in *MCM6* gene and associated with lactase-persistence in
443 Eurasia; rs3827760 in *EDAR* gene; rs17822931 in *ABCC11*; seven SNPs located in the
444 *FADS1/2* gene complex.
445 For predicted standing height, polygenic scores (PS) were computed based on a set of 670
446 SNPs[74].

447 # Demographic Analyses

448 ***MSMC2*** analyses[75] were performed on the phased-imputed dataset following the author's
449 recommendation, including masking for chromosome mappability on the hs37d5 human
450 reference genome and for sample-specific sequencing quality (Supp. Information - Section 6).
451 ***fastsimcoal2*** analyses[25] were carried out on seven different data sets of newly sequenced
452 individuals, on the neutral SFS and with neutral mutation rate defined for each dataset (see Supp.
453 Information - Section 6). We used 50 independent runs with 100 expectation conditional
454 maximization (ECM; 150 in one model) cycles per run and 500,000 coalescent simulations per
455 estimation of the expected SFS. Confidence intervals for the maximum-likelihood parameters
456 point estimates were computed with a parametric bootstrap approach.

457 # Reporting summary

458 Further information on research design is available in the Nature Research Reporting Summary
459 linked to this paper.

460 # Code availability

461 ATLAS pipeline commit 6df90e7 is available at:
462 https://bitbucket.org/wegmannlab/atlas-pipeline/src/master/

463 # Data availability

464 (Sequences will be made available at the European Nucleotide Archive.)

References

1. Childe, V. G. *The Dawn of European Civilization*. (Kegan Paul, Trench, Trübner & Co., 1925).
2. Bar-Yosef, O. The Natufian culture in the Levant, threshold to the origins of agriculture. *Evolutionary Anthropology: Issues, News, and Reviews* **6**, 159–177 (1998).
3. Tanno, K.-I. & Willcox, G. How fast was wild wheat domesticated? *Science* **311**, 1886 (2006).
4. Willcox, G. & Stordeur, D. Large-scale cereal processing before domestication during the tenth millennium cal BC in northern Syria. *Antiquity* **86**, 99–114 (2012).
5. Peters, J., von den Dreisch, A. & Helmer, D. The upper Euphrates-Tigris basin: Cradle of agro-pastoralism? in *The first steps of animal domestication: new archaeozoological approaches* (eds. Helmer, D., Peters, J. & Vigne, J.-D.) 96–123 (Oxbow, 2005).
6. Stiner, M. C., Buitenhuis, H. & Duru, G. A forager–herder trade-off, from broad-spectrum hunting to sheep management at Aşıklı Höyük, Turkey. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8404–8409 (2014).
7. Baird, D. *et al.* Agricultural origins on the Anatolian plateau. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E3077–E3086 (2018).
8. Özdoğan, M. Archaeological Evidence on the Westward Expansion of Farming Communities from Eastern Anatolia to the Aegean and the Balkans. *Current Anthropology* **52**, S415–S430 (2011).
9. Weninger, B. *et al.* Neolithisation of the Aegean and Southeast Europe during the 6600-6000 calBC period of Rapid Climate Change. *Documenta Praehistorica* **41**, 1–31 (2014).
10. Perlès, C., Quiles, A. & Valladas, H. Early seventh-millennium AMS dates from domestic seeds in the Initial Neolithic at Franchthi Cave (Argolid, Greece). *Antiquity* **87**, 1001–1015 (2013).
11. Douka, K., Efstratiou, N., Hald, M. M., Henriksen, P. S. & Karetsou, A. Dating Knossos and the arrival of the earliest Neolithic in the southern Aegean. *Antiquity* **91**, 304–321 (2017).
12. Shennan, S. *The First Farmers of Europe: An Evolutionary Perspective*. (Cambridge University Press, 2018).
13. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
14. Hofmanová, Z. *et al.* Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 6886–6891 (2016).
15. Feldman, M. *et al.* Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).
16. Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
17. Kılınç, G. M. *et al.* The Demographic Development of the First Farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).
18. Broushaki, F. *et al.* Early Neolithic genomes from the eastern Fertile Crescent. *Science* **353**, 499–503 (2016).
19. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
20. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).
21. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
22. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
23. Günther, T. *et al.* Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).
24. Brace, S. *et al.* Ancient genomes indicate population replacement in Early Neolithic Britain. *Nature Ecology & Evolution* **3**, 765–771 (2019).
25. Excoffier, L., Dupanloup, I., Huerta-Sanchez, E., Sousa, V. C. & Foll, M. Robust demographic inference from genomic and SNP data. *PLoS Genet.* **9**, e1003905 (2013).
26. Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in

Europe. *Science* **336**, 466–469 (2012).
27. Marcus, J. H. *et al.* Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nat. Commun.* **11**, 939 (2020).
28. Ringbauer, H., Novembre, J. & Steinruecken, M. Detecting runs of homozygosity from low-coverage ancient DNA. *bioRxiv* (2020).
29. Ceballos, F. C. *et al.* Human inbreeding has decreased in time through the Holocene. *bioRxiv* (2020) doi:10.1101/2020.09.24.311597.
30. Gazal, S. *et al.* Inbreeding coefficient estimation with dense SNP data: comparison of strategies and application to HapMap III. *Hum. Hered.* **77**, 49–62 (2014).
31. Burger, J. *et al.* Low Prevalence of Lactase Persistence in Bronze Age Europe Indicates Ongoing Strong Selection over the Last 3,000 Years. *Curr. Biol.* (2020) doi:10.1016/j.cub.2020.08.033.
32. Fumagalli, M. *et al.* Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* **349**, 1343–1347 (2015).
33. Buckley, M. T. *et al.* Selection in Europeans on Fatty Acid Desaturases Associated with Dietary Changes. *Mol. Biol. Evol.* **34**, 1307–1318 (2017).
34. Hofmanová, Z. Palaeogenomic and biostatistical analysis of ancient DNA data from Mesolithic and Neolithic skeletal remains. (University of Mainz, 2016).
35. Kozłowski, J. K. & Kaczanowska, M. Gravettian/Epigravettian sequences in the Balkans and Anatolia. *Mediterranean Archaeology and Archaeometry* **4**, 5–18 (2004).
36. *The Central European Magdalenian: Regional Diversity and Internal Variability*. (Springer, 2015).
37. Bar-Yosef, O. Climatic Fluctuations and Early Farming in West and East Asia. *Curr. Anthropol.* **52**, S175–S193 (2011).
38. Bogaard, A. *et al.* Agricultural innovation and resilience in a long-lived early farming community: the 1,500-year sequence at Neolithic to early Chalcolithic Çatalhöyük, central Anatolia. *Anatolian Studies* vol. 67 1–28 (2017).
39. Horejs, B. *et al.* The Aegean in the Early 7th Millennium BC: Maritime Networks and Colonization. *J World Prehist* **28**, 289–330 (2015).
40. Özbal, R. & Gerritsen, F. Barcın Höyük in Interregional Perspective: An Initial Assessment. in *Concluding the Neolithic. The Near East in the Second Half of the Seventh Millennium BC* (ed. Marciniak, A.) 287–305 (Lockwood Press, 2019).
41. Efstratiou, N., Biagi, P. & Starnini, E. The Epipalaeolithic site of Ouriakos on the island of Lemnos and its place in the Late Pleistocene peopling of the east Mediterranean region. *Adalya* **17**, 1–13 (2014).
42. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
43. Bortolini, E. *et al.* Early Alpine occupation backdates westward human migration in Late Glacial Europe. *Genomics* e0217996 (2020).
44. Perlès, C. *The Early Neolithic in Greece: The First Farming Communities in Europe*. (Cambridge University Press, 2001).
45. Perlès, C. An alternate (and old-fashioned) view of Neolithisation in Greece. *Documenta Praehistorica* **30**, 99–113 (2003).
46. Buitenhuis, H. *et al.* The faunal remains from levels 3 and 2 of Aşıklı Höyük: Evidence for emerging management practices. in *The Early Settlement at Aşıklı Höyük. Essays in Honor of Ufuk Esin* (eds. Özbaşaran, M., Duru, G. & Stiner, M.) 281–324 (Yayınları, 2018).
47. Ergun, M., Tengberg, M., Willcox, G. & Douché, C. Plants of Aşıklı Höyük and changes through time: First archaeobotanical results from the 2010--14 excavation seasons. in *The Early Settlement at Aşıklı Höyük. Essays in Honor of Ufuk Esin* (eds. Özbaşaran, M., Duru, G. & Stiner, M.) 191–217 (Yayınları, 2018).
48. *Europe's First Farmers*. (Cambridge University Press, 2000).
49. Nikitin, A. G. *et al.* Interactions between earliest Linearbandkeramik farmers and central European hunter gatherers at the dawn of European Neolithization. *Sci. Rep.* **9**, 19544 (2019).
50. Lipson, M. *et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
51. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East.

572 *Nature* **536**, 419–424 (2016).
573 52. Lahr, M. M. & Foley, R. A. Towards a theory of modern human origins: geography,
574 demography, and diversity in recent human evolution. *Am. J. Phys. Anthropol.* **Suppl**, 137–176
575 (1998).
576 53. Verdugo, M. P. *et al.* Ancient cattle genomics, origins, and rapid turnover in the Fertile
577 Crescent. *Science* **365**, 173–176 (2019).
578 54. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex
579 sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2012).
580 55. Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142 diverse
581 populations. *Nature* **538**, 201–206 (2016).
582 56. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
583 *arXiv [q-bio.GN]* (2013).
584 57. The 1000 Genomes Project Consortium. A global reference for human genetic variation.
585 *Nature* vol. 526 68–74 (2015).
586 58. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-
587 generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
588 59. Link, V. *et al.* ATLAS: Analysis Tools for Low-depth and Ancient Samples.
589 doi:10.1101/105346.
590 60. Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of
591 ancient human remains using DNA shotgun sequencing. *Journal of Archaeological Science* vol. 40
592 4477–4482 (2013).
593 61. Fu, Q. *et al.* A revised timescale for human evolution based on ancient mitochondrial
594 genomes. *Curr. Biol.* **23**, 553–559 (2013).
595 62. Rasmussen, M. *et al.* An Aboriginal Australian genome reveals separate human dispersals
596 into Asia. *Science* **334**, 94–98 (2011).
597 63. Matthey-Doret, R. & Whitlock, M. C. Background selection and FST : Consequences for
598 detecting local adaptation. *Mol. Ecol.* **28**, 3902–3914 (2019).
599 64. Pouyet, F., Aeschbacher, S., Thiéry, A. & Excoffier, L. Background selection and biased gene
600 conversion affect more than 95% of the human genome and bias demographic inferences. *Elife* **7**,
601 (2018).
602 65. Delaneau, O., Zagury, J.-F., Robinson, M. R., Marchini, J. L. & Dermitzakis, E. T. Accurate,
603 scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019).
604 66. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat.*
605 *Genet.* **48**, 1279–1283 (2016).
606 67. Nei, M. & Li., W. H. Mathematical model for studying genetic variation in terms of
607 restriction endonucleases. *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5269–5273 (1979).
608 68. Frichot, E. & François, O. LEA: An R package for landscape and ecological association
609 studies. *Methods in Ecology and Evolution* vol. 6 925–929 (2015).
610 69. Browning, B. L. & Browning, S. R. Detecting identity by descent and estimating genotype
611 error rates in sequence data. *Am. J. Hum. Genet.* **93**, 840–851 (2013).
612 70. Navarro-Gomez, D. *et al.* Phy-Mer: a novel alignment-free and reference-independent
613 mitochondrial haplogroup classifier. *Bioinformatics* **31**, 1310–1312 (2015).
614 71. Ralf, A., González, D. M., Zhong, K. & Kayser, M. Yleaf: Software for Human Y-
615 Chromosomal Haplogroup Inference from Next-Generation Sequencing Data. *Molecular Biology and*
616 *Evolution* vol. 35 1291–1294 (2018).
617 72. Walsh, S. *et al.* The HIrisPlex system for simultaneous prediction of hair and eye colour from
618 DNA. *Forensic Sci. Int. Genet.* **7**, 98–115 (2013).
619 73. Chaitanya, L. *et al.* The HIrisPlex-S system for eye, hair and skin colour prediction from
620 DNA: Introduction and forensic developmental validation. *Forensic Sci. Int. Genet.* **35**, 123–135
621 (2018).
622 74. Chan, Y. *et al.* Genome-wide Analysis of Body Proportion Classifies Height-Associated
623 Variants by Mechanism of Action and Implicates Genes Important for Skeletal Development. *Am. J.*
624 *Hum. Genet.* **96**, 695–708 (2015).
625 75. Schiffels, S. & Wang, K. MSMC and MSMC2: The Multiple Sequentially Markovian
626 Coalescent. *Methods Mol. Biol.* **2090**, 147–166 (2020).

Acknowledgements

We are grateful to Martina Unterländer and Aleksandra Žegarac for help with sample preparation. Lara Cassidy and Kay Prüfer kindly provided access to unpublished raw sequencing data. We thank Ourania Palli and Franz Pieler for useful archaeological information.

We also acknowledge the use of the sequencing platform at the University of Berne for whole genome sequencing services and support, the IBU cluster of the University of Berne for NGS data analysis (https://www.bioinformatics.unibe.ch/), as well as the UBELIX HPC cluster of the University of Berne (http://www.id.unibe.ch/hpc) for demographic model analyses, and the super-computer Mogon of Johannes Gutenberg University Mainz (https://hpc.uni-mainz.de).

NM was supported by a Swiss SNSF grant No. 310030_188883 to LE and by a Seal of Excellence Fund grant from the University of Berne (SELF2018-04). IS and VL were supported by a Swiss SNSF grant No. 31003A_173062 to DW. MB was supported by a Marie Skłodowska-Curie individual fellowship (GA: 793893, 'ODYSSEA'). ZH was supported by EMBO Long-Term Fellowship (EMBO ALTF 445-2017). JBu, CP and SK were funded by the German Science Foundation (BU 1403/6-1). JBu and CP were funded by the Humboldt foundation. JBu, SS, SF, and ZH received funding from Marie-Curie-Actions ITN “BEAN”. SS was supported by ERC BIRTH project (GA: 640557) and Serbian Ministry of Science (GA: III47001). CP, JBu, LW, EG and YD were co-financed by the Greek-German bilateral agreement (GSRT and BMBF) project “BIOMUSE” (MIS: 5030121). LW was additionally funded by Mainz University.

Contributions

JBu, DW, and LE initiated and designed the project. SS, MB, CP, ST, NK, FG, AZL, JPec, JPet, EL, MT-N provided samples and/or archaeological and anthropological context. LW, SF, SK produced data. IS, VL and AT curated data. NM, LE, AK, EG, VP, JBu, JBl, YD, ZH, IS, AP, CRB, and DW illustrated and analysed data. NM, MB, DW, JBu, and LE wrote the paper with the help of all co-authors.

Competing interests

The authors declare no competing interests.

Supplementary Information

- Supplementary Information
- Supplementary Tables Legend
- Supplementary Tables 1-5