Population Genetics and Signatures of Selection in Early Neolithic European Farmers

Ainash Childebayeva,*,1,2 Adam Benjamin Rohrlach,1,2 Rodrigo Barquera,1,2 Maïté Rivollat,1,4 Franziska Aron,1 András Szolek,5,6 Oliver Kohlbacher,5,7,8,9 Nicole Nicklisch,10,11 Kurt W. Alt,10,11 Detlef Gronenborn,12 Harald Meller,11 Susanne Friederich,11 Kay Prüfer,1,2 Marie-France Deguilloux,4 Johannes Krause,1,2 and Wolfgang Haak *,1,2

1Archaeogenetics Department, Max Planck Institute for the Science of Human History, Kahlaische Straße 10, D-07745 Jena, Germany
2Archaeogenetics Department, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany
3ARC Centre of Excellence for Mathematical and Statistical Frontiers, School of Mathematical Sciences, The University of Adelaide, Adelaide, Australia
4Université de Bordeaux, CNRS, PACEA-UMR 5199, 33615 Pessac, France
5Applied Bioinformatics, Dept. of Computer Science, University of Tübingen, Tübingen, Germany
6Department of Immunology, Interfaculty Institute for Cell Biology, University of Tübingen, Tübingen, Germany
7Institute for Bioinformatics and Medical Informatics, University of Tübingen, Tübingen, Germany
8Translational Bioinformatics, University Hospital Tübingen, Tübingen, Germany
9Biomolecular Interactions, Max Planck Institute for Developmental Biology, Tübingen, Germany
10Center of Natural and Cultural Human History, Danube Private University, Krems-Stein, Austria
11State Office for Heritage Management and Archaeology Saxony-Anhalt - State Museum of Prehistory, Halle (Saale), Germany
12Römisch-Germanisches Zentralmuseum, Leibniz Research Institute for Archaeology, Ernst-Ludwig-Platz 2, 55116 Mainz, Germany

*Corresponding authors: E-mails: ainash_childebayeva@eva.mpg.de; wolfgang_haak@eva.mpg.de.

Abstract

Human expansion in the course of the Neolithic transition in western Eurasia has been one of the major topics in ancient DNA research in the last 10 years. Multiple studies have shown that the spread of agriculture and animal husbandry from the Near East across Europe was accompanied by large-scale human expansions. Moreover, changes in subsistence and migration associated with the Neolithic transition have been hypothesized to involve genetic adaptation. Here, we present high quality genome-wide data from the Linear Pottery Culture site Derenburg-Meerestieg II (DER) (N = 32 individuals) in Central Germany. Population genetic analyses show that the DER individuals carried predominantly Anatolian Neolithic-like ancestry and a very limited degree of local hunter-gatherer admixture, similar to other early European farmers. Increasing the Linear Pottery culture cohort size to ~100 individuals allowed us to perform various frequency- and haplotype-based analyses to investigate signatures of selection associated with changes following the adoption of the Neolithic lifestyle. In addition, we developed a new method called Admixture-informed Maximum-likelihood Estimation for Selection Scans that allowed us test for selection signatures in an admixture-aware fashion. Focusing on the intersection of results from these selection scans, we identified various loci associated with immune function (JAK1, HLA-DQB1) and metabolism (LMF1, LEPR, SORBS1), as well as skin color (SLC24A5, CD82) and folate synthesis (MTHFR, NBPF3). Our findings shed light on the evolutionary pressures, such as infectious disease and changing diet, that were faced by the early farmers of Western Eurasia.

Key words: selection, neolithization, ancient DNA, Linear Pottery culture.

Introduction

The Neolithic transition, a shift from a foraging to a farming-based subsistence, marks one of the most substantial social, ecological, and economic transformations in the prehistory of West Eurasia. The process of neolithization has interested and fascinated archaeologists, anthropologists, demographers, epidemiologists, specialists in evolutionary medicine, and the general public for more than a 100 years (Lubbock 1866; Dawkins 1894; Bogucki 1999; Johnson and Earle 2000; Chamberlain 2006; Bocquet-Appel and Bar-Yosef 2008). In retrospect, the Neolithic was a transformative period in which a mostly sedentary lifestyle, domestication of plants and animals, and the associated relative independence from nature with
reliance on sustained food supplies came into being. A manufacturing economy with good yields, surplus, and stockpiling caused an increase in fertility, with earlier age at weaning and subsequent shorter birth intervals which brought about an exponential population increase (Deevey 1960; Bocquet-Appel and Bar-Yosef 2008). The negative effects of neolithization manifested in an increased prevalence and transmission of infectious, metabolic, and nutritional deficiency diseases, favored by sedentary lifestyles, growing populations, and close contact with domesticated animals (Barrett et al. 1998; Cordain 1999; Harper and Armelagos 2010; Gering et al. 2019).

Archaeological research has shown that agriculture and animal husbandry originated in the Near East around 10,000 BCE (Zeder 2011), from where these practices spread to North Africa, Europe, and South Asia. With the help of ancient DNA (aDNA) studies, it was shown that the spread of agriculture across Europe was not a transfer of ideas and technologies but was instead mediated by the expansion of the early farmers (EFs), that is, through demic diffusion (Haak et al. 2010, 2015; Gamba et al. 2014; Lazaridis et al. 2014; Skoglund et al. 2014, 2012; Mathieson et al. 2015; Hofmanová et al. 2016). By ~6,000 BCE, sedentary farming was established in the Aegean (Horejs et al. 2015), southeastern Europe and the Carpathian Basin (Bánhffy 2019), and direct connections have been established between the early European farmers (EEFs) from Neolithic Hungary and Greece and the EFs from western Anatolia (Mathieson et al. 2015; Hofmanová et al. 2016; Lazaridis et al. 2016; Lipson et al. 2017). Confirming previous archaeological findings (Zvelebil 2001; Gronenborn 2014), genetic studies have reported evidence for two routes of EF expansion from the Near East: the inland Central European route; and the coastal (Mediterranean) route to western Europe, Iberia (Oalde et al. 2015; Brunel et al. 2020; Rivollat et al. 2020), and eventually the Atlantic Archipelago (Cassidy et al. 2016; Brace et al. 2019; Rohrlach et al. 2021).

The Linear Pottery Culture (or LBK after German Linearbandkeramik) is associated with early accounts of farming in Neolithic Europe (Zvelebil 2004; Bánhffy 2019; Petrasch 2020). The LBK culture originated in the western Carpathian Basin in today’s Hungary and Slovakia between 5500 and 5400 BCE and spread across the European loess plains to the Paris Basin in Western Europe (Bickle 2009) and Ukraine in eastern Europe. Archaeologically, LBK emerged from interactions between Starčevo-Körös-Criș complex farmers and local Mesolithic hunter-gatherers (HG) (Gronenborn 2007). Numerous LBK sites have been found in Central Germany, particularly in the Mittelelbe-Saale (MES) region, a biogeographical region that was attractive to the early Neolithic farmers due to its fertile soils, waterways, and adequate levels of precipitation.

The emergence of agriculture in Western Eurasia has been linked to new environmental and cultural pressures. During this time, HG groups in the area of the Fertile Crescent transitioned to a more sedentary lifestyle associated with increased population sizes, a change in diet and, importantly, increased exposure to infectious diseases from animals, for example, Salmonella enterica (Armelagos et al. 1991; Barrett et al. 1998; Harper and Armelagos 2010; Key et al. 2020). Successful settlement in new geographic regions, such as central Europe, would have required rapid adaptation to new environmental, economic, and social conditions. Lower meat consumption and increased caries lesions in teeth indicative of cereal consumption have been reported for the early Neolithic compared to later time periods (Nicklisch et al. 2016; Münster et al. 2018). Moreover, increased prevalence of cribra orbitalia and porotic hyperostosis, which indicate either a significant burden of infectious diseases, low quality diet, or a combination of both (Walker et al. 2009), was found at LBK sites (Ash et al. 2016; Nicklisch 2017). A previous study of selection in Europe, using aDNA, in general reported evidence of selection for light skin color, infectious disease resistance, and fatty acid metabolism, among other signals (Mathieson et al. 2015). In addition, Neolithic Aegeans have been shown to carry alleles associated with reduced skin pigmentation and type 2 diabetes susceptibility, and a number of inflammatory disease-associated loci (Hofmanová et al. 2016). Taken together, early Neolithic farmers likely faced selective pressures from the increased exposure to pathogens and the change in diet.

Here, we report new genome-wide data from humans remaining associated with an LBK burial site Derenburg-Meerensieng II (Wernigerode, Saxony-Anhalt, Germany). The Derenburg (DER) burial ground is located in Saxony-Anhalt, close to the Holtemme tributary of the Bode River, in the northern Harz foreland (fig. 1). Previous studies using stable isotopes have shown that mean adult human values for carbon and nitrogen isotopes were typical for the region and suggested a mixed farming diet including domesticated plant and animal products (Oelze et al. 2011; Münster et al. 2018). A previous study of mitochondrial DNA (mtDNA) lineages from the EEF, including the individuals in this study, showed that the mtDNA haplogroups of LBK individuals and their frequency distribution is more similar to the present-day population of Anatolia and the Near East (Haak et al. 2010). This indirectly argued for demic diffusion from Anatolia to Europe and a degree of genetic continuity between the two regions, which was later confirmed using genome-wide studies (Mathieson et al. 2015; Hofmanová et al. 2016; Lazaridis et al. 2016; Lipson et al. 2017; Rivollat et al. 2020).

One of our aims was to provide an updated genomic portrait of DER in the context of genome-wide variation of the EEF. We explored the genomic ancestry of LBK individuals comparing them to other EEF and Anatolian EFS, as well as local European Western Hunter-Gatherers (WHGs) on a broader scale. We also investigated the intra-site biological relatedness and explored the demography of LBK groups by comparing DER individuals to previously published individuals from nearby sites, such as Halberstadt (only ~10–15 km apart), and other LBK sites from Germany, Poland, Austria, and Hungary (Lipson et al. 2017; Mathieson et al. 2018; Rivollat et al. 2020).
New Approaches

Leveraging the substantially increased cohort size of the EEFs, we performed various scans of selection to determine specific genes and pathways that were under selection in LBK individuals, as well as the EFs in general. To our knowledge, genetic adaptation associated with the Neolithic lifestyle has not been formally tested using aDNA. We performed allele frequency—[LSBL and our own newly developed method called Admixture-informed Maximum-likelihood Estimation for Selection Scans (AIMLESS)] and haplotype-based selection scan (XP-EHH), and analyses of the HLA class I and II alleles, to investigate whether the LBK population was under an evolutionary pressure as a result of the Neolithization and population migration/expansion. Moreover, we performed selection tests comparing EF individuals to modern-day African populations and the local HGs to test general adaptation to the Neolithization in Western Eurasia.

The ability to detect true loci that are under selection is based on how well the population structure is accounted for, especially in a test that depends on comparisons between populations. Thus, we developed AIMLESS to test for selection in an ancestry-aware fashion. This method is a modification of the admixture-aware linear model-based test developed by Mathieson et al. (2015); however, our test is specifically designed for two population admixture scenarios. AIMLESS is similar to other likelihood ratio test (LRT)-based tests such as AdaptMix (Mendoza-Revilla et al. 2021) and Ohana (Cheng et al. 2022) but does not rely on a priori ancestry estimation and determines the mixture components for the ancestries considered for each single-nucleotide polymorphism (SNP) independently instead of using a genome-wide estimate.

Results and Discussion

Data Generation and Authentication

We generated genome-wide SNP capture data for ~1.24 million variant sites (1,240k SNP array) across the genome (Fu et al. 2015; Mathieson et al. 2015), mitochondrial genome capture (Maricic et al. 2010), Y-chromosome capture (Fu et al. 2015; Mathieson et al. 2015), and mitochondrial capture (Rohrlach et al. 2021) and 3 Mb immune-capture data (Immel et al. 2021) from N = 32 individuals. In addition, we sequenced the complete genomes of two DER individuals (DER002 2.26X and DER009 3.07X), and a local HG individual from Bad Dürenberg, Germany (BDB001 13x), for which genome-wide SNP capture data was previously reported (Rivollat et al. 2020). A subset of DER individuals (N = 10) was radiocarbon dated to ca. 5,300–4,800 cal. BCE (supplementary table S1, Supplementary Material online) or ~7,000 years before present (BP), which is consistent with the younger phase of the LBK period in Germany (Gronenborn et al. 2014).

The authenticity of human aDNA from the DER individuals was confirmed using MapDamage (Ginolhac et al. 2011), wherein all aDNA fragments were shown to have deamination patterns consistent with degradation over time, and unambiguous sex determination. Male samples were checked for X chromosome contamination from modern sources, using ANGS (Kornelusssen et al. 2014). Mitochondrial DNA contamination was estimated using Schmutzi (Renault et al. 2015) and contamMix (Fu et al. 2013). No significant contamination (>5%) from modern sources was determined in the DER cohort (supplementary table S2, Supplementary Material online). Based on the genetic sex determination, we identified 18 females and 14 males (54.5% female).

Genome-Wide Data Analysis

We genotyped the 1,240k SNP positions using pileupCaller (https://github.com/strichff/sequenceTools/tree/master/src-pileupCaller) obtaining pseudohaploid calls for each variant position on the capture. We merged the 1,240k SNPs with the modern-day Human Origins data set, as well as relevant ancient samples (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data, version 44.3), to perform a principal component analysis (PCA). Owing to the incomplete nature of the ancient data, ancient individuals were projected onto the genetic variation of 1363 modern West Eurasians (fig. 1A). Based on the position in PC space, we found that the DER individuals fall together with published LBK individuals, as well as contemporaneous early Neolithic individuals from southeastern Europe and western Anatolia (fig. 1A). We then performed an unsupervised ADMIXTURE analysis and found that at K = 12 DER represent a mixture of an ancestry component maximized in Anatolian Neolithic individuals and a smaller proportion of ancestry component maximized in WHGs (fig. 1B).

Following the PCA and ADMIXTURE analyses, we calculated various F-statistics to formally test the ancestry components identified in the admixture analysis and inferred from the position on the PCA plot. We first performed an f4-test of the form f4(Whacht, HG; DER, Anatolia_N) to test whether DER individuals carried additional HG ancestry compared with Anatolia_N (fig. 2A). Significantly negative f4-values (|z| > 3) for models in which England_Mesolithic, Villabruna, Loschbour HG, France_Rochedane, and France_Chauderdes were used as HG, indicated excess WHG ancestry in DER. We also obtained a significantly negative f4-value for the Iron Gates HGs (Mathieson et al. 2018).

In a follow-up f4-analysis of the form f4(Whacht, HG; EEF, Anatolia_N) we iterated through a number of EEF groups and tested the following two HG sources specifically: Loschbour HG (WHG) and Iron Gates HGs (supplementary fig. S1, Supplementary Material online). In addition, we tested BDB001 (Rivollat et al. 2020) as a local HG ancestry proxy from the MES region. The DER individuals showed significantly negative f4-results suggesting more HG ancestry compared to Anatolia_N in the first two cases (Loschbour HG and Iron Gates), but not for the local HG proxy BDB001. This suggests that DER individuals had received additional HG admixture upon their expansion into Europe, but not HG ancestry immediately related to the
local BDB001 source. When comparing all LBK groups, many were cladal with Anatolia_N, indicating a rather swift expansion with only occasional contribution of additional (and local) HG ancestry.

In turn, when comparing the HG sources to each other, we could not identify which source contributed the most to DER (supplementary fig. S2, Supplementary Material online). However, we did see an increased affinity of DER to the Loschbour HG representing the WHG ancestry, compared with Iron Gates HGs who have additional eastern HG ancestry (supplementary fig. S2, Supplementary Material online) (Mathieson et al. 2018).

We then used qpAdm to quantify the relative contribution of HG ancestry represented by the Loschbour HG and the Anatolia_N ancestry represented by Neolithic individuals from Barcin, western Anatolia (Mathieson et al. 2018). When modelled individually, we found that DER individuals had between 0% and 20% (SE 6.6%) European HG ancestry (fig. 2B, supplementary table S3, Supplementary Material online). When modeled as a group, we found that on average DER individuals had ~7% HG and ~93% western Anatolian Neolithic ancestries, indicating limited HG contribution (fig. 2C).

Using the same set of sources (Anatolia_N and Loschbour HG), we estimated the time of admixture between the Anatolian farmers and WHGs using DATES v. 753 (Moorjani et al. 2016) to have occurred 24.286 ± 2.784 generations ago (supplementary fig. S3, Supplementary Material online), or ~680 calendar years before the time of the DER individuals assuming a generation time of 28 years (Fenner 2005), which falls at the beginning of the Starčevo-Körös-Criș complex (~6,200–5,200 BCE) of the southeastern European Neolithic (Porčić et al. 2020).

According to both the model-free ADMIXTURE and the qpAdm results with a defined set of sources, we can show very limited HG ancestry influx in the DER LBK group in addition to the HG ancestry already carried by Anatolian farmers. This suggests minimal interaction between European HG and expanding early LBK farmers from southeastern Europe during the Neolithic, corroborating findings by other studies (Lipson et al. 2017; Mathieson et al. 2018; Rivollat et al. 2020).

Uniparentally Inherited Markers
Mitochondrial DNA and Y-chromosome haplogroups of the DER individuals are generally consistent with the genetic variation reported from early Neolithic farmers (supplementary table S1, Supplementary Material online) (Brandt et al. 2013; Szécsényi-Nagy et al. 2015), and the mitochondrial genomes match the HVR-I haplotypes reported earlier (Haak et al. 2010). We primarily found G2a2 and H2 Y haplogroups associated with Neolithic farmers (Lacan et al. 2011; Haak et al. 2015; Rivollat et al. 2020; Rohrlach et al. 2021). We also observed one male

Fig. 1. Population genetic analyses of LBK-associated individuals from Derenburg. (A) PCA of DER individuals and relevant published ancient groups, projected on West Eurasian genetic variation. (B) Unsupervised ADMIXTURE results ($k=12$) of a subset of individuals shown in panel A. (C) Map of Europe and zoom in on LBK sites in today’s Germany. Map tiles in panel (C) by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.
with the Y-chromosome I haplotype that is generally common among European HGs and therefore considered as a signal of male HG ancestry contribution but also reported in lower frequency from Iberian and French Neolithic individuals (Lipson et al. 2017; Rivollat et al. 2020).

**Biological Relatedness**

Based on the analysis of biological relatedness via READ (Monroy Kuhn et al. 2018) and lcMLkin (Lipatov et al. 2015), we identified three first-degree relationships (two-parent offspring, one sibling pair) and second-degree relationships.
restricted to (a) particular biological kin group(s). That the site was used over a longer period of time and not among a total of 32 genotyped DER individuals suggest a median relatedness of 0.12, and the subadults (P value = 0.35), none of which significantly differed from each other (supplementary fig. S4, Supplementary Material online).

We performed Wilcoxon Rank Sum tests on the mean measure of relatedness for each individual to the rest of the group as calculated from READ to see whether there is a general skew towards males or females from the site being more closely related to each other, which would indicate either potential patri- or matrilocality. We compared the pair-wise relatedness of all males to all females (P value = 0.34), all adult males to females (P value = 0.12), and the subadults (P value = 0.35), none of which significantly differed from each other (supplementary fig. S4, Supplementary Material online).

Considering that the DER burial site is archaeologically a closed site, we estimated the relatedness of the 2nd degree pairs, as several alternatives were possible.

We performed Wilcoxon Rank Sum tests on the mean measure of relatedness for each individual to the rest of the group as calculated from READ to see whether there is a general skew towards males or females from the site being more closely related to each other, which would indicate either potential patri- or matrilocality. We compared the pair-wise relatedness of all males to all females, all adult males to females, and the subadults, none of which significantly differed from each other (supplementary fig. S4, Supplementary Material online).

The biological relatedness at DER is similar to other LBK sites from Germany, for example, Halberstadt-Sonntagsfeld, where one first- and two first- or second-degree related pairs were observed of 24 total individuals (Lipson et al. 2017). Considering that the DER burial site is archaeologically a closed find and that the temporal spread of the calibrated years before present (calBP) is around 300 years, 6848.5–7141calBP (supplementary table S1, Supplementary Material online), the few first- and second-degree relationships among a total of 32 genotyped DER individuals suggest that the site was used over a longer period of time and not restricted to (a) particular biological kin group(s).

**LBK Inter-site Comparisons**

We further tested the observations on LBK population structure inferred from biological relatedness results using inbreeding coefficients and effective population size comparisons of LBK sites in Germany with other already published EEFs. We used READ (Monroy Kuhn et al. 2018) to test for relatedness between individuals from DER and other LBK sites in Germany, mainly already published data from Halberstadt-Sonntagsfeld (Germany_LBK_HBS, N = 29) and Stuttgart-Mühlhausen (Germany_LBK_SMH, N = 29) (Lipson et al. 2017; Rivollat et al. 2020). To do this, we first estimated the median relatedness in each German LBK group separately and found them to be comparable with each other (LBK_HBS = 0.24; LBK_SMH = 0.25; LBK_DER = 0.24). We then performed a joint biological relatedness analysis of all German LBK sites using READ and found no evidence of first- or second-degree relatedness between individuals from different sites.

To investigate this further, we estimated the effective population size (N_e) in LBK as implemented in (Fernandes et al. 2021). Using the maximum-likelihood to fit N_e, we estimated the effective population size to have been ~5,000 individuals (3,688–6,778 95% CI) for all LBK sites from Germany, which is consistent with effective population size estimates for a relatively large population (Tenesa et al. 2007; Melé et al. 2012). Together with the results from our analysis of the lack of inbreeding and inter-site biological relatedness, this further suggests that the early Neolithic LBK farmers were part of a larger population, did not practice close biologically related mating, and prevented inbreeding. In comparison, based on limited biological relatedness among individuals in intramural burials, it has been suggested that Neolithic Anatolia was not a strictly kin-based society (Pilloud and Larsen 2011; Yaka et al. 2021).

**Phenotypic Analysis**

We determined genotype likelihoods for a select number of SNPs associated with genetic adaptation, including lactase persistence, metabolic adaptation, among others, as well as SNPs associated with phenotypic traits as determined by the H-Irislex5 platform (Walsh et al. 2014; Chaitanya et al. 2018) (supplementary fig. S5, Supplementary Material online). Based on the analysis of the SNPs associated with phenotypes, we see the highest frequencies of brown (33%) eye color, intermediate skin color (21%), and brown (21%) hair color. We report these data being fully aware of the fact that the analysis of phenotypic variation can be a contentious issue, due to the fact that it was developed based on modern individuals of predominately European ancestry, and does not perform as well in admixed individuals and populations outside of Europe (Dembinski and Picard 2014, 2016; Yun et al. 2014). Thus, we note that these data, and specifically eye, skin, and hair color, are frequent requests of museums and other organizations to illustrate reconstruction life histories of past societies.

**HLA Class-I and II Allele Analysis**

HLA Class-I and II haplotypes in the DER individuals were analyzed using Optitype (Szolé et al. 2014). Based on the analysis, we report the highest allele frequencies for HLA-A to be A*24:02 (0.44) and A*02:01 (0.38), B*51:01 (0.31) and B*27:05 (0.25) for HLA-B, and C*02:02 (0.25) and C*15:02 (0.23) for HLA-C (supplementary tables S4 and S5, Supplementary Material online). In comparison, Late
Neolithic individuals from the Wartberg culture show similar distributions of the HLA class I alleles, with HLA-B*27:05 at a frequency of 23% and HLA-C*02:02 at 17% (Immel et al. 2019), which is higher than the modern frequencies of the two alleles in Europeans being 3.73% (Wu et al. 2021) and 5.6% (UK pop 3 from Gonzalez-Galarza et al. 2021), respectively. Interestingly, the two highest frequency HLA-B alleles in our sample, HLA-B*51:01 and HLA-B*27:05, have been associated with Behçet’s disease susceptibility (Verity et al. 1999; Kirino et al. 2013) and ankylosing spondylitis (AS) (Cauli et al. 2013; International Genetics of Ankylosing Spondylitis Consortium (IGAS) 2013; Chen et al. 2017). Based on the skeletal analysis of the individuals, the carriers of the HLA allele B*27:05 among the DER individuals did not show skeletal signs of AS, even though ~90% of the AS patients are HLA-B*27 (Brewerton et al. 1973; Schlosstein et al. 1973; Woodrow and Eastmond 1978). Although only 1–5% of the HLA-B*27 positive individuals develop AS in the general population (van der Linden et al. 1983; Reveille et al. 2010), it is recognized as one of the most important monogenic associations for this disease known to date (van der Linden et al. 1983; Reveille et al. 2010; Karnes et al. 2017). It is possible that some of the carriers in our sample could have still developed the condition later in life (Feldtkeller et al. 2003). The most frequent HLA class II alleles are HLA-DRB1*11:01 (0.42), -DRB3*02:02 (0.48), -DQA1*05:05 (0.44), and -DQB1*03:01 (0.52) (supplementary tables S4 and S5, Supplementary Material online). All these alleles show similar frequencies as they are all part of the same HLA class II haplotype (43.75%), which was reported in relatively high frequencies in present-day central European populations such as Czech Republic (15.30%) (Zajacova et al. 2016), but also in Nganasan from Siberia (18.80%) (Uinuk-Ool et al. 2004).

Selection Scans

The transition from hunting and gathering to agriculture and animal husbandry in the Neolithic has been linked to new selective pressures due to an increased pathogen burden (Armelagos et al. 1991; Key et al. 2020) and a shift to a cereal-based diet together with decreased nutritional diversity (Luca et al. 2010). We performed two types of selection scans (frequency- and haplotype-based) to detect signatures of selection associated with Neolithization in the newly generated DER cohort together with already published data. We first describe the settings and cut-offs used in each analysis and then address the intersection of the results in a joint discussion.

We first used a locus-specific branch length (LSBL) test for selection (Shriver et al. 2004). The LSBL test is based on the pair-wise genetic distances between three populations, wherein each SNP is considered independently. SNPs with the longest Fst branches in one population compared to the other two are considered as potential selection candidates. We performed three different LSBL tests to determine whether the transition from hunting and gathering to an agricultural lifestyle and the population expansion from the Near East was associated with adaptation by natural selection to new environments in higher latitudes of Europe and to the Neolithic way of life in general.

First, we compared three cohorts consisting of LBK (N = 97), Neolithic individuals from western Anatolia (N = 37), and autochthonous HGs as local control group (N = 86), (see supplementary fig. S6, Supplementary Material online for the PCA with the individuals used in our cohorts, and supplementary table S6, Supplementary Material online for the list of the individuals). We chose to compare the LBK to the Anatolian farmers because these populations are known to be closely related based on our PCA and Admixture analyses, as well as previous studies (Mathieson et al. 2015; Lazaridis et al. 2016). SNPs with the largest pair-wise Fst values between Anatolian farmers and the LBK, as well as between LBK and HGs were prioritized, since these SNPs could indicate selection for the new environmental or lifestyle conditions associated with the expansion into and across Europe, and ongoing Neolithization, respectively. Based on the LSBL analysis, we prioritized N = 9549 genetic loci under the empirical quantile < 0.01 (top 1%) that was calculated based on the distribution of the LSBL statistic following previous studies (Bigham et al. 2010).

Second, to test for selection since the deeper split of LBK and Anatolia_N from WHG, we merged the ancient farmer populations (LBK and Anatolia) into a single EF group, and compared it with WHG, and modern Esan (ESN) individuals from the 1KG data set (The 1000 Genomes Project Consortium 2015) as a non-European farming group comparison. In a third analogous LSBL analysis, we also compared EF to WHG and an African HG cohort represented by Biaka and Mbuti from the SGDJP and HGDP data sets, respectively (Mallick et al. 2016; Bergström et al. 2020). During this deeper time frame, we expect the Neolithization in Western Eurasia to have had a substantial effect due to the change in subsistence and increased population densities and disease susceptibility.

We also performed an ancestry-aware selection scan using our newly developed method AIMLESS. For each position on the genome for which we had allele frequencies for all three groups (LBK, Anatolia_N, and WHG), we performed a maximum-likelihood estimation (MLE) of allele frequencies in the LBK based on the observed data (unconstrained model), which we compared with the MLE based on the most likely admixture of Anatolia_N and WHG ancestries (constrained model). We performed a likelihood ratio test comparing two models and determined P values based on a χ² distribution with two degrees of freedom and considered loci with P value < 1e-8 as significant based on simulation studies (supplementary figs. S7 and S8, Supplementary Material online).

To look for more recent signals of adaptation (Szpiech and Hernandez 2014) to the Neolithic lifestyle through the comparison of LBK to WHG, we also explored a haplotype-based selection test called Cross Population Extended Haplotype Homozygosity (XP-EHH), which can detect selective sweeps in one population compared...
with another (Sabeti et al. 2007). To do so, we first used GLIMPSE (Rubinacci et al. 2021) to impute and phase the genome-wide data from DER and already published LBK individuals (Lazaridis et al. 2014; Mathieson et al. 2015; Lipson et al. 2017; Rivollat et al. 2020), as well as Iron Gates HGs (Mathieson et al. 2018). To understand the coverage cutoff for imputation, we performed a down-sampling experiment, wherein we down-sampled a high-coverage (92 × 1,240k SNP captured LBK individual from Stuttgart Mühlhausen, which was obtained as a result of routinely using this library as positive control in capture experiments. Based on our test of the genotype agreement between the down-sampled and diploid data, we decided on a cutoff of 0.5 × (LBK: N = 31, Iron Gates: N = 27) for the inclusion of the imputed samples in the downstream selection scans with selscan (Szpiech and Hernandez 2014) (supplementary fig. S9, Supplementary Material online). We identified 52,648 sites with the absolute XP-EHH value >2.

Owing to the limitations of missing data inherent in aDNA and to further overcome the potential issue of false positive results based on imputation or factors such as drift and population structure, we prioritized results that were overlapping between XP-EHH, LSBL, and AIMLESS, since it has been shown that a combination of multiple tests is more powerful to detect true selection signals (Grossman et al. 2010). An important caveat is that we used the 1 KG data as modern reference panel (The 1000 Genomes Project Consortium 2015) to impute the ancient data. This is not ideal since imputation with a modern reference can introduce spurious genotypes and make the individuals imputed more similar to the reference panel used. We are also aware that the European gene pool underwent further transformation during the Bronze Age (Allentoft et al. 2015; Haak et al. 2015). This is why the selection results based on imputed data have to be analyzed and interpreted with caution.

As predicted, we identified many SNPs associated with immune function and metabolism among the LSBL top 1% hits from the comparison of LBK versus Anatolia_N and WHG (supplementary table S7, Supplementary Material online). To determine if any biological pathways were over-represented among the top SNPs identified by the LSBL test of LBK versus Anatolia_N and WHG, we used Gowinda (Kofer and Schlotterer 2012) for a gene-set enrichment analysis. We identified 110 significant Gene Ontology (GO) terms after a correction for multiple comparisons using the false discovery rate cutoff of 0.05. Among the significant results, we found 8 of 110 GO terms associated with lipid metabolism (OR: 1.82 compared with all GO terms) (GO0061365: positive regulation of triglyceride lipase activity, GO0051006: positive regulation of lipoprotein lipase activity, GO0060193: positive regulation of lipase activity, GO0015910: long-chain fatty acid transport, GO0015908: fatty acid transport, among others) and 6 out of 110 with the immune system (OR: 0.65 compared with all GO terms) (GO0006911: phagocytosis, GO0002376: immune system process, GO0050900: leukocyte migration, GO0001776: leukocyte homeostasis, and others) (supplementary table S8, Supplementary Material online). We summarized and visualized the GO terms with REVIGO (Supek et al. 2011) using the default parameters (supplementary fig. S10, Supplementary Material online). Among the GO term “superclusters” we identified: “long-chain fatty acid transport”, “leukocyte homeostasis”, “positive regulation of lipoprotein lipase activity”, “immune system process”.

Overall, we found 227 overlapping positions between the three LSBL tests (supplementary tables S9–S11, Supplementary Material online). Notably, those included SNPs associated with APOM rs805297 (linked to rheumatoid arthritis and lipid signaling), BDNF rs11030104 (BMI), IGSF21 rs12031938 and ILDR2 rs10489574 (immune response), LMFI rs12933840 and rs8062983 (lipid metabolism), MHC SNPs rs2844482 and rs9277027 (immune function), and PDE2A rs11235559 (susceptibility to viral infections) (Fumagalli et al. 2010). Our findings from the LSBL comparisons highlight the immune and lipid metabolism pathways as strong selection candidates. This is not surprising, since immune function adaptation has been shown before in previous analyses of natural selection in Western Europe from the Neolithic to modern time (Mathieson et al. 2015; Chekalin et al. 2019). To our knowledge, this is the first time that the continuous selection on the immunity related pathways is shown with regards to the Neolithic expansion from the Near East to Europe.

Among the SNPs that were identified by the LSBL tests for deeper selection in the early Neolithic farmers, we find loci associated with skin color (SLC24A5, CD82) adaptation and folate synthesis (MTHFR, NBP3) (supplementary tables S9 and S10, Supplementary Material online). Notably, SLC24A5, has been also identified by previous studies of selection in Europeans using aDNA (Mathieson et al. 2015; Ju and Mathieson 2021). Skin color variation is one of the most striking features of human phenotypic diversity. Skin color is highly correlated with latitude and the resulting ultraviolet radiation (UVR) intensity, and populations closer to the equator have darker skin color, whereas the populations located away from the equator have lighter skin color (Jablonski and Chaplin 2000). It has been long hypothesized that skin color variation has evolved as a balance between vitamin D synthesis and folate synthesis, as well as skin cancer protection. On one hand, vitamin D is synthesized upon UVR exposure (Jablonski and Chaplin 2010), and, on the other hand, folate has to be protected from photolysis from UVR (He et al. 2009). Our findings support the idea of co-evolution of folate synthesis and lighter pigmentation in the EF, compared to the HGs and the African outgroup populations. Moreover, our study narrows down the timing of selection on SLC24A5 to after the split between the European HGs and the ancestors of the early Neolithic farmers. Of note, European HGs are characterized by a darker skin color, even though they occupied higher latitudes with low UVR exposure. The current hypothesis is that the diet based on meat/fish served as a sufficient source of vitamin D in higher

Downloaded from https://academic.oup.com/molbev/article/39/6/msac108/6586604 by MP evolutionary Anthropology user on 23 June 2022
latitudes for the HGs, similar with what has been shown in the Inuit and other Arctic indigenous populations (Kolahdooz et al. 2013; Schaebel et al. 2015).

We analyzed SNPs that are associated with pigmentation in the ancient populations separately (LBK, Anatolia_N, WHG) (table 1), since we hypothesized that migration to a higher latitude from Anatolia to central Europe could be associated with positive selection for light skin color phenotype. Based on the comparison between LBK and Anatolia_N, we did not see a major shift in skin color SNPs that are involved in determining lighter skin pigmentation, suggesting that this adaptation already happened in Anatolia/the Near East during the transition to a sedentary farming lifestyle.

The results from the XP-EHH test confirmed well-known selection signals linked to skin color adaptation in the gene SLC24A5 (Lamason et al. 2005) (supplementary fig. S11, Supplementary Material online). Selection signature around the SLC24A5 core SNP rs1426654 (xpehh = −2.05) (supplementary fig. S12, Supplementary Material online). The FADS1/FADS2 locus has been identified before by previous studies of selection using ancient individuals. It has been shown that FADS was under selection both before and after the Neolithic transition (Ye et al. 2017), with the ancestral allele under selection before (Mathieson 2020) and the derived after (Martiniano et al. 2017).

A novel gene of interest that was also identified by both LSBL and XP-EHH is LMF1 (fig. 4A). LMF1 or lipase maturation factor 1 is involved in the lipase system, and is required for the maturation of the lipoprotein lipase, hepatic lipase, and endothelial lipase [for review see (Péterfy 2012)]. LMF1 activity is associated with levels of triglycerides, and its mutations are linked to hypertriglyceridemia, or the presence of high levels of triglycerides.

### Table 1. SNPs Associated with Pigmentation and Other Phenotypes of Interest.

| Gene              | Phenotype            | SNP       | REF | ALT | EFF | LBK     | Anatolia | WHG | AFR | EAS | EUR |
|-------------------|----------------------|-----------|-----|-----|-----|---------|----------|-----|-----|-----|-----|
| FADS1/FADS2       | Fatty acid Metabolism| rs174546  | C   | T   | T   | 0.70    | 0.71     | 0.97| 0.02| 0.57| 0.35|
| TYR               | Light skin (WE)      | rs10831496| A   | G   | A   | 0.28    | 0.42     | 0.13| 0.17| 0.25| 0.68|
| TYR               | Light skin (WE)      | rs1042602 | A   | C   | A   | 0.17    | 0      | 0   | 0.01| 0.00| 0.37|
| OCA2              | Light skin (EE)      | rs1800414 | C   | T   | C   | 0       | 0      | 0   | 0.01| 0.60| 0   |
| OCA2              | Light skin (WE)      | rs1800404 | C   | T   | T   | 0.73    | 0.85     | 0.83| 0.13| 0.39| 0.79|
| OCA2              | Blue eyes (WE)       | rs12913832| G   | A   | G   | 0.42    | 0.30     | 0.55| 0.03| 0.00| 0.64|
| APBA2             | Light skin (WE)      | rs4424881 | C   | T   | C   | 0.90    | 0.95     | 0.71| 0.10| 0.45| 0.87|
| SLC24AS           | Light Skin (WE)      | rs1426654 | A   | G   | A   | 0.89    | 1       | 0.47| 0.07| 0.01| 1.00|
| SLC24AS           | Blue eyes (WE)       | rs2470102 | A   | G   | A   | 0.90    | 1       | 0.44| 0.07| 0.25| 0.99|
| MC1R              | Light skin (EE)      | rs2228479 | A   | G   | A   | 0.03    | 0       | 0.08| 0.00| 0.29| 0.07|
| MHC               | Immunity             | rs2269424 | A   | G   | A   | 0.73    | 0.89     | 0.68| 0.01| 0.14| 0.26|

Fig. 3. Manhattan plot of the results of the selection scans. Chromosomes are shown on the x axis, whereas the -log10(P values) from AIMLESS are represented on the y axis. Significant loci that are shared between AIMLESS, LSBL (LBK vs. Anatolia_N and HG), and XP-EHH are highlighted in green. Loci that do not overlap between the three tests are shown in gray. The loci related to immunity and metabolic pathways are highlighted in red.
and lipase deficiency (Péterfy et al. 2007). LMF1 is present at high levels in mice embryos, suggesting a potential, although currently unknown, role in embryonic development (Ehrhardt et al. 2014). Plant-based diets are characterized by a lower lipid content (Yokoyama et al. 2017), and efficient lipid metabolism could have been under selection especially in LBK individuals, who have been shown to have higher rates of caries than individuals in the Late Neolithic, indicative of the reliance on cereals (Nicklisch et al. 2016). Another novel selection candidate is JAK1 (fig. 4B). The JAK-STAT pathway is involved in the tuning of immune response, and is associated with immune disorders (Witalisz-Siepracka et al. 2018). Janus kinases are involved in immune cell signaling, and complete deletions of Jak1 and Jak2 are not compatible with life in mice (Ghoreschi et al. 2009). Polymorphisms in JAK1 have been linked to blood cell type composition (Astle et al. 2016), white blood cell count (Kichaev et al. 2019).

Lastly, we explored all 971 loci that are shared between the three selection tests AIMLESS, XP-EHH and LSDL (LBK vs. Anatolia N and HGs). These include SNPs associated with LEPR, IL13, HLA-A, HLA-B, HLA-DQB2, CHRAC1, BMP1, SORBS1, ELMO1, and others (supplementary table S12, fig. 3), among which SNPs associated with immune function and metabolism such as HLA-DQB1, SORBS1, and LEPR stood out (supplementary table S12, Supplementary Material online). The HLA-DQ molecule is a major histocompatibility complex (MHC) class II protein, expressed in B lymphocytes, dendritic cells, and macrophages. Diseases that are associated with the genes HLA-DQA1 and -DQB1 coding for HLA-DQ include Celiac Disease (Zubillaga et al. 2002; Megiorni and Pizzuti 2012), autoimmune diseases (Spurkland et al. 1991; Badenhoop et al. 1995), and viral resistance (Huang et al. 2019). Interestingly, HLA-DQB1 showed a strong signature of balancing selection in regions where Celiac Disease is common (Solberg et al. 2008; Sams and Hawks 2014), indicating its involvement in the response to pathogens, potentially offsetting the negative selective pressures from Celiac Disease (Hale 2017). SORBS1 is associated with obesity and the development of type 1 and 2 diabetes (Lin et al. 2001; Germain et al. 2015), as well as reduced Tuberculosis bacterial growth (Agrawal et al. 2016). LEPR encoding the leptin receptor is another gene of interest, due to its association with obesity-related outcomes in modern individuals (Clement et al. 1998; Farooqi et al. 2007), as well as its involvement in the immune system (De Rosa et al. 2015). Importantly, LEPR has been reported as a selection candidate using both ancient and modern data (Voight et al. 2006; Colbran et al. 2021).

Conclusions
On the basis of comparative analyses of genome-wide genetic variation at \( \sim 1.24M \) SNPs in EEF, we find similar ancestry proportions at Derenburg-Meerentie, as we see in other LBK sites across Germany and the broader LBK distribution zone (Hofmanová et al. 2016; Lipson et al. 2017; Rivollat et al. 2020). Overall, we find that LBK individuals closely resemble western Anatolian early Neolithic farmers, suggesting a rapid population expansion from western Anatolia to Central Europe. Moreover, the results from our LBK individuals suggest that there was no, or very limited, admixture between the local European HG and the incoming farmers during the pioneer phase of the early Neolithic in Central Europe. We show that all EF belonged to a relatively large source population compared to the
local HGs, and found no signal of recent inbreeding, which means that larger population sizes were initially maintained. The relative genetic similarities between EFs allowed us to group them together, leveraging the increased cohort size for selection scans. Despite there being a limited difference in time between the LBK and Anatolian Neolithic individuals, and potential ongoing gene flow from Anatolia and southeastern Europe, the overlapping results of the three selection scans we explored and of the HLA class-I and II analysis suggest that Early Neolithic individuals in central Europe were experiencing an ongoing metabolic and immune function adaptation to the Neolithic lifeways as a result of the increased disease burden in group sizes that exceeded those of HGs. Moreover, we see an adaptation to higher latitudes in Neolithic individuals compared with modern Africans and HGs, as evidenced by allele frequency changes of variants associated with pigmentation and folate synthesis. The Neolithization has been considered one of the strongest natural selective forces acting on the human genome. Our results show that there are multiple signals of selection that can be observed in the LBK and the EFs in general in association with the transition to an agricultural lifestyle and sedentism. However, our findings also highlight the challenges of studying selection using aDNA, especially in light of complex admixture or demographic histories and limited time that elapsed since the split of populations under study. We also stress the need for increased cohort sizes and refinement of statistical tools that can take these confounding factors into account.

Sample Preparation
Samples were processed using the following aDNA extraction protocol (Velsko et al. 2020), with the UDG-half treatment. Samples (N = 37) were first shotgun sequenced followed by the 1,240k capture on the samples (N = 32) with more than 0.1% endogenous DNA for select SNPs (supplementary table S1, Supplementary Material online). The raw fastq files were processed and mapped to the human reference hs37d5 with default parameters from EAGER (Peltzer et al. 2016). Minimum mapping and base qualities of 30 were used.

Contamination and aDNA Authentication
aDNA status was authenticated using MapDamage2.0 (Jonsson et al. 2013). Two base pairs were trimmed from the UDGHalf bam files following the aDNA damage pattern using trimBam as implemented in BamUtil 1.013 (Jun et al. 2015). X-chromosome contamination estimation in males was performed using ANGSD (Korneliussen et al. 2014) following the software manual. First step: “angsd -r X:5000000-154900000 -doCounts 1 -iCounts 1 -minMapQ 30 -minQ 30”, second step: “contamination -a angsdCounts -h HapMapChrX.gz 2 > out”. Contamination estimates from samples with more than ~200 SNPs on the X chromosome covered were considered as reliable (supplementary table S1, Supplementary Material online). Mitochondrial DNA contamination was estimated using Schmutzi (Renaud et al. 2015) and contamMix (Fu et al. 2013).

Genotyping
Pseudohaploid genotyping was performed using SAMTools mpileup (Li and Durbin 2009) and PileupCaller (https://github.com/stenschiff/sequenceTools/blob/master/src/SequenceTools/PileupCaller.hs) for the 1,240k sites. Pseudohaploid genotypes were merged with the Human Origin Affymetrix dataset of modern individuals around the globe (Patterson et al. 2012), as well as published ancient West Eurasian individuals (https://reich.hms.harvard.edu/ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data, version 44.3). SmartPCA from EIGENSOFT v7.2.1 (http://www.hsph.harvard.edu/alkes-price/software/) was used to perform a PCA projection of ancient samples on top of the variation of the modern data. Unsupervised admixture was performed on k components from 1 to 17 with random seeding after LD pruning in plink with the following settings (-indep-pair-wise 200 25 0.4). Cross-validation (CV) values were compared, and the k component 11 with the lowest CV error was selected.

F-statistics
R package admixr was used to perform f-statistics and qpAdm (Petr et al. 2019). We performed f4 tests to determine if there were differences in HG ancestry in the individuals from DER compared with other EEF groups, as well as to test whether LBK individuals had additional HG ancestry
compared to individuals from Anatolia Neolithic (Bar8 and Bar31 (Hofmanová et al. 2016)). Based on the admixture analysis, we formally tested the admixture components using apaAdm as implemented in the admixr package (Petr et al. 2019) using a 12 sample outgroup (Mbunt, Papuan, Onge, Han, Karitiana, Mota, Ust-Ishim, Russia_Mal’ta1, Vestonice, CHG, Israel_Natufian, Villabruna) + additional European HG sources (Villabruna, Russia_EHG, GoyetQ2). The admixture event between WHG and Anatolian farmers was determined using DATES v. 753 (Moorjani et al. 2016).

**Biological Relatedness and ROH Analyses**

READ (Monroy Kuhn et al. 2018) and icMLkin (Lipatov et al. 2015) were used to perform tests of biological relatedness. ROH analysis was performed using HapROH on samples with more than 600k SNPs based on the 1,240k capture (Ringbauer et al. 2020).

**Mitochondrial and Y Haplotype Analysis**

To determine mitochondrial DNA haplogroups, bam files were mapped to the mtDNA rCRS sequence (Andrews et al. 1999). Haplogrep 2.0 was used to assign mitochondrial haplogroups. Y-haplotype analysis was performed with alleles of different loci.

**Shotgun Data**

Three individuals (DER002, DER009, BDB001) were shotgun sequenced (PE75) to high coverage.

**HLA Typing**

A development version of OptiType 2 (Szolek et al. 2014) was used to determine HLA Class I and II alleles from ≥30 bp single-end sequencing data retrieved from the immunocapture libraries (https://github.com/FRED-2/ OptiType, tag DER). Fifteen of 432 allele calls were overruled in favor of runner-up alleles based on anomalous coverage patterns induced by short reads cross-mapping with alleles of different loci.

**Selection Scans**

LSBL was used on the 1,240k genotype data to determine sites in the genome that have undergone recent selective sweeps. This method relies on pair-wise Fst values for three populations to identify loci with extreme Fst difference between population of interest and two other groups used for comparison. Resulting LSLB values were ranked and empirical P values were determined. Resulting SNPs were annotated using the R package BioMart (Durinck et al. 2009).

We also performed a pathway enrichment analysis using gowinda (Kofer and Schlotterer 2012) comparing the top 1% LSLB hits to all loci analyzed (1,240k SNPs after QC). This method performs gene-set enrichment base on GO classifications and designed specifically for GWA studies. Gowinda was designed to take into account the assumptions that are often violated in genome-wide associations studies, such as longer genes having more SNPs, and overlapping genes being sampled in clusters (Kofer and Schlotterer 2012). Gowinda performs internal permutations by randomly sampling SNPs from the total set of SNPs. By recording the associated genes and performing the permutation multiple times, gowinda establishes an empirical null distribution, and the significant overrepresentation of candidate SNPs is determined compared with the empirical null distribution (Kofer and Schlotterer 2012).

**AIMLESS Admixture-Informed Maximum-Likelihood Estimation for Selection Scans**

For site i on the genome, we calculate the allele frequency for two source populations and a target population, denoted $f^i_1, f^i_2,$ and $f^i$ respectively. We then calculate the optimal value of the mixing parameter $\gamma \in [0, 1]$ such that we minimize the value of the function

$$f(\gamma) = \left(f^i_1 - \gamma f^i_2 + (1 - \gamma) f^i\right)^2,$$

and define the constrained allele frequency estimator to be $\hat{f}^i_{\gamma} = \gamma \hat{f}^i_1 + (1 - \gamma) \hat{f}^i_2$. Note that if $\hat{f}^i_2 = \hat{f}^i$, then we simply let $\gamma = 1$.

We then calculate log-likelihood values for the unconstrained model and the constrained model assuming that $X_i \sim B(N_i, \hat{f}^i_{\gamma})$ for the constrained model, and that $X_i \sim B(N_i, \hat{f}^i)$ for the unconstrained model, denoted $l_1$ and $l_0$. Naturally, as $f$ is the maximum-likelihood estimator for a binomial observation, $l_0 \geq l_1$.

Finally, we perform a LRT, with test statistic $L_1 = -2(l_1 - l_0)$, where $A_1 \sim X^2$. Via a simulation study based on the observed allele frequencies in Anatolia_N and WHG populations, we identified a P value cutoff of $1 \times 10^{-4}$ that returned a false discovery rate of zero, with the highest positive discovery rate.

Note that we filter sites for which we have observed allele counts from at least 10 individuals from each population, and for which the allele frequency in the source populations is $>0$.

**Haplotype-Based Scan**

Genome-wide 1,240k capture data from Derenburg, LBK and Iron Gates HG individuals were imputed using GLIMPSE following the standard procedures (Rubinacci et al. 2021). Genotype likelihoods were computed using bcftools mpileup (Li 2011). The GLIMPSE_phase function was used to run imputation in chunks with window size of 2,000,000 bps and buffer size of 200,000 bps. The chunks were then ligated using GLIMPSE_ligate and haplotypes were sampled using GLIMPSE_sample. The 1000G reference panel was used for imputation (The 1000 Genomes Project Consortium 2015). Samples with endogenous DNA coverage on the 1,240k capture of above 0.5× were included in the analysis. REHH package was used to run haplotype-based selection scans (Gautier and Vitalis 2012).
Supplementary Material

Supplementary data are available at Molecular Biology and Evolution online.

Acknowledgments

We thank the members of the Archaeogenetics Department of the Max Planck Institute for the Science of Human History, especially Ellen Späth for sampling, the Population Genetics group for helpful feedback, and Guido Brandt, Antje Wissgott and Stephen Clayton for sequencing and handling of the raw data. We are indebted to State Office for Heritage Management and Archaeology Saxony-Anhalt—State Museum of Prehistory, Germany, and all archaeologists and students involved in the excavations at Denenberg-Meerestieg II. We would also like to thank Iain Mathieson, Gaspard Kerner, and Vanessa Villalba-Mouco for helpful discussions during the preparation of the manuscript. The study was funded by the Max Planck Society, and the French (ANR) and German (DFG) Research Foundations, INTERACT project, ANR-17-FRAL-0010, DFG-HA-5407/4-1, 2018-2021 (to M.-F.D., W.H.), and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program under grant agreement no. 771234 – PALEoRIDER (to W.H.).

Author Contributions

Conceptualization, investigation, data curation, and writing – original draft preparation: W.H. and A.C.; data generation: F.A. and R.B.; formal analysis: A.C. and A.B.R.; funding acquisition: W.H. and M.-F.D.; writing – review & editing: A.C., W.H., and A.B.R.

Data Availability

Genomic data (BAM and fastq format) are available at the European Nucleotide Archive under accession number PRJEB52488, genotypes in eigenstrat format can be found at https://edmond.mpdl.mpg.de/dataset.xhtml?persistentId=doi:10.17617/3.HOKISI. The code for AIMLESS is available at https://edmond.mpdl.mpg.de/dataset.xhtml?persistentId=doi:10.17617/3.HOKISI.

References

The Australo-Anglo-American Spondyloarthritis Consortium (TASC). 2010. Genome-wide association study of ankylosing spondylitis identifies non-mhc susceptibility loci. Nat Genet. 42(2):123–127.

International Genetics of Ankylosing Spondylitis Consortium (IGAS). 2013. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet. 45(7):730–738.

The 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. Nature 526(7571):68–74.

Agrawal N, Bhattacharyya C, Mukherjee A, Ullah U, Pandit B, Rao KVS, Majumder PP. 2016. Dissecting host factors that regulate the early stages of tuberculosis infection. Tuberculosis 100: 102–113.

Allentoft ME, Sikora M, Sjögren KG, Rasmussen S, Rasmussen M, Stenderup J, Damgaard PB, Schroeder H, Ahlstrom T, Vinner L, et al. 2015. Population genomics of bronze age eurasia. Nature 522(7555):167–172.

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet. 23(2):147.

Armelagos GJ, Goodman AH, Jacobs KH. 1991. The origins of agriculture: population growth during a period of declining health. Popul Environ. 13(1):9–22.

Ash A, Francken M, Pap I, Tvrz ́y Z, Wahl J, Pinhasi R. 2016. Regional differences in health, diet and weaning patterns amongst the first neolithic farmers of central Europe. Sci Rep. 6:29458.

Astle WJ, Elding H, Jiang T, Allen D, Ruklisa D, Mann AL, Mead D, Bouman H, Riveros-Mckay F, Kostadima MA, et al. 2016. The allelic landscape of human blood cell trait variation and links to common complex disease. Cell 167(5):1415–1429.e19.

Badenhoop K, Wallfish PG, Rau H, Fischer S, Nikolay A, Bogner U, Schleusener H, Usadel KH. 1995. Susceptibility and resistance alleles of human-leukocyte antigen (hla) dqalpha1 and hla dqbeta1 are shared in endocrine autoimmune-disease. J Clin Endocrin Metab. 80(7):2112–2117.

Bánffy E. 2019. First farmers of the carpathian basin: changing patterns in subsistence, ritual and monumental figurines. Oxbow Books.

Barrett R, Kuzawa CW, McDade T, Armelagos GJ. 1998. Emerging and re-emerging infectious diseases: the third epidemiologic transition. Annu Rev Anthropol. 27(1):247–271.

Bergström A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecpe P, Chen Y, Felkel S, Hallast P, Kamm J, 2020. Insights into human genetic variation and population history from 929 diverse genomes. Science 367(6484):eaay5012.

Bickpe B. 2009. Creating communities: new advances in central European neolithic research. Oxbow Books.

Bigham A, Bauchet M, Pinto D, Mao X, Akey JM, Scherer SW, Julian CG, Wilson MJ, Lopez Herraez D, et al. 2010. Identifying signatures of natural selection in tibetan and andean populations using dense genome scan data. PloS Genet. 6(9):e1001116.

Bocquet-Appel J-P, Bar-Yousef O. 2008. The neolithic demographic transition and its consequences: Springer Science & Business Media.

Bogucki P. 1999. The origins of human society. Blackwell Publishers Oxford.

Braud S, Diekmann Y, Booth TJ, van Dorp L, Faltskova Z, Rohland N, Mallick S, Oladele I, Ferry M, Michel M, et al. 2019. Ancient genomes indicate population replacement in early neolithic britain. Nat Ecol Evol. 3(5):765–771.

Brandt G, Haak W, Adler CJ, Roth C, Szecsenyi-Nagy A, Karimnia S, Seebauer P, et al. 2013. Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. Science 342(6155): 257–261.

Brewerton DA, Hart FD, Nicholls A, Caffrey M, James DC, Sturrock RD. 1973. Ankylosing spondylitis and hl-a 27. Lancet 1(7809): 904–907.

Brunel S, Bennett EA, Cardin L, Garraud D, Barrand Emam H, Beylier A, Boulestin B, Chenal F, Ciesielski E, Convertini F, et al. 2020. Ancient genomes from present-day france unveil 7,000 years of ancient European mitochondrial diversity. Science 367(6484):eaay5012.

Cauli A, Shaw J, Giles J, Hatano H, Rysnik O, Payeli S, McHugh K, Dessole G, Porru G, Desogus E, et al. 2013. The arthritis-associated
descended from neolithic aegaeans. Proc Natl Acad Sci U S A. 113(25):6886–6891.

Horejs B, Milić B, Ostmann F, Thanheiser U, Weninger B, Galik A. 2015. The aegae in the early 7th millennium bc: maritime networks and colonization. J World Prehistory 28(4):289–330.

Huang JY, Xu R, Wang M, Liao Q, Huang K, Shan ZG, You QZ, Li CY, Rong X, Fu YS. 2019. Association of hla-dqb1*0301 and drb1*11:01 with spontaneous clearance of hepatitis C virus in Chinese Li ethnicity, an ethnic group genetically distinct from Chinese Han ethnicity and infected with unique HCV subtype. J Med Virol 91(10):1830–1836.

Immel A, Key FM, Szolek A, Barquera R, Robinson MK, Harrison GF, Palmer WH, Spyrou MA, Susat J, Krause-Kyora B, et al. 2021. Analysis of genomic DNA from medieval plague victims suggests long-term effect of yersinia pestis on human immunity genes. Mol Biol Evol. 38(10):4059–4076.

Jabłoński NG, Chaplin G. 2000. The evolution of human skin coloration. J Hum Evol. 39(1):57–106.

Jabłoński NG, Chaplin G. 2010. Colloquium paper: human skin pigmentation as an adaptation to UV radiation. Proc Natl Acad Sci U S A. 107(Suppl 2):8962–8968.

Johnson AW, Earle TK. 2000. The evolution of human societies: from foraging to agrarian state. Stanford University Press.

Jonsson H, Gíngílsón A, Schubert M, Johnson PL, Orlando L. 2013. Mapdamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29(13):1682–1684.

Ju D, Mathieson I. 2021. The evolution of skin pigmentation-associated variation in west Eurasia. Proc Natl Acad Sci U S A. 118(11):e2009227118.

Jun G, Wing MK, Abecasis GR, Kang HM. 2015. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. Genome Res. 25(6):918–925.

Karnes JH, Bastarache L, Shaffer CM, Gaudieri S, Xu Y, Glazer AM, Mosley JD, Zhao S, Raychaudhuri S, Mallal S, et al. 2017. Phenome-wide scanning identifies multiple diseases and disease severity phenotypes associated with hla variants. Sci Transl Med. 9(389):eaai8708.

Key FM, Posth C, Esquivel-Gómez LR, Hübler R, Spyrou MA, Neumann GU, Furtwängler A, Sabin S, Burri M, Wissgott A, et al. 2020. Emergence of human-adapted salmonella enterica is linked to the neolithization process. Nat Ecol Evol. 4(3):324–333.

Kichaev G, Bhattach G, Loh PR, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL. 2019. Leveraging polygenic functional enrichment to improve gwas power. Am J Hum Genet. 104(1):65–75.

Kirino Y, Bertsias G, Ishigatsubo Y, Mizuki N, Tugal-Tutkun I, Seyahi E, Ozyazgan Y, Sachi FS, Erber R, Inoko H, et al. 2013. Genome-wide association analysis identifies new susceptibility loci for behcet’s disease and epistasis between hla-b*51 and erap1. Nat Genet. 45(2):202–207.

Kofler R, Schlotterer C. 2012. Gowinda: unbiased analysis of gene set enrichment for genome-wide association studies. Bioinformatics 28(15):2084–2085.

Kolahdooz F, Barr A, Roache C, Sheehy T, Corriveau A, Sharma S. 2013. Dietary adequacy of vitamin d and calcium among inuit and inuvialuit women of child-bearing age in arctic canada: a growing concern. PLoS One 8(11):e78987.

Kornellussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinformatics 15:356.

Lacan M, Keyser C, Ricaut FX, Brucato N, Duranthon F, Guillaume J, Crubezy E, Ludes B. 2011. Ancient DNA reveals male diffusion through the neolithic mediterranean route. Proc Natl Acad Sci U S A. 108(24):9788–9791.

Lamason RL, Mohideen MA, Mest JR, Wong AC, Norton HL, Aros MC, Juryncz MJ, Mao X, Humphreville VR, Humblet JE, et al. 2005. Slc24a5, a putative cation exchanger, affects pigmentation in zebrafish and humans. Science 310(5755):1782–1786.

Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, Fernandes D, Novak M, Gamaara B, Sirak K, et al. 2016. Genomic insights into the origin of farming in the ancient Near East. Nature 536(7617):419–424.

Lazaridis I, Patterson N, Mitternik A, Renaud G, Mallick S, Kirasnov K, Sudmant PH, Schraiger JG, Castellano S, Lipson M, et al. 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 513(7518):409–413.

Li H. 2011. A statistical framework for snp calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. Bioinformatics 27(21):2987–2993.

Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-wheeler transform. Bioinformatics 25(14):1754–1760.

Lin WH, Chiu KC, Chang HM, Lee KC, Tai TY, Chuang LM. 2001. Molecular scanning of the human sorbin and sh3-domain-containing-1 (sorbs1) gene: positive association of the t228a polymorphism with obesity and type 2 diabetes. Human Mol Genet. 10(17):1753–1760.

Lipatov M, Sanjeev K, Pato R, Veeramah KR. 2015. Maximum likelihood estimation of biological relatedness from low coverage sequencing data. bioRxiv 023374.

Lipatov M, Szczesniny-Nagy A, Mallick S, Posa A, Stegmar B, Keel V, Rohland N, Swara, M, McVerry B, Ferry M, Michel M, et al. 2017. Parallel palaeogenomic transects reveal complex genetic history of early European farmers. Nature 551(7680):368–372.

Lubbock J. 1866. Address delivered to the section of “primaval antiques” at the London Meeting of the Archaeological Institute, July, 1866. Archaeol J. 23(1):190–211.

Luca F, Perry GH, Di Rienzo A. 2010. Evolutionary adaptations to dietary changes. Annu Rev Nutr. 30:291–314.

Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A, et al. 2016. The simons genome diversity project: 300 genomes from 142 diverse populations. Nature 538(7624):201–206.

Maricic T, Whitten M, Pääbo S. 2010. Multiplexed DNA sequence capture of mitochondrial genomes using pcr products. PLoS One 5(11):e14004.

Martiniano R, Cassidy LM, O’Maoiliúin R, McLaughlin R, Silva NM, Manco L, Fidalgo D, Pereira T, Coelho MJ, Serra M, et al. 2017. The population genomics of archaeological transition in west iberia: investigation of ancient substructure using imputation and haplotype-based methods. PLoS Genet. 13(7):e1006852.

Mathieson I. 2020. Limited evidence for selection at the fads locus in early European farmers. Mol Biol Evol. 37(7):2029–2033.

Mathieson I, Alpaslan-Roodenberg S, Posth C, Szczesniny-Nagy A, Rohland N, Mallick S, Olalde I, Broomandkhoshbacht N, Candilio F, Cheronet O, et al. 2018. The genomic history of southeastern europe. Nature 555(7695):197–203.

Mathieson I, Lazaridis I, Rohland N, Mallick S, Patterson N, Roodenberg SA, Harney E, Stewardson K, Fernandes D, Novak M, et al. 2015. Genome-wide patterns of selection in 230 Ancient Eurasians. Nature 528(7583):499–503.

Megiorini F, Pizzuti A. 2012. Hla-dqa1 and hla-dqb1 in celiac disease predisposition: practical implications of the hla molecular typing. J Biomed Sci. 19:88.

Mélo M, Javed A, Pybus M, Zalloua P, Haber M, Comas D, Netea MG, Balanovský O, Balanovská E, Jin L, et al. 2012. Recombination gives a new insight in the effective population size and the history of the old world human populations. Mol Biol Evol 29(1):25–30.

Mendoza-Revilla J, Chacón-Duque JC, Fuentes-Guajardo M, Ormond L, Wang K, Hurtado M, Villes V, Granja V, Acuña-Alonzo V, Jaramillo C et al. 2021. Disentangling signatures of selection before and after European colonization in Latin American. bioRxiv. 2021.2011.2015.467418.
Monroy Kuhn JM, Jakobsson M, Günther T. 2018. Estimating genetic kin relationships in prehistoric populations. PLoS One 13(4): e0195491.

Moorjani P, Sankararaman S,Fu Q, Przeworski M, Patterson N, Reich D. 2016. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. Proc Natl Acad Sci U S A. 113(20):5652–5657.

Münster A, Knipper C, Oelze VM, Nicklisch N, Stecher M, Schlenker B, Gansmeier R, Fragata M, Friederich S, Dresely V, et al. 2018. 4000 years of human dietary evolution in central Germany, from the first farmers to the first elites. PLoS One 13(3):e0194862.

Nicklisch N. 2017. Spuren suche am skelett: Paläodemografische und epidemiologische untersuchungen an neolithischen und frühbronzezeitlichen bestattungen aus dem mittelelbe-saale-gebiet im kontext populationsdynamischer prozesse: kulturwandel. Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt, Landesmuseum für ....

Nicklisch N, Gansmeier R, Siebert A, Friederich S, Meller H, Alt KW. 2016. Holes in teeth – dental caries in neolithic and early bronze age populations in Central Germany. Ann Anat. 203:90–99.

Oelze VM, Siebert A, Nicklisch N, Meller H, Dresely V, Alt KW. 2011. Early Neolithic diet and animal husbandry: stable isotope evidence from three Linearbandkeramik (LBK) sites in Central Germany. J Archaeol Sci. 38(2):270–279.

Olalde I, Schroeder H, Sandoval-Velasco M, Vinner L, Löbon I, Ramirez O, Civit S, Garcia Borja P, Salazar-Garcia DC, Talamo S, et al. 2015. Uncertainty in genetic origin for early farmers from mediterranean cardinal and central european lbk cultures. Mol Biol Evol. 32(12):3312–3314.

Patterson N, Moorjani S, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. Genetics 192(3):1065–1093.

Peltzer A, Jäger G, Herbig A, Seitz A, Kniep C, Krause J, Niselstiel K. 2016. Eager: efficient ancient genome reconstruction. Genome Biol. 17:60.

Pétery M. 2012. Lipase maturation factor 1: a lipase chaperone involved in lipid metabolism. Biochim Biophys Acta 1821(5): 790–794.

Pétery M, Ben-Zeev O, Mao HZ, Weissglas-Volkov D, Aouizerat BE, Rohrlach AB, Papac L, Childebayeva et al. · Ringbauer H, Novembre J, Steinrücken M. 2016. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. Proc Natl Acad Sci U S A. 113(20):5652–5657.

Pétery M, Ben-Zeev O, Weissglas-Volkov D, Aouizerat BE, Rohrlach AB, Papac L, Childebayeva et al. 2020. Ancient genome-wide DNA from france highlights the complexity of interactions between mesolithic hunter-gatherers and neolithic farmers. Sci Adv. 6(22):eaaz5344.

Rohrlach AB, Papac L, Childebayeva et al. · Rivollat M, Villalba-Mouco V, Neumann GU, Penske S, Skourtanioti E, van der Loosdrecht M, Akar M, et al. 2021. Using y-chromosome capture enrichment to resolve haplogroup h2 shows new evidence for a two-path Neolithic expansion to Western Europe. Sci Rep. 11(1):15005.

Rubinacci S, Ribeiro DM, Hofmeister RJ, Delanoue O. 2021. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. Nat Genet. 53(1):120–126.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie X, Byrne EH, McCarroll SA, Gaudet R, et al. 2007. Genome-wide detection and characterization of positive selection in human populations. Nature 449(7164):913–918.

Sams A, Hawks J. 2014. Celiac disease as a model for the evolution of multifactorial diseases in humans. Human Biol. 86(1):19–35.

Schaebel LK, Bonefeld-Jorgensen EC, Lauberg P, Vestergaard H, Andersen S. 2015. Vitamin d-rich marine inuit diet and markers of inflammation – a population-based survey in greenland. J Nutr Sci. 4(e40).

Schröder H, Sandoval-Velasco M, Vinner L, Löbon I, Ramírez O, Civit S, García Borja P, Salazar-García DC, Talamo S, et al. 2015. Uncertainty in genetic origin for early farmers from Mediterranean Cardinal and Central European LBK Cultures. Mol Biol Evol. 32(12):3312–3314.

Patterson N, Moorjani S, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. Genetics 192(3):1065–1093.

Peltzer A, Jäger G, Herbig A, Seitz A, Kniep C, Krause J, Niselstiel K. 2016. Eager: efficient ancient genome reconstruction. Genome Biol. 17:60.

Pétery M. 2012. Lipase maturation factor 1: a lipase chaperone involved in lipid metabolism. Biochim Biophys Acta 1821(5): 790–794.

Pétery M, Ben-Zeev O, Mao HZ, Weissglas-Volkov D, Aouizerat BE, Pullinger CR, Frost PH, Kane JP, Malloy MJ, Reue K, et al. 2007. Mutations in LMF1 cause combined lipase deficiency and severe hypertriglyceridemia. Nat Genet. 39(12):1483–1487.

Petr M, Vernot B, Kelso J. 2019. Admixr-r package for reproducible analyses using admixtools. Bioinformatics 35(17):3194–3195.

Petrach J. 2020. Settlements, migration and the break of tradition: the settlement patterns of the earliest Bandkeramik and the LBK and the formation of a neolithic lifestyle in Western Europe. Quaternary Int. 560:524–528.

Pilloud MA, Larsen CS. 2011. "Official" and "practical" kin: inferring social and community structure from dental phenotype at Neolithic Çatalhöyük, Turkey. Am J Phys Anthropol. 145(5):519–530.

Porčić M, Blagojević T, Pendić J, Stefanović S. 2020. The timing and tempo of the neolithic expansion across the central balkans in the light of the new radiocarbon evidence. J Archaeol Sci Rep. 33:102528.

Renaud G, Slon V, Duggan AT, Kelso J. 2015. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. 16:224.

Reveille JD, Sims AM, Danoy P, Evans DM, Leo P, Pointron JJ, Jin R, Zhou X, Brud LA, Appleton LH, et al. 2010. Genome-wide association study of ankylosing spondylitis identifies non-hmC susceptibility loci. Nat Genet. 42(2):123–127.

Ringbauer H, Novembre J, Steinrücken M. 2020. Human parental relatedness through time – detecting runs of homozygosity in ancient DNA. bioRxiv 2020.2005.2031.126912.

Ringbauer H, Novembre J, Steinrücken M. 2021. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. Nat Commun. 12(1):5425.

Rivollat M, Jeong C, Schiffls S, Küçükkalipci I, Pemonge MH, Rohrlach AB, Alt KW, Binder D, Friederich S, Ghesquière E, et al. 2020. Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. Sci Adv. 6(22):eaaaz5344.
Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A map of recent positive selection in the human genome. *PLoS Biol*. 4(3):e72.

Walker PL, Bathurst RR, Richman R, Gjerdrum T, Andrushko VA. 2009. The causes of porotic hyperostosis and cribra orbitalia: a reappraisal of the iron-deficiency-anemia hypothesis. *Am J Phys Anthropol.* 139(2):109–125.

Walsh S, Chaitanya L, Clarisse L, Wirken L, Draus-Barini J, Kovatsi L, Maeda H, Ishikawa T, de Knijff P, et al. 2014. Developmental validation of the hirisplex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. *Forensic Sci Int Genet.* 9:150–161.

Woodrow JC, Eastmond CJ. 1978. Hla b27 and the genetics of ankylosing spondylitis. *Ann Rheum Dis.* 37(6):504–509.

Wu X, Wang G, Zhang L, Xu H. 2021. Genetics of ankylosing spondylitis-focusing on the ethnic difference between east asia and europe. *Front Genet.* 12:671682.

Yaka R, Mapelli I, Kaptan D, Doğu A, Chyleński M, Erdal OD, Köptekin D, Vural KB, Bayliss A, Mazzucato C, et al. 2021. Variable kinship patterns in neolithic anatolia revealed by ancient genomes. *Curr Biol.* 31(11):2455–2468.e118.

Ye K, Gao F, Wang D, Bar-Yosef O, Keinan A. 2017. Dietary adaptation of fads genes in europe varied across time and geography. *Nat Ecol Evol.* 1:167.

Yokoyama Y, Levin SM, Barnard ND. 2017. Association between plant-based diets and plasma lipids: a systematic review and meta-analysis. *Nutr Rev.* 75(9):683–698.

Zeder MA. 2011. The origins of agriculture in the near east. *Curr Anthropol.* 52(S4):S221–S235.

Zvelebil M. 2001. The agricultural transition and the origins of neolithic society in Europe. *Documenta Praehistorica* 28:1–26.

Zvelebil M. 2004. The many origins of the lbk. In: LBK Dialogues. *Studies in the formation of the Linear Pottery Culture*. Oxford: Archaeopress. p. 183–205.