
Further information on publisher’s website:
https://doi.org/10.1038/s41467-019-09209-7

Publisher’s copyright statement:
This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

Additional information:

Use policy
The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full DRO policy for further details.
Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia

Michal Feldman1, Eva Fernández-Domínguez2, Luke Reynolds3, Douglas Baird4, Jessica Pearson4, Israel Hershkovitz5,6, Hila May5,6, Nigel Goring-Morris7, Marion Benz8, Julia Gresky9, Raffaela A. Bianco1, Andrew Fairbairn10, Gökhan Mustafaoğlu11, Philipp W. Stockhammer1,12, Cosimo Posth1,12, Wolfgang Haak1, Choongwon Jeong1 & Johannes Krause1

Anatolia was home to some of the earliest farming communities. It has been long debated whether a migration of farming groups introduced agriculture to central Anatolia. Here, we report the first genome-wide data from a 15,000-year-old Anatolian hunter-gatherer and from seven Anatolian and Levantine early farmers. We find high genetic continuity (~80–90%) between the hunter-gatherers and early farmers of Anatolia and detect two distinct incoming ancestries: an early Iranian/Caucasus related one and a later one linked to the ancient Levant. Finally, we observe a genetic link between southern Europe and the Near East predating 15,000 years ago. Our results suggest a limited role of human migration in the emergence of agriculture in central Anatolia.
The practice of agriculture began in the Fertile Crescent of Southwest Asia as early as 10,000 to 9000 BCE. Subsequently, it spread across western Eurasia while increasingly replacing local hunting and gathering subsistence practices, reaching central Anatolia by c. 8300 BCE.6,13

Recent genetic studies have shown that in mainland Europe, farming was introduced by an expansion of early farmers from Anatolia that replaced much of the local populations4,5. Such mode of spread is often referred to as the demic diffusion model. In contrast, in regions of the Fertile Crescent such as the southern Levant and the Zagros Mountains (located between present-day eastern Iraq and western Iran), the population structure persists throughout the Neolithic transition6, indicating that the hunter-gatherers of these regions locally transitioned to a food-producing subsistence strategy.

Central Anatolia has some of the earliest evidence of agricultural societies outside the Fertile Crescent3 and thus is a key region in understanding the early spread of farming. While archeological evidence points to cultural continuity in central Anatolia, due to the lack of genetic data from pre-farming individuals, it remains an open question whether and to what scale the development of the Anatolian Neolithic involved immigrants from earlier farming centers admixing with the local hunter-gatherers.

Likewise, pre-farming genetic links between Near-Eastern and European hunter-gatherers are not well understood, partly due to the lack of hunter-gatherer genomes from Anatolia. Genetic studies have suggested that ancient Near-Eastern populations derived a substantial proportion of their ancestry from a common outgroup of European hunter-gatherers and East Asians4,6,7. This deeply branching ancestry often referred to as Basal Eurasian likely diverged from other Eurasians before the latter received Neanderthal gene flow6. Interestingly, a previous study reported that European hunter-gatherers younger than 14,000 years ago tend to show an increased affinity with present-day Near Easterners compared to older European hunter-gatherers8, although how this affinity formed is not well understood.

Here, we report new genome-wide data from eight prehistoric humans (Fig. 1a, Table 1, and Supplementary Table 1), including the first Epipaleolithic Anatolian hunter-gatherer sequenced to date (labeled AHG; directly dated to 13,642–13,073 cal BCE, excavated from the site of Pınarbaşı, Turkey), five early Neolithic Aceramic Anatolian farmers (labeled AAF; c. 8300–7800 BCE, one directly dated to 8269–8210 cal BCE3, from the site of Boncuklu, Turkey), adding to previously published genomes from this site9, and two Early Neolithic (PPNB) farmers from the southern Levant (one labeled KFH2, directly dated to c. 7700–7600 cal BCE, from the site of Kfar HaHoresh, Israel; and the second labeled BAJ001, c. 7027–6685 cal BCE, from the site of Ba‘a, Jordan). These data comprise a genetic record stretching from the Epipaleolithic into the Early Holocene, spanning the advent of agriculture in the region.

We find that the AHG is genetically distinct from other reported late Pleistocene populations. We reveal that Neolithic Anatolian populations derive a large fraction of their ancestry from the Epipaleolithic Anatolian population, suggesting that farming was adopted locally by the hunter-gatherers of central Anatolia. We also detect distinct genetic interactions between the populations of central Anatolia and earlier farming centers to the east, during the late Pleistocene/early Holocene and describe a genetic link with European hunter-gatherers that predates 15,000 years ago.

Results
Genetic continuity and detected admixtures in Anatolia. We extracted DNA from the ancient human remains and prepared it for next-generation sequencing10,11, which resulted in human DNA yields lower than 2% (Supplementary Data 1), comparable with low DNA preservation previously reported in the region8,9. To generate genome-wide data despite the low DNA yields, we performed in-solution DNA enrichment targeting 1.24 million genome-wide single-nucleotide polymorphisms (SNPs) (“1240k capture”12), which resulted in 129,406 to 917,473 covered SNPs per individual. We estimated low mitochondrial contamination levels for all eight individuals (1–6%; see Methods and Supplementary Table 2) and could further test the males for nuclear contamination, resulting in low estimates (0.05–2.23%; Supplementary Table 2). For population genetic analyses, we merged genotype data of the new individuals with previously published datasets from 587 ancient individuals and 254 present-day populations (Supplementary Data 2).

To estimate how the ancient individuals relate to the known west Eurasian genetic variation, we projected them onto the top two dimensions (PC1, PC2) of present-day principal component analysis (PCA)6 (Fig. 1b). Strikingly, the AHG individual is positioned near both AAF and later Anatolian Ceramic farmers12 (7000–6000 cal BCE). These three prehistoric Anatolian populations (AHG, AAF, and ACF), representing a temporal transect spanning the transition into farming, are positioned along PC1 between Mesolithic western European hunter-gatherers (WHG)4,7,12 who are at one extreme of PCA and Levantine Epipaleolithic Natufians6 who are at the other. Along PC2, ancient Anatolians, WHG, and Natufians have similar coordinates. The newly reported Levantine Neolithic farmers (BAJ001 and KFH2) are positioned near the previously published Levantine Neolithic farmers6 (Supplementary Note 2). In ADMIXTURE analysis AHG, AAF, and ACF are inferred as a mixture of two components that are each maximized in Natufians and WHG, consistent with their intermediate positions along PC1 in PCA (Supplementary Figure 1).

Inspired by our qualitative observations in PCA and ADMIXTURE analyses, we applied formal statistical frameworks to describe the genetic profiles of the three Anatolian populations and to test and model genetic differences between them. We first characterized the ancestry of AHG. As expected from AHG’s intermediate position on PCA between Epipaleolithic/Neolithic Levantines and WHG, Patterson’s D-statistics13 of the form $D(\text{AHG, WHG; Natufian/Levant}_N, \text{Mbuti}) \leq 4.8$ SE (standard error) and $D(\text{AHG, Natufian/Levant}_N; \text{WHG, Mbuti}) \leq 9.0$ SE (Supplementary Table 3) indicate that AHG is distinct from both the WHG and Epipaleolithic/Neolithic Levantine populations and yet shares extra affinity with each when compared to the other. Then, we applied a qpAdm-based admixture modeling to integrate these D- statistics. qpAdm is a generalization of Di f f -statistics that test whether the target population and the admixture model (i.e., a linear combination of reference populations) are symmetrically related to multiple outgroups13. By doing so, it tests whether the proposed admixture model is adequate to explain the target gene pool and provides admixture coefficient estimates. We find an adequate two-way admixture model with $p = 0.158$, in which AHG derives around half of his ancestry from a Neolithic Levantine-related gene pool ($48.0 \pm 4.5\%$; estimate $\pm 1$ SE) and the rest from the WHG-related one (Supplementary Tables 4, 5). While these results do not suggest that the AHG gene pool originated as a mixture of Levant_N and WHG, both of which lived millennia later than AHG, it still robustly supports that AHG is genetically intermediate between WHG and Levant_N. This cannot be explained without gene flow between the ancestral gene pools of those three groups. This supports a late Pleistocene presence of both Near-Eastern and European hunter-gatherer-related ancestries in central Anatolia. Notably, this genetic link with the Levant pre-dates the advent of farming in this region by at least five millennia.
In turn, AAF are slightly shifted on PC2 compared to AHG, to the direction where ancient and modern Caucasus and Iranian groups are located. Likewise, when compared to AHG by $D(AAF, AHG; test, Mbuti)$, the AAF early farmers show a marginal excess affinity with early Holocene populations from Iran or Caucasus and with present-day south Asians, who have also been genetically linked with Iranian/Caucasus ancestry$^{14,15}$ (e.g., $D = 2.3$ and $2.7SE$ for CHG and Vishwabrahmin, respectively; Fig. 2a, Supplementary Figures 2, 3, and Supplementary Data 3). Accordingly, a mixture of AHG and Neolithic Iranians provides a good fit to AAF in our $qPAdm$ modeling ($\chi^2p = 0.296$), in which AAF derive most of their ancestry ($89.7 \pm 3.9\%$) from a population related to AHG (Supplementary Tables 4 and 6). A simpler model without contribution from Neolithic Iranians (i.e., modeling AAF as a sister clade of AHG) shows a significant reduction in model fit ($\chi^2p = 0.014$). This suggests a long-term
genetic stability in central Anatolia over five millennia despite changes in climate and subsistence strategy. The additional Neolithic Iranian-related ancestry (10.3 ± 3.9%) presumably diffused into central Anatolia during the final stages of the Pleistocene or early Holocene, most likely via contact through eastern Anatolia. This provides evidence of interactions between eastern and central Anatolia in the Younger Dryas or the first millennium of the Holocene, currently poorly documented archeologically.

In contrast, we show that the later ACF individuals share more alleles with the early Holocene Levantines than AAF do, as shown by positive $D(ACF, AAF; Natu\textsc{iers}, Mbuti) \geq 3.8$ SE (Fig. 2b, Supplementary Figures 4, 5, and Supplementary Data 3). Ancient Iran/Caucasus populations and contemporary South Asians do not share more alleles with ACF ($|D| < 1.3$ SE). Likewise, $qpAdm$ modeling suggests that the AAF gene pool still constitutes more than 3/4 of the ancestry of ACF 2000 years later (78.7 ± 3.5%; Supplementary Tables 4 and 7) with additional ancestry well modeled by the Neolithic Levantines ($\chi^2 p = 0.115$) but not by the Neolithic Iranians ($\chi^2 p = 0.076$; the model estimated infeasible negative mixture proportions) (Supplementary Tables 4 and 7). These results suggest gene flow from the Levant to Anatolia during the early Neolithic. In turn, Levantine early farmers (Levant\textsc{Neol}) that are temporally intermediate between AAF and ACF could be modeled as a two-way mixture of Natufians and AHG or AAF (18.2 ± 6.4% AHG or 21.3 ± 6.3% AAF ancestry; Supplementary Tables 4 and 8 and Supplementary Data 4), confirming previous reports of an Anatolian-like ancestry contributing to the Levantine Neolithic gene pool. These two distinct detected gene flows support a reciprocal genetic exchange between the Levant and Anatolia during the early stages of the transition to farming.

**Genetic links between Pleistocene Europe and the Near East.** AHGs and inhabited a region that connects Europe to the Near East. However, pre-Neolithic interactions between Anatolia and Southeastern Europe are so far not well documented archeologically. Interestingly, a previous genomic study showed that present-day Near Easterners share more alleles with European hunter-gatherers younger than 14,000BP (“Later European HG”) than with older ones (“Earlier European HG”). With ancient genomic data available, we could directly compare the genetic affinity of European hunter-gatherers with Near-Eastern hunter-gatherers (AHG and Natufian) using the $D$-statistic of the form $D(\text{European hunter-gatherers, Kostenki14}; \text{AHG}/\text{Natufian}, \text{Mbuti})$. We compared the European hunter-gatherers to the 37 thousand-year-old individual Kostenki14$^{14,17}$ representing the oldest available European genome with genetic affinity to later European hunter-gatherers (Fig. 3a and Supplementary Data 5). As is the case for present-day Near Easterners, this statistic is significantly positive for all European hunter-gatherers younger than 14,000BP. Most of the Later European HGs belong to a largely homogeneous gene pool referred to as the “Villabruna cluster,”$^{18}$ named after its oldest available member from an Epigravettan site in northern Italy. Our results suggest that the non-Basal Eurasian ancestry of ancient Anatolians and Levantines derived from a gene pool related to the Villabruna cluster prior to its expansion within Europe observed after 14,000BP.

Among the Later European HG, recently reported Mesolithic hunter-gatherers from the Balkan peninsula, which geographically connects Anatolia and central Europe (“Iron Gates HG”)$^{18}$, show the highest genetic affinity to AHG and the second highest one to Natufians, as shown in the positive statistic $D(\text{Iron Gates HG}, \text{European hunter-gatherers}; \text{AHG}/\text{Natufian}, \text{Mbuti})$ (Supplementary Figures 6 and 7). This affinity is surprising considering that Iron Gates HG have been previously modeled as a mixture of WHG (~85%) and eastern European hunter-gatherers (EHG; ~15%)$^{18}$, the latter of which shares a much lower affinity with ancient Near Easterners in respect to other European HG (Fig. 3a). Since the previously reported WHG and EHG model did not fit well ($\chi^2 p = 0.0003$) and since Iron Gates HG harbored Near-Eastern-like mitochondrial groups, an affinity with Anatolians beyond the WHG + EHG model has been hypothesized$^{18}$. Accordingly, we find that Iron Gates HG can be modeled as a three-way mixture of Near-Eastern hunter-gatherers (25.8 ± 5.0% AHG or 11.1 ± 2.2% Natufian), WHG (62.9 ± 7.4% or 78.0 ± 4.6%, respectively) and EHG (11.3 ± 3.3% or 10.9 ± 3%,

### Table 1 An overview of ancient genomes reported in this study

<table>
<thead>
<tr>
<th>ID</th>
<th>Library name</th>
<th>Analysis group</th>
<th>Estimated date</th>
<th>Site</th>
<th>Sampled tissue</th>
<th>Total sequenced reads ($\times 10^6$)</th>
<th>Human DNA (%)</th>
<th>Mean coverage (fold)</th>
<th>Genetic sex</th>
<th>mt</th>
<th>Ychr</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZBC</td>
<td>IP001/B/C0101</td>
<td>AHG</td>
<td>13,642–13,073 cal BCE</td>
<td>Pinarbaşı Intermediate phalanx Boncuklu Petroş</td>
<td>126.7</td>
<td>33</td>
<td>2.9</td>
<td>Male</td>
<td>K2b</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>ZHAG</td>
<td>BON004/A0101</td>
<td>AAF</td>
<td>8300–7800 BCE</td>
<td>Boncuklu third molar Boncuklu Petroş</td>
<td>92.0</td>
<td>38</td>
<td>1.48</td>
<td>Female</td>
<td>N1a1a1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZMOJ</td>
<td>BON014/A0101</td>
<td>AAF</td>
<td>8300–7800 BCE</td>
<td>Boncuklu third molar Boncuklu Petroş</td>
<td>77.9</td>
<td>27</td>
<td>0.8</td>
<td>Male</td>
<td>K1a</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>ZKO</td>
<td>BON001/A0101</td>
<td>AAF</td>
<td>8300–7800 BCE</td>
<td>Boncuklu third molar Boncuklu Petroş</td>
<td>84.8</td>
<td>31</td>
<td>0.9</td>
<td>Male</td>
<td>U3</td>
<td>G2a2b2b</td>
<td></td>
</tr>
<tr>
<td>ZHJ</td>
<td>BON024/A0101</td>
<td>AAF</td>
<td>8300–7800 BCE</td>
<td>Boncuklu third molar Boncuklu Petroş</td>
<td>87.7</td>
<td>38</td>
<td>0.76</td>
<td>Female</td>
<td>U3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZHAI</td>
<td>BON034/A0101</td>
<td>AAF</td>
<td>8269–8210 cal BCE</td>
<td>Boncuklu third molar Boncuklu Petroş</td>
<td>75.4</td>
<td>30</td>
<td>0.69</td>
<td>Female</td>
<td>U3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KFH2</td>
<td>KF002/A0101</td>
<td>Levant\textsc{Neol}</td>
<td>7712–7589 cal BCE</td>
<td>Kfar Hahoersh Ba'ja</td>
<td>342.0</td>
<td>8</td>
<td>0.16</td>
<td>Female</td>
<td>N1a1b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAJ01</td>
<td>BAJ001/A0101</td>
<td>Levant\textsc{Neol}</td>
<td>7027–6685 cal BCE</td>
<td>Kfar Hahoersh Ba'ja</td>
<td>17.3</td>
<td>45</td>
<td>0.75</td>
<td>Female</td>
<td>N1b1a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For each individual the analysis group is given (AHG — Anatolian hunter-gatherer; AAF — Anatolian Aceramic farmers; Levant\textsc{Neol} — Levantine early farmer). When $^{13}C$ dating results are available, the date is given in cal BCE in 2-sigma range, otherwise a date based on the archeological context is provided (detailed dating information is provided in Supplementary Note 1 and Supplementary Table 1). The proportion of human DNA and the mean coverage on 1240k target sites are listed. Detailed information on the uniparental analysis can be found in Supplementary Note 1 and Supplementary Data 6.
respectively); ($\chi^2 p = 0.308$ and $\chi^2 p = 0.589$ respectively; Supplementary Tables 4 and 9).

To further test the model of Near-Eastern gene flow into the ancestors of Iron Gates HG as an explanation of the extra affinity between them, we utilized the Basal Eurasian ancestry that was widespread in early Holocene and late Pleistocene Near-Eastern populations and their descendants but undetectable in European hunter-gatherers, as a tracer for gene flow from the Near East. To estimate the Basal Eurasian ancestry proportion (“$\alpha$”), we followed a previously established qpAdm-based approach that uses an African reference (the ancient Ethiopian Mota genome) as a proxy (Supplementary Table 10). We estimated $\alpha$ to be 24.8
Fig. 2 Differences in genetic affinities between the ancient Anatolian populations. We plot the highest and lowest 40 values of $D$(population 1, population 2; test, Mbuti) on the map. Circles mark ancient populations and triangles present-day ones. “Test” share more alleles with population 1 when values are positive and with population 2 when negative. The detected gene flow direction is illustrated in the upper schematics; the illustrated route represents the shortest one between the proximate source and the target and should not be interpreted as the historic route of the gene flow. The statistics and SEs are found in Supplementary Figures 2-5 and Supplementary Data 3. a Early Holocene Iranian and Caucasus populations, as well as present-day South Asians, share more alleles with Aceramic Anatolian farmers (AAF) than with Anatolian hunter-gatherers (AHG), measured by positive $D$(AAF, AHG, test, Mbuti). The top 10 values with $\pm 1\sigma$ and $\pm 3\sigma$ SE are shown in the upper box. b Ancient Levantine populations share more alleles with Anatolian Ceramic farmers (ACF) than with AAF, measured by positive $D$(ACF, AAF; test, Mbuti). The top 10 values with $\pm 1\sigma$ and $\pm 3\sigma$ SE are shown in the lower box. Source data are provided as a Source Data file.

Uniparental markers and phenotypic analysis. The uniparental marker analysis placed AHG within the mitochondrial subhaplogroup K2b and within the Y-chromosome haplogroup C1a2, both rare in present-day Eurasians (Table 1 and Supplementary Data 6). Mitochondrial Haplogroup K2 has so far not been found in Paleolithic hunter-gatherers. However, Y-haplogroup C1a2 has been reported in some of the earliest European hunter-gatherers. The early farmers belong to common early Neolithic mitochondrial (N1a, U3 and K1a) and Y chromosome types (C and G2a), with the exception of the Levantine BAJ001, which represents the earliest reported individual carrying the mitochondrial N1b group (Table 1 and Supplementary Data 6).

We examined alleles related to phenotypic traits in the ancient genomes (Supplementary Data 7). Notably, three of the AAF carry the derived allele for rs12193832 in the HERC2 (hect domain and RLD2) gene that is primarily responsible for lighter eye color in Europeans. The derived allele is observed as early as 14,000–13,000 years ago in individuals from Italy and the Caucasus, but had not yet been reported in early farmers or hunter-gatherers from the Near East.

Discussion

By analyzing genome-wide-data from pre-Neolithic and early Neolithic Anatolians and Levantines, we describe the demographic developments leading to the formation of the Anatolian early farmer population that later replaced most of the European hunter-gatherers and represents the largest ancestral component in present-day Europeans.

We report a long-term persistence of the local AHG gene pool over seven millennia and throughout the transition from foraging to farming. This demographic pattern is similar to those previously observed in earlier farming centers of the Fertile Crescent and differs from the pattern of the demic diffusion-based spread of farming into Europe. Our results provide a genetic support for archeological evidence, suggesting that Anatolia was not merely a stepping stone in a movement of early farmers from the Fertile Crescent into Europe, but rather a place where local hunter-gatherers adopted ideas, plants, and technology that led to agricultural subsistence.

Interestingly, while we observe a continued presence of the AHG-related gene pool throughout the studied period, a pattern of genetic interactions with neighboring regions is evident from as early as the Late Pleistocene and early Holocene. In addition to the local genetic contribution from earlier Anatolian populations, Anatolian Aceramic farmers inherit about 10% of their genes from a gene pool related to the Neolithic Iranian/Caucasus while later ACF derive about 20% of their genes from another distinct gene pool related to the Neolithic Levant.

Wide temporal gaps between available genomes currently limit our ability to distinguish the mode of transfer. Obtaining additional genomic data from these regions as well as from geographically intermediate populations of eastern Anatolia and the greater Mesopotamia region could help determine how these genetic changes happened in central Anatolia: for example, whether by a short-term massive migration or a low-level back-ground gene flow in an isolation by distance manner.

To the west, we observe a genetic link between the Anatolian and European Pleistocene hunter-gatherers, which extends the temporal frame of the previously reported genetic affinity between late Pleistocene Europeans and present-day Near-Eastern populations. Especially, Mesolithic Southeastern European hunter-gatherers (Iron Gates HG) show a strong genetic affinity with AHG. Our analysis on their Basal Eurasian ancestry proportions, although limited in resolution, suggests that a Near-Eastern gene flow from AHG into the ancestors of Iron Gates HG may not be sufficient to explain this affinity. Two additional scenarios, both involving gene flow from the ancestors of Iron Gates HG to the ancestors of AHG, can help explain the extra affinity between Iron Gates HG and AHG. One assumes a secondary gene flow from Southeastern Europe to Anatolia after the initial formation of the Near-Eastern gene pool as a mixture of the Basal Eurasian and the Villabruna-related gene pools. The other assumes that Iron Gates HG are indeed the most closely related group among European hunter-gatherers to the Villabruna-related ancestry in ancient Near Easterners. Further sampling in Anatolia and Southeastern Europe is needed to specify the spatiotemporal extent of the genetic interactions that we observe.

Methods

aDNA analysis. We extracted and prepared DNA for next-generation sequencing in two different dedicated ancient DNA (aDNA) facilities (Liverpool and Iena).

In Liverpool, UK, sampling and extraction steps for the individuals from Pinarbaşı and Boncuklu were carried out in the aDNA labs at the Liverpool John Moores University. The outer layer of the bone was removed using powdered aluminum oxide in a sandblasting instrument. Then, the bone was ultraviolet (UV) irradiated for 10 min on each side and ground into fine powder using a cryogenic grinder Freezer/Mill. DNA was extracted from 100 mg of bone powder following an established protocol. The extraction included incubation of the bone powder
in 1 ml of extraction buffer (0.45 M EDTA, pH 8.0, and 0.25 mg ml⁻¹ proteinase K) at 37 °C for over 12–16 h. Subsequently, DNA was bound to a silica membrane using a binding buffer containing guanidine hydrochloride and purified in combination with the High Pure Viral Nucleic Acid Large Volume Kit (Roche). DNA was eluted in 100 μl of TET (10 mM Tris-HCl, 1 mM EDTA, pH 8.0, and 0.05% Tween-20). One extraction blank was taken along. The extracts were then shipped to Jena, Germany where downstream processing was performed.

In Jena, Germany, all pre-amplification steps were performed in dedicated aDNA facilities of the Max Planck Institute for the Science of Human History (MPI-SHH). The inner ear part of the petrous bones of the individuals from Kfar HaHoresh and Ba’ja was sampled by drilling and DNA was extracted from 76 to 109 mg of the bone powder. An extraction of ~100 mg pulverized bone from the Pinaraşı individual ZBC was done in the Jena facility in addition to the Liverpool extraction (the sequenced data from the two extracts of individual ZBC were merged in downstream analysis after passing the quality control step). All extractions followed the same protocol as cited for Liverpool. A 20 µl aliquot from each extract was used to prepare an Illumina double-stranded, double-indexed DNA library following established protocols. Deaminated cytosines that result from DNA damage were partially removed using uracil-DNA glycosylase and endonuclease VIII, but still retained in terminal read positions as a measure of

Fig. 3 Genetic links between Near-Eastern and European hunter-gatherers. a Genetic affinity between Near-Eastern and European hunter-gatherers increases after 14,000 years ago as measured by the statistic D(European HG, Kostenki14; Natufian/AHG, Mbuti). Vertical lines mark ± 1 SE. Data points for which D > 3 SE are outlined. Kostenki14 serves here as a baseline for the earlier European hunter-gatherers. Statistics including all analyzed European hunter-gatherers are listed in Supplementary Data 5. Individuals marked with an asterisk did not reach the analysis threshold of over 30,000 single-nucleotide polymorphisms (SNPs) overlapping with Natufian/AHG. b Basal Eurasian ancestry proportions (α) as a marker for Near-Eastern gene flow. Mixture proportions inferred by qpAdm for the Anatolian hunter-gatherer (AHG) and the Iron Gates hunter-gatherers (Iron Gates HG) are schematically represented. The lower schematic shows the expected α in Iron Gates HG under assumption of unidirectional gene flow, inferred from α in the AHG source population. The observed α for Iron Gates HG is considerably smaller than expected; thus, the unidirectional gene flow from the Near East to Europe is not sufficient to explain the affinity between Iron Gates HG and AHG. Source data are provided as a Source Data file.
aDNA authentication. A negative library control (H2O) was taken along for each experiment. Unique combinations of two indexes (8 bp length each) were assigned to each library. The indexes were then attached through a ten-cycle amplification reaction using the Pfu Turbo CX Hotstart DNA Polymerase (Agilent), the PCR products purified using a Qiagen MinElute kit (Qiagen), and then eluted in TET (10 mM Tris-HCl, 1 mM EDTA, pH 8.0, and 0.05% Tween-20). Subsequently, indexed libraries were amplified using Herculase II Fusion DNA polymerase, following the manufacturer’s protocol, to a total of 10^9 DNA copies per reaction and again purified using a Qiagen MinElute kit (Qiagen) and eluted in TET (10 mM Tris-HCl, 1 mM EDTA, pH 8.0 and 0.05% Tween-20). Finally, all samples were diluted and pooled (10 nM) for sequencing. The indexed amplified libraries were also used for two previously published downstream in-solution enrichments: a protocol targeting 1,257,207 genome-wide SNPs (“1240k capture”) and one targeting the entire human mitochondrial genome.

The “1240k capture” is an established in-solution enrichment assay based on hybridization of the indexed libraries to DNA probes. The targeted SNP panel is a combination of the two separate SNP sets first reported by Haak et al. and by Fu et al. and further described by Mathieson et al. Each of the −1.2 million target SNPs, we used four distinct 52-bp-long probes: two flanking the target SNP from each side and the other two centered on the SNP matching with a mismatch in one of the index bases.

Both the initial shotgun and target-enriched libraries were single-end sequenced on an Illumina HiSeq 4000 platform (1 × 75 + 2 cycles). Sequenced reads were demultiplexed allowing one mismatch in each index and further processed using EAGER (v 1.92.54). First, adapter sequences were clipped and reads shorter than 80 bp were discarded using AdapterRemover (v 2.2.0). Adapter-clipped reads were subsequently mapped with the BWA aln/sam programs (v 0.7.12) against the UCSC genome browser’s human genome reference hg19 with a lenient stringency parameter (“*=0”). We retained reads with Phred-scaled mapping quality scores ≥20 and ≥30 for the whole genome and the mitochondrial genome, respectively. Duplicate reads were subsequently removed using DeDup v 0.12.229. Pseudo-diploid genotypes were generated for each individual using pileupCaller, which randomly draws a high quality base (Phred-scaled base quality score ≥30) mapping to each targeted SNP position (https://github.com/stschiff/sequenceTools). To prevent false SNP calls due to retained DNA damage, two terminal positions in each read were clipped prior to genotyping. The genotyping produced between 129,406 and 917,473 covered targeted SNPs and a mean coverage ranging between 0.16 and 2.9 fold per individual (Table 1).

Dataset. We merged the newly reported ancient data and data reported by Mathieson et al. 2018 with a dataset that has been described elsewhere. This dataset includes 587 published ancient genomes from 759 individuals, representing world-wide present-day populations, and genomes from 2,706 individuals, representing world-wide prehistoric present-day populations, that were genotyped on the Affymetrix Axiom™ Genome-Wide Human Origins 1 array ("HO dataset") with a total of 597,573 SNP sites in the merged dataset. To minimize bias from differences in analysis pipelines, we re-processed the raw read data deposited for previously published Neolithic Anatolian genomes (labeled “Tepheciyub” and “Bonacciu, pub”) in the same way as described for the newly reported individuals.

aDNA authentication and quality control. We estimated authenticity of the ancient data using multiple measures. First, blank controls were included and analyzed for extractions as well as library preparations (Supplementary Data 8). Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.

Second, we assessed levels of DNA damage in the mapped reads using mapDamage (v 2.0)37. Third, we estimated human DNA contamination on the mitochondrial genome targeting the entire human mitochondrial genome27.
each of the reported SNPs was confirmed by visually inspecting the bam pileup in Geneious (v11.0.8)59. The resulting consensus sequences were then analyzed with HaploFind32 and HaploParser37 to assign haplogroups and double-checked with the rCRS oriented version of Phylotree33.

Y-chromosome analysis. To assign Y-chromosome haplogroups we used yHaplo64. Each male individual was genotyped at 13,508 ISOGG consortium SNP positions (strand-ambiguous SNPs were excluded) by randomly drawing a single base mapping to the SNP position, using the same quality filters as for the HO dataset. In addition to the yHaplo automated haplogroup designations, we manually verified the presence of derived alleles supporting the haplogroup assignment.

Phenotypic traits analyses. We tested for the presence of alleles related to biological traits that could be of interest in the geographical and temporal context of the reported ancient populations, including lactose persistence55,56, Malaria resistance55,58, glucose-6-phosphate dehydrogenase deficiency59,60, and skin pigmentation36,61,62. The allele distribution for the SNP positions listed in Supplementary Data 7 was tabulated for each individual using Samtools mpileup (v 1.3).

Carbon dating. The phalanx bone from individual ZBC (Pinarbaşı) and the petrous bone from individual KFH2 (Kfar HaHoresh) were each sampled and directly radiocarbon dated at the CEZ Archaeometry gGmbH, Mannheim, Germany (Supplementary Table 1). Collagen was extracted from the bone samples, purified by ultrafiltration (fraction >30kDa), freeze-dried, and combusted to CO2 in an elemental analyzer. CO2 was converted catalytically to graphite. The dating was performed using the MCADAS-AMS of the Klaus-Tschira-Archäometrie-Zentrum. The resulting 14C ages were normalized to d13C = −25.03 and calibrated using the dataset INTCAL1364 and the software SwissCal 1.065.

Received: 20 June 2018 Accepted: 15 February 2019
Published online: 19 March 2019

References
Acknowledgements

We thank G. Brandt, A. Wissgott, F. Aron, M. Burri, C. Freund, and R. Stahl (MPI-SHH) for their support in laboratory work, M. Oredry for graphic support, A. Gibson for help in proofreading, and the members of the population genetics group in the Department of Archaogenetics, SHH-MPI for their input and support. This work was supported by the Max Planck Society. Experimental work at LJMU and LR stay at the MPI-SHH were funded with an internal grant from LJMU Faculty of Science to EFD. The funding for the Ba‘a project was granted by the German Research Foundation (GZ 80 1599/14–1) and ex oriente e.V., Berlin. Klar Halloresh fieldwork was supported (to N.G.-M.) by the Israel Science Foundation (Grants 846/01, 558/04, 755/07, 1161/10), The National Geographic Society (Grant 8625/09), and The Irene Levi Sala CARE Foundation. The Konya Plain fieldwork was funded by The British Institute in Ankara, British Academy (Research Development Award BR100077), a British Academy Large Research Grant LRG 35439, Australian Research Council (grants DP0663385 and DP120100969), National Geographic award GEFNE 1-11, University of Oxford (Wainwright Fund), Australian Institute for Nuclear Science and Engineering (AINSE awards AINGRA05051 and AINGRA10069), Wenner-Gren Foundation for Anthropological Research (Postdoctoral Research Grant 2008 The Origins Of Farming In The Konya Plain, Central Anatolia, Institute for Field Research.

Author contributions


Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-09209-7.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: Nature Communications thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019