Socio-cultural practices may have affected sex differences in stature in Early Neolithic Europe

Samantha L Cox[1,2*], Nicole Nicklisch[3], Michael Francken[4], Joachim Wahl[5], Harald Meller[6], Wolfgang Haak[7], Kurt W Alt[3], Eva Rosenstock[8], and Iain Mathieson[1*]

[1]*Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA*
[2]*Physical Anthropology Section, Penn Museum, University of Pennsylvania, Philadelphia, PA, USA*
[3]*Center of Natural and Cultural Human History, Danube Private University, Krems-Stein, Austria*
[4]*State Office for Cultural Heritage Management Baden-Württemberg, Osteology, Konstanz, Germany*
[5]*Institut für Naturwissenschaftliche Archäologie, AG Paläoanthropologie, Tübingen, Germany*
[6]*State Office for Heritage Management and Archaeology Saxony-Anhalt, State Museum of Prehistory, Halle, Germany*
[7]*Department of Archaeogenetics, Max Planck Institute for the Science of Human History, Jena, Germany*
[8]*Bonn Center for ArchaeoSciences, Universität Bonn, Bonn, Germany*
[*]*Correspondence to SLC (coxsl@sas.upenn.edu) or IM (mathi@pennmedicine.upenn.edu)*

Abstract

The rules and structure of human culture impact health as much as genetics or environment. To study these relationships, we combine ancient DNA (n=230), skeletal metrics (n=391), paleopathology (n=606), and dietary stable isotopes (n=873) to analyze stature variation in Early Neolithic Europeans from North Central, South Central, Balkan, and Mediterranean regions. In North Central Europe, stable isotopes and linear enamel hypoplasias (LEH) indicate high environmental stress across sexes, but female stature is low, despite polygenic scores identical to males, and suggests cultural factors preferentially supported male recovery from stress. In Mediterranean populations, sexual dimorphism is reduced, indicating male vulnerability to stress and no strong cultural preference for males. Our analysis indicates that biological effects of sex-specific inequities can be linked to cultural influences at least as early as 7000 years ago, and culture, more than environment or genetics, drove height disparities in Early Neolithic Europe.

1 Introduction

Human skeletal variation reflects varying combinations of genetic, cultural, and environmental factors. While there are many links between culture and health in the modern world, the history and evolution of these relationships are not always well established. Due to the entanglement of these factors, our ability to draw conclusions about their effects has been limited in archaeological data. With the recent advent of ancient DNA sequencing technology, genetic information from ancient populations has become increasingly available. However, attempting to analyze changing patterns of

variation based solely on genetic data is difficult—genotypes do not necessarily equate to phenotypes due to the effect of the environment.[1] Similarly, while it is tempting to predict phenotypic changes in ancient people based on their genetic variation, this is currently challenging as genetic effects are not always transferable across populations.[2] Our solution is to integrate these complementary fields to construct multidisciplinary analyses with phenotype, genotype, culture, and environment data from ancient human populations. This approach allows us to begin to separate the effects of these variables and reveal the interactions between genes, environment, and culture which are critical in shaping human health and variation.

Many traits of interest, including height, are highly polygenic, with thousands of independent genetic variants contributing significantly to heritability. One common approach to addressing the role of genetics in morphological change is to compare patterns of phenotypic variation with genetic ancestry or genome-wide patterns of genetic variation.[3–5] However, even for highly polygenic traits like height, genome-wide variation may not be directly relevant, leading to spurious associations between genetic effects, ancestry, and environmental confounds. For example, if a population is tall and has a high proportion of ancestry from Neolithic sources, it could be concluded that Neolithic ancestry is associated with "genetic tallness"; however, the effects could equally be non-genetic and related to lifestyle changes associated with agriculture. An alternative approach is to focus only on genetic variation that is known to be associated with a specific trait.[6,7] Effect sizes for these trait-related variants estimated from genome-wide association studies (GWAS) of present-day individuals can be combined with genetic data from ancient individuals to calculate polygenic risk scores (PRS), which can be thought of as estimated genetic values for the phenotype. In European ancestry populations, polygenic scores for height can explain up to 25% of phenotypic variation in present-day individuals,[8] and 6-8% of variation in ancient individuals.[9,10] On a broad scale, temporal changes in polygenic score over time in Europe are qualitatively consistent with changes in stature as inferred from the skeletal record,[11] while local deviations from this pattern provide evidence of environmental effects.[10,11]

Analyses of human populations over tens of thousands of years involve individuals that are diverse in genetic ancestry, environment, and culture and it is challenging to exclude the possibility of confounding by unmeasured variables. We therefore focus specifically on the European Early Neolithic. One of the most studied periods in prehistory, it represents a fundamental shift in technology, culture, and genetics. In particular, the *Linearbandkeramik* (LBK) culture of Central Europe is one of the most comprehensively documented Early Neolithic cultures, with an abundance of excavated settlements and cemeteries.[12] LBK groups tended to choose settlement locations based on the presence of rich loess soils for farming, and the northern edge of these soils appears to delineate the northern limit of LBK sites.[13,14] Bioarchaeological evidence indicates broad regional differences between individuals from northern settlements in this agricultural boundary zone vs southern settlements in a climate zone that was more comfortable for Neolithic crops.[15,16] Based on this, we divided our Central European group into Northern (above 50°N latitude) and Southern

(below 50°N) populations. The Mesolithic hunter-gatherer population in Central Europe made a limited genetic contribution to the LBK population, whose members harbor only traces of hunter-gatherer admixture.[7,17,18] Contemporary populations from southeastern Europe have similarly low levels of hunter-gatherer ancestry.[19] In contrast, Neolithic southern European populations associated with the Cardial and Impressed Ware cultures followed a separate migration route (Figure 1), occupied a milder climate zone, and carried more Mesolithic ancestry.[17,20] Individuals in this region tend to be shorter than those from Central Europe and combined with their admixed ancestry this has led to suggestions of a genetic basis for decreased statures in this region.[7,21]

By comparing and contrasting four closely related archaeological populations, we aim to investigate how differences in environment and genetics combine to produce observed phenotypes. We collected genetic data, skeletal metrics, paleopathology, and dietary stable isotopes to begin separating the effects of each on Neolithic stature trends. By specifically investigating and controlling for the effects of genetics in these samples, we are able to provide nuanced interpretations of height variation, gain a better understanding of the aspects of height which are controlled by genetics or environment, and show evidence for sex-specific cultural effects which modify the genetically predicted patterns. We illustrate the strengths of leveraging multidisciplinary datasets, and indicate caution when analyzing genotype-phenotype relationships without complete data, especially for traits which are not preserved in the archaeological record and cannot be directly tested. This integrated analysis highlights the role of plasticity in morphology, and establishes culturally mediated disparities at least as early as the European Neolithic.

2 Results

2.1 Distributions of statures, polygenic scores, and stable isotope values

We collected either genetic, dietary stable isotope, paleopathology, or skeletal metric data from 1269 individuals associated with the archaeological LBK culture in Northern and Southern Central Europe dated to between 7700-6900 years before present (BP), as well as 139 individuals from the southeastern (Balkan) and 127 individuals from the southern (Mediterranean) regions dated between 8000-6000 BP. All individuals included in the analysis have at least one of the four data types available (Materials and Methods; Figure 2d; Supplementary Figure 1). We classify individuals as "male" or "female" based on either chromosomal sex or the morphological category they most closely resemble. We recognize that these two categories do not represent the full range of sex and intersex morphologies in humans; however, to our knowledge, there currently are no methods by which to identify other morphologies in the archaeological record. Though ideally morphological sex variation would be analyzed on a continuous scale, in this study we incorporate previously published data and thus use the word "dimorphism" to reflect that we analyze only two sex categories.

Observed patterns of femur length vary between sexes and populations. We find no evidence of difference between male femora of the Central and Balkan regions (North vs. South Central:

p=0.221, β=0.419cm, SE=0.341; North Central vs Balkan: p=0.632, β=0.444cm, SE=0.925), but Mediterranean males are significantly shorter (p=0.003, β=-1.314cm, se=0.433). Conversely, female femora show a different pattern, with no significant difference when comparing Mediterranean to South Central (p=0.645, β=-0.251cm, SE=0.542), and Balkan (p=0.491, β=-0.703cm, SE=1.019) populations, but significantly shorter values comparing North Central to South Central populations ($p < 0.001$, β=-1.988cm, SE=0.560; Figure 3a). Differences between male and female femur lengths are highly significant in all populations (p<0.001, β=4.128cm, SE=0.228). In contrast to the differences in femoral lengths, we see no difference in polygenic scores for height between all populations (pairwise t-tests min. p>0.9, DF=41-71) using the clumping/thresholding PRS construction (Figure 3b). This result is expected given the extensive overlap between the 95% confidence intervals for the population PRS means (Figure 3b). PRS constructed with LDpred show Mediterranean individuals to be shorter than the other populations (max. p=0.0293, DF=41-71; Supplementary Figure 5). However, a PRS constructed using summary statistics derived from between-sibling analysis[22] finds no difference in genetic values between populations with either PRS construction method (clumping/thresholding: min. p=0.099, DF=41-71; LDpred: min. p=0.143, DF=41-71), so we conclude that apparent lower Mediterranean PRS may be due to population stratification in the GWAS data and may not reflect a true genetic difference. There are no significant differences between male and female PRS in any population (min. p=0.7647, β=0.076cm, SE=0.254; Figure 3b), providing no evidence for a genetic basis to this dimorphism.

Signatures of δ^{13}C and δ^{15}N suggest different dietary patterns in each of the analyzed groups (Figure 3c, d). Both the Mediterranean and Balkan groups are significantly distinct from the Central in δ^{13}C (max. p<0.001, β=0.541‰, SE=0.136) while we find no difference in δ^{15}N values in the Central and Mediterranean groups (p=0.101, β=0.7‰, SE=0.42557), but distinct in the Balkans (p<0.001, β=5.0488‰, SE=0.237). Generally, the Balkan population is characterized by high δ^{15}N values, while Mediterranean populations show high δ^{13}C relative to the Central Europeans (Figure3c). The exception to this pattern is a cluster of individuals, classified as Balkan in our analysis, which overlaps with the North Central population as well as some of the Mediterranean. These points represent individuals from present-day Greece and indicate that the diets of these peoples might better be classified as Mediterranean than Balkan. Nitrogen values are higher in males compared to females (ANOVA F=21.0, DF=1,622, p<0.001) differ across regions (F=390.2, DF=3,622, p<0.001; Figure 3d), and the interaction between sex and region is significant (F=3.5, DF=3,622, p=0.016).

2.2 Patterns of non-genetic factors in Central Europe

The most dramatic observation is the difference in female stature and consequent sexual dimporphism in Northern compared to Southern Central Europe. Female femora in the North are significantly shorter than female femora in the South (p<0.001, β=-1.988cm, SE=0.560), while male femora are not significantly different (p=0.221; β=0.419cm, SE=0.341; Figure 3a). On aver-

age, male femora from the North are about 13% longer than female femora, Southern Central and Balkan male femora are about 9% and 11% longer respectively, and Mediterranean male femora are only 5% longer (Figure 3a). These values are reduced slightly when calculated using estimated statures instead of femora (North Central: 10%, South Central: 7%, Balkans: 8%, Mediterranean: 4%), possibly due to error associated with stature estimation (see Ref. 9) and body proportions, or because the relationship between femur length and stature is different between males and females. Where we have both genetic and metric data for the same individuals, there is a qualitative relationship between femur length and PRS; PRS tends to increase as femur lengths increase (Supplementary Figure 3b). However, the effect of PRS on femur length is not significant, likely due to the small number of individuals with both types of data available in the sample (n=55; p=0.059, β=0.781, SE=0.405cm/SD, BF=0.278).

Overall, trends in dietary stable isotopes show that individuals in Southern Central Europe have significantly higher nitrogen values (δ^{15}N; p<0.001; β=0.865‰, SE=0.143) and lower carbon (δ^{13}; p<0.001; β=-0.383‰, SE=0.082) as compared to the North. However, while males in both regions qualitatively have higher nitrogen, the interaction between sex and region is not significant (p=0.872; β=-0.0316‰, SE=0.196) Figure 4a), indicating that the difference between male and female values in each region is not significant, though the Bayes factor (BF=-0.012) suggests this might be due to a lack of power. Carbon values are similar between the sexes (p=0.473; β=0.081‰, SE=0.113, BF=-1.749). For individuals with both stature and stable isotope values, we find no statistically significant relationship between femur length and δ^{15}N (p=0.489; β=-0.223‰, SE=0.322, BF=-1.495) or δ^{13}C (p=0.114; β=0.550‰, SE=0.346, BF=-0.651) in either Central group, separately or combined.

We do, however, find a statistically significant relationship between presence of linear enamel hypoplasias (LEH) and shorter femora, suggesting that LEH may reflect an underlying variable in childhood that also affects stature (p=0.0214; β=-1.275cm, SE=0.548; Figure 4c). Both males and females from the North are more likely to have LEH than individuals living in the South (logistic regression p=0.0018; β=-1.386, se=0.444). Indeed, over 50% of the Northern sample have LEH while they are only present in about 20% of the Southern (Figure 4b). There is no significant difference between the number of males and females with LEH in either region (p=0.832; β=0.113, SE=0.534, BF=-1.425). Though the interaction effect between sex and LEH on femur length is not significant (p=0.072; β=1.390, SE=0.765, BF=-0.121), qualitatively the effect of LEH on femur length appears greater in females than in males and Bayes Factors suggest our sample does not have the power to detect the interaction (Figure 4c). When the sexes are analyzed separately, females with LEH do have significantly shorter femora than those without (p=0.022; β=-1.334cm, SE=0.570, BF=0.963), which is not the case for males (p=0.741; β=0.186cm, SE=0.562, BF=-1.264). We hypothesize that the relationship between LEH and femur length is driven by females, but our sample size is too small to formally show this. Incidence of cribra orbitalia is also significantly higher in the Northern region than in the Southern (p<0.001, β=-1.890, SE=0.449), though we do not see a relationship with

femur length (p=0.609; β=-0.324cm, SE=0.633, BF=-1.245). We find no evidence of differences in porotic hyperostosis between the North and South (p=0.524; β=-0.265, SE=0.417, BF=-1.345), nor do we detect a relationship between porotic hyperostosis and femur length (p=0.878; β=0.0567cm, SE=0.369, BF=-1.747).

In summary, comparison of Northern and Southern Central Europe identifies no predicted genetic difference in stature, which is consistent with male but not female femur length. This suggests a non-genetic basis for reduced female stature. Stable isotope data and skeletal stress indicators suggest lower protein intake and more general stress in the North; however, males and females overall appear equally affected by these variables. Despite a similar number of hypoplasias in both sexes, shorter femora in females suggest that increased general stress, due to other unmeasured environmental or cultural factors, leads to a female-specific reduction in stature.

2.3 Patterns of genetic ancestry in the Mediterranean

In contrast to Northern Central Europe, Mediterranean Neolithic males are shorter than other groups, but females are not. PCA indicates that individuals from the Central regions and the Balkans share similar genetic ancestry while those from the Mediterranean are distinct (Figure 2c; unimputed PCA in Supplementary Figure 4a), a difference known to be due to higher levels of hunter-gatherer ancestry in the Mediterranean.[17] We therefore additionally compared our samples to Mesolithic individuals of Western Hunter-Gatherer (WHG) ancestry, as well as individuals from early Neolithic Anatolia. These two groups represent source populations for the two largest ancestry components in Europe at this time.[7,17]

On the PCA plots of these extended data, Neolithic Anatolians cluster with Central and Balkan groups. While Mediterraneans are near the farmer cluster, they are shifted towards the WHG (Figure 5a; unimputed PCA in Supplementary Figure 4b). ADMIXTURE analysis of all six populations supports this conclusion, showing significantly increased proportions of WHG ancestry in the Neolithic Mediterranean as compared with Central Europe (p<0.001; β=0.354%, SE=0.438). The average proportion of WHG ancestry in the Mediterranean is 11.4% (SE=0.53%); in the Balkans, 5.3% (SE=se=0.57%); in the South Central, 4.1% (SE=0.16%); and in the North Central, 1.1% (SE=0.09%). If there are significant PRS differences between Mediterranean and other populations, they would likely be linked to this greater WHG ancestry and reflect genetic differences between WHG and other populations.

Computing PRS using clumping/thresholding, we find that the WHG have the lowest PRS of any population in our data (p<0.001, β=-0.903SD, SE=0.232), while Anatolians are similar to the Balkan and Central Europeans. Among individuals, proportions of WHG ancestry are significantly associated with lower PRS (p=0.006, β=-0.594%/SD, SE=0.214, BF=1.481). However, when we compute PRS with an infinitesimal LDpred2 model, Mediterranean PRS is intermediate between Neolithic Europeans and WHG. When we repeat the LDpred analysis using summary statistics computed from between-sibling GWAS,[22] we find that the direction of Hunter-Gatherer PRS flips,

and they have significantly greater PRS than the other groups p=0.002; β=0.701, SE=0.225; Supplementary Figure 3a). The inconsistency of these results shows that the apparent PRS difference between WHG and Neolithic populations is highly sensitive to the PRS construction and summary statistics. This may indicate uncorrected population stratification in the non-sibling GWAS.[23,24] We therefore conclude that there is no strong evidence for a genetic difference in stature between Mediterranean and other Neolithic populations.

3 Discussion

Understanding the causes of past stature variation not only allows us to understand ancient communities, but may also provide us with insights into the origin and evolution of modern health patterns. However, interpretations of human stature variation through time remain confounded by the difficulty of separating genetic and environmental effects, obscuring trends. Recently, several researchers have begun to compile multivariate datasets for the purposes of understanding human stature[10,25–27]; however, many of these analyses do not directly take genetic effects into account,[25,27] or cover very broad temporal or geographic ranges.[10,26] In contrast, we aim to understand these processes on a finer scale to better interpret the outcomes of biological and environmental/cultural interactions. For instance, previous studies of stature variation from the Mesolithic to Neolithic indicated that Neolithic individuals were not achieving their genetic height potential,[10] but our analyses suggest that this effect might be heterogeneous, pertaining more to some locations and portions of society than others. Finally, while interesting in its own right, height can also serve as a model trait for how to incorporate genetics and anthropological data into studies of human morphology and variation. Here, by integrating genetic, cultural, and environmental data, we are able to begin teasing apart the contributions of genetic and non-genetic factors in producing the observed phenotypic variation. We also illustrate the existing limitations of interpreting genetic data.

We note several limitations in our data which restrict our analysis and interpretations. While we have a narrowly defined geographic and temporal sample in our Central European group, limitations of existing data required us to broaden our inclusion criteria in other regions to strike a balance between controlling variables and having statistical power. Though our South Central group is geographically more broad than the North, they fall within a specific time range and genetic and osteological samples are well-matched geographically. However, comparative materials from the Mediterranean and Balkan regions necessitated a broader time period to compile a sample size that would be statistically meaningful. Restricting these two groups to either time-match Central Europe or to geographically match data types results in very small samples with little power to address key questions (Supplementary Figure 7). The range of these populations is relatively broad and our approach is averaging trends across the whole region, likely obscuring local nuances. Due to these limitations, we have focused the majority of our interpretation on the Central European populations

where we have the most and highest quality data, and have thus restricted our discussion of the Mediterranean and Balkan populations to only those results that we felt were adequately supported. Future, more detailed studies of these regions would have the potential to reveal fine-scale dynamics like those we observe in Central Europe.

Overall, the Central and Balkan groups are genetically homogenous with similar levels of hunter-gatherer admixture and polygenic scores, while Mediterranean individuals have more hunter-gatherer ancestry (consistent with previous observations[7,17,18]).This may be associated with lower PRS, though this relationship is not robust and may simply reflect residual population stratification in the GWAS. None of our populations show evidence for substantial genetic differences in height between sexes (Supplemetary Figure 2), which is expected given that there is little evidence for sex differences in ancestry, or of sex-specific genetic effects on stature.[28–30] We can therefore largely exclude a genetic contribution to differences in stature between North Central individuals and other groups, while we find no strong evidence for a genetic contribution to shorter Mediterranean stature.

3.1 Central European dimorphism reflects the effects of culture

Dietary differences between Southern and Northern Central European populations may indicate environmental stress in the North. In the early European Neolithic, the expansion of agriculture is thought to have been largely limited by poor soils and climate, as colder temperatures and decreased daylight made it increasingly difficult to grow early cereals (wheat, barley) and pulses (peas).[31,32] The boundary to which these plants could be grown has been estimated to coincide with the northern limits of the LBK culture,[13,33] and the majority of our Northern sites are concentrated near this climate edge in areas of fertile loess soils.[14] However, as there are many nuances which affect the interpretation of stable isotope values, especially between populations, differences between our Northern and Southern groups may not be as dramatic as they appear. An examination of isotope values from herbivorous animals in our study regions (using data from Refs. 16,34–38) indicates that baseline values of δ^{15}N are elevated in the South Central region as compared to the North, potentially due to differences in climate and the use of manure as fertilizer. Therefore some portion of the difference between Northern and Southern nitrogen values might be attributable to variation in climate and farming practices rather than diet. Differences in carbon values between populations can be similarly sensitive to environment and we feel that interpreting the carbon results would be difficult without a more in-depth isotopic analysis. However, isotopic values from other studies show a higher proportion of plant vs animal foods in the North, particularly low protein domesticated cereal grains.[15,35,39] Additionally, the available archaeological evidence supports some level of dietary difference between the Northern and Southern regions. While remains of cattle and dairy production are documented in Northern LBK sites,[40,41] there is less archaeological evidence for the presence of other wild or domesticated animals that are seen in the South, indicating the people of this region were highly reliant on plant foods.[39,42] We conclude that our observed differences in Northern and Southern stable isotope values probably reflect both dietary factors and differences in climate or

farming practices.

It is therefore not surprising that people of the Northern Central region exhibit evidence of increased stress potentially due to unreliable and lower quality food resources. Lower protein consumption could be an indicator of dietary stress and has been linked to decreased stature.[43] Diet can affect dimorphism in some cases,[44] but the isotopic signatures of males and females in the North Central and South Central regions are very similar, suggesting that this factor alone does not explain reduced female stature in the North. Femur length and isotope values for individuals are not significantly associated in our data, an indication that either diet has little effect on Neolithic stature or stable isotopes do not capture the elements of diet relevant to height. Alternatively, it is possible the range of variation in our data is too small to see this effect, or our sample of individuals with both metric and isotopic data is not large enough. In addition, we only analyzed adult samples and while the isotopic values of weaned children in the LBK fall within the range of adults,[42] it is possible that there could be sex differences in childhood diets affecting femur growth. Future studies incorporating collagen from long bones or teeth, rather than from ribs as we have here, would give dietary evidence with greater time depth, and might be able to provide more nuanced interpretations in the absence of a known-sex sub-adult population.

Paleopathological analysis also indicates increased stress in the Northern population in the form of increased incidence of linear enamel hypoplasia and cribra orbitalia. The causes of LEH formation are varied and their appearance in the bioarchaeological record is generally interpreted as a non-specific indication of childhood stress.[45] Other archaeological sites have reported a high instance of LEH with high sexual dimorphism ratios in areas of Neolithic Europe, though the cause and meaning of these patterns was not explored (e.g. Ref. 46 and references therein). It has been suggested that cribra orbitalia might also reflect childhood stress, specifically anaemias, even when seen in adults.[47] Our results are consistent with others who have considered the same paleopathologies and found a qualitative relationship between presence of paleopathology and shorter femora.[10] In our data, incidence of both LEH and cribra orbitalia are higher in Northern compared to Southern Central Europe, but are not different between sexes in either group. The association between shorter femora and presence of LEH appears to be driven by females, suggesting a moderating factor causing a female-specific effect despite equal incidence of LEH in both sexes.

While we see a general increase in stress shared between sexes in North Central Europe, typical population-level stress responses usually show male vulnerability and female buffering effects.[48–50] Though the exact causes and mechanisms are not well understood, female biology tends to have a less extreme response, or is "buffered", to many diseases[51–53] and environmental changes[54] compared to males. Our data indicate an opposite pattern in Central Europe, and no evidence of a variable which acts upon females alone. However, the Northern population shows sexual dimorphism that is extreme by present-day standards. In most modern global populations the ratio of male to female height is 1.06-1.08[55] (ratios in Ref. 55 range up to 1.12, but population locations or cultural affiliations are not given, see Ref. 44), though it is difficult to know how to compare height versus

femur length ratios as the transformation from metrics to stature scales differently in males and females. Based on 147 European individuals from the past 100 years (using data from Ref. 56), we find that the height ratio is very similar to the ratio of femur length—typically within 1%. We therefore conclude that dimorphism ratios in Southern Central (1.09) and Balkan (1.11) Europeans are elevated and the ratio in the North Central region is exceptionally high (1.14). Few modern populations have height dimorphism ratios as high as 1.10, and those that we could find in the literature come from India[57] and the United Arab Emirates,[58] both countries known for their cultural preferences and biases for male children.[59]

We therefore hypothesize that the effects of high environmental stress in the North were modulated by culture. Other researchers have noted specific situations in which culture buffers males against environmental effects and creates vulnerability in females: there is an association between decreased female stature and polygyny in cultures around the globe;[60] female height was more influenced by economic conditions during infancy and early childhood than males in lower-class 19th-century Europe;[61] sexual dimorphism ratios in modern Chile decreased after the institution of social and government programs to combat gender inequality;[62] 20th-century female stature decreased in India during times of environmental stress due to sexually disproportionate investment of scarce resources;[59] and son preference has been shown to decrease the height of female children in Indian families regardless of birth order.[63] In LBK sites, strontium isotope values show that females are more likely to be non-local compared to males, suggesting patrilocality and potential differences in cultural treatment of females.[14,64] In parallel to our evidence for higher biological variation in females, ongoing discussion about the relationship between biological sex and the formation of gendered identities in the LBK suggests more variation in the roles and identities of females compared to males.[65] We therefore suggest that culturally mediated differences led to sex-specific stress responses in Neolithic Central Europe *via* cultural practices which either directly decreased female stature or, more likely, supported catch-up growth preferentially in males. Though dimorphism ratios in the South Central and Balkan regions are not as extreme as in the North, they are elevated and also consistent with this pattern of male bias, but response is likely less exaggerated due to lower environmental stress conditions.

3.2 Mediterranean dimorphism may have an environmental basis

In the Early Neolithic Mediterranean population we see decreased male stature and low dimorphism ratios (1.05) relative to other Neolithic populations. Mediterranean populations are genetically distinct from other Early Neolithic groups with a higher proportion of WHG ancestry. In some analyses, WHG ancestry proportion correlates with lower PRS for height. However, PRS in the Mediterranean and WHG populations are sensitive to PRS construction method likely due to residual population stratification in the GWAS. These inconsistent results mean that we can neither confirm nor exclude the possibility of a genetic contribution to differences in stature between the Mediterranean and other Early Neolithic populations, though on balance we find the likelihood for a substantial genetic

contribution to be low. Even if it were not, the genetic effects alone could not explain the reduced dimorphism ratio, emphasizing the need to also consider cultural/environmental effects.

While the dimorphism ratio in the Mediterranean Neolithic is low, it is not outside the range of present-day populations.[55] In fact, while males are relatively short, the longest average female femur lengths of our data are in the Mediterranean. This reduction in dimorphism is commonly seen in populations where the sexes experience an equal stress burden: as males tend to be more sensitive, decreasing their height, females are biologically buffered and stature remains consistent.[48–50,66] Although we do not have paleopathological stress data for the Mediterranean individuals in our sample, published values for other Neolithic Mediterranean populations are generally similar to those for South Central Europe,[67–69] with exceptions.[70] Dietary isotopes indicate that the Mediterranean diet differs in some aspects, with increased δ^{13}C values compared to the other Neolithic populations, but similar δ^{15}N values. Our data indicates similar protein intake and low-level stress as other Neolithic populations, but do not suggest any clear hypothesis for the difference in male stature between the Mediterranean and other Neolithic groups. Possible differences in Mediterranean body proportions which are not captured by femur length should be mentioned as a caveat, though this likely would not be enough to account for the differences in stature compared to the rest of Europe, and would not affect observed dimorphism within the population. Our hypothesis is that the Mediterranean experienced similar levels of environmental stress as other Neolithic groups, but that they did not share the cultural practices which preferentially supported males and increased female vulnerability.

3.3 Conclusion

By integrating genetic and anthropological data, we are able to begin to understand the contributions of genetics and environment to human variation, allowing us to better interpret the genetic, environmental, and cultural landscapes of Neolithic Europe. Our results are consistent with a model in which sexually dimorphic differences in femur length are culturally and environmentally driven: relatively low dimorphism in the Mediterranean caused by female buffering to environmental stress and less cultural male preference, and high dimorphism in Northern Central Europe caused by the interaction of relatively high environmental stress and strong cultural male preference. Some analyses suggest that differences in average femur length between Central/Southeastern Europe and the Mediterranean are associated with differing genetic ancestries, but lack of robustness, uncertainty about the transferrability of polygenic scores, and questions of residual population stratification prevent us from interpreting this conclusively. In this study we focused on the European Early Neolithic because of its relative genetic, cultural, and environmental homogeneity, but, with more data, these methods could be extended to other populations, traits, and timescales to further explore the effects of human culture on biological variation. Using this approach, we gain a deeper understanding of the relationship between phenotypic plasticity, culture and genetic architecture, which constrain the mechanisms by which human biology adapts to environment.

4 Materials and Methods

We collected a combination of genetic, dietary stable isotope, skeletal metric, and paleopathological (stress) data from 1282 individuals from the Central European Early Neolithic associated with the archaeological LBK culture, approximately 7700-6900 BP (Figure 2, Supplementary Table 1). As there is archaeological evidence for broad regional variation within the LBK and our sampled sites form clear geographic groups, we divided these individuals into two regions based on geographical location, those to the north of 50°N latitude (North Central) and those to the south (South Central) (Figure 2a-b; North Central n=203, n femur length=133, n isotopes=100, n aDNA=67, n stress=83; South Central n=1067, n femur length=187, n isotopes=670, n aDNA=72, n stress=523). Each individual has at least one of the data types, and while some individuals have multiple data types, the overlaps are small (Supplementary Figure 1).

To provide wider context, we also compared Central individuals to other Neolithic populations from southern European (Mediterranean) and southeastern European (Balkan) regions, and restricted to individuals dated to 8000-6000 BP. We chose these regions as the Neolithic transition occurs at similar times and is associated with populations closely related to Central Europe. The acceptable date range for inclusion in the study was expanded from that which defines the LBK as these dates encompass comparable Early Neolithic phases in other parts of Europe while maximizing the number of eligible individuals. There could be a possibility that the later Balkan and Mediterranean individuals were more adapted to Neolithic life than the Central European groups, as these samples cover a longer time period, but we found no statistical within-population differences in our variables between the early and late ranges of our time span (minimum p=0.08). We excluded areas such as Scandinavia and Britain, where Neolithic technologies were not generally adopted until a later date. For the final analysis, we included 127 Mediterranean (n femur length=60, n isotopes=25, n aDNA=42) and 139 Balkan (n femur length=12, n isotopes=78, n aDNA=49) individuals (Figure 2). Unfortunately, there is a wide range of recording and reporting used for skeletal stress indicators, and it was not possible to build a statistically powerful dataset in these two populations for comparison; as a result, we did not analyze paleopathology in these populations. Finally, we collected genetic data from Mesolithic hunter-gatherer (n=25, 14000-7080BP, south of 48°N) and Anatolian Neolithic (n=21) individuals for additional comparison.

4.1 Genetic data

We obtained genetic data for a total of 276 individuals.[7,18–20,71–87] Most data were generated by targeting a set of 1.24 million SNPs (the "1240k" capture reagent).[17,75] For each individual, we randomly selected a single allele from each of the 1240k sites. Coverage in our dataset is low (median coverage=0.33; coverage above 0.60 n=71), and typically, it is not possible to directly infer diploid genotypes, potentially limiting PRS performance. Imputation of missing genotypes has been shown to help improve polygenic predictions for low coverage ancient samples,[9] and we therefore

imputed diploid genotypes using the two-stage method described in that paper, restricting to SNPs in the 1240k set.

We calculated polygenic scores as previously described.[9] Briefly, we used standing height summary statistics generated by the Neale Lab from 456,000 individuals of European ancestry in the UK Biobank[88] for analyses of combined-sex PRS, and summary statistics from male- and female-only UK Biobank GWAS also generated by the Neale Lab.[88] To test the potential effects of residual population structure in our data, we also computed PRS using additional summary statistics from a between-sibling GWAS (n=99,997).[22] We intersected the sites from each of these datasets with those on the 1240k array and then further restricted to HapMap3 SNPs (SNPs n=405,000). We computed polygenic scores using both a clumping/thresholding approach (r^2=0.3, p-value cutoff=10^{-6}, 100kb windows in *plink2*[89]), and an infinitesimal *LDpred2* model using their pre-computed LD reference panel.[90] Finally, we computed polygenic scores using the --score command in *plink2*. In order to maximize the possibility of detecting sex-specific effects, we generated sex-specific PRS using three different approaches: 1) calculating PRS for all individuals using the female summary statistics; 2) calculating PRS for all individuals using the male summary statistics; and 3) calculating PRS for males and females separately using their respective summary statistics. While approach 3 seems at first to be the best for detecting these effects, observed patterns potentially become difficult to interpret due to differences in scaling between male and female PRS calculated as separate datasets.

We computed principal components for both unimputed and imputed data using *smartpca*,[91] projecting ancient individuals onto principal component axes defined by 777 present-day West Eurasian individuals.[92] We also estimated K=2 unsupervised ADMIXTURE[93] components for unimputed ancient individuals after first LD pruning using the command --pairwise-indep 200 25 0.4 in *plink2*.

4.2 Osteology and stable isotope data

We aggregated skeletal metric data from both published[56,94–98] and unpublished (n=28) sources. Maximum femur lengths were recorded when available, otherwise we estimated femur length from published stature estimates.[9] Estimated femur lengths correlate highly with stature estimates, but decrease the error that results from combining different estimations methods. The method from Ref. 99 provides separate equations for estimating the statures of northern vs. southern Europeans when using the tibia, due to differences in body proportions between the regions. There are two Mediterranean samples for which we estimated the length of the femur based on statures which used the southern tibia equation. Ref. 99 does not provide regional equations for femur estimation, so for these two individuals, we estimated femur length using the reverse of this region-agnostic femur equation.

For the individuals in this study who do not have genetic data, morphology was used to estimate sex. We acknowledge that there are a range of sex and intersex morphologies present in humans, however, current methods of skeletal sex estimation only classify morphologies as associated with male or female categories. We use the term "dimorphism" to reflect that we analyze two sex clas-

sifications in this study, rather than as a statement about the nature of human variation. The majority of individuals in the current study have been taken from previous publications, and we used the sexes which had been estimated by those authors. For the individuals in our study which have not been previously published, sex was determined by co-authors using a 5-point scale on the cranium and pelvis as described by Ref. 100. For all individuals, sexes determined as probable male or probable female were coded in our study as either male or female as appropriate. Subadults and those with indeterminate morphologies were coded as NA, resulting in these individuals being dropped from the sex-specific analyses. The majority of sexes for individuals with metric data were determined by, or supervised by, co-authors and the remainder (n=13) either have genetic sexes or come from Ref. 56 which we consider a reliable source. Despite generally high accuracy for morphological sex determination, some level of uncertainty always remains, mainly due to variation in sexual dimorphism and preservation of the remains.[101] Sex estimations for our sample have all been performed in the last 20 years, and the majority within the last 5 years, meaning the researchers who performed them should be aware of avoiding the biases which can affect sex-ratios in the estimations of older data. Our dataset is large enough that small errors in classification of sex should not make substantial differences to results or interpretation, but the potential for inaccurate morphological sex estimations must always be considered in any osteological analysis. A large portion of our paleopathology data comes from tables S3 and S6 of Ref. 102, in which there are many instances of the same individual listed in both tables, but with discordant sex estimations. As we could not determine the reason for these discrepancies, we used the sex which was reported in the original publications cited as sources for their data. The few individuals (n=3) for whom this could not be resolved were treated as indeterminate and coded as NA. Ages were determined based on the average of the age range reported for each individual in their original publications.

For the paleopathological data in Central Europe, we took data from published sources,[94,95,102–105] as presence/absence of linear enamel hypoplasia (LEH), porotic hyperostosis, and cribra orbitalia. These three pathologies are often used by anthropologists as indicators of general, non-specific stress experienced by individuals or populations. While the exact eitiologies of these pathologies are generally not known, they have been shown to change through time within and between populations, and often correlate with environmental, social, or cultural shifts. Linear enamel hypoplasias are horizontal defects in tooth enamel that form during episodes of childhood stress severe enough to interrupt growth for some period of time, usually associated with dietary deficiency or infectious disease.[45] Individuals can exhibit one or multiple LEH on single or multiple teeth and in order to minimize errors from differences in reporting LEH in the literature, we have simply recorded whether an individual had any LEH (present) or none (absent). Porotic hyperostosis and cribra orbitalia are both porous lesions that are distinguished by their appearance on either the cranial vault or roof of the eye orbit respectively. The eitiologies of these are mostly unknown and though they are traditionally associated with amaemias, there are also a number of other conditions that can produce the same type of lesions. Medically, there is little evidence of these pathological changes

despite their prevalence in the bioarchaeological record.[106] Similar to LEH, we have recorded these as either present or absent for each individual in order to standardize between reporting conventions across publications.

While sensitive to confounding factors such as climate, vegetation, and individual metabolism,[107] δ^{13}C and δ^{15}N stable isotope data can be used to reconstruct aspects of diet.[108] Here, carbon values are indicative of dietary plant resources and of the terrestrial vs marine vs limnic provenance of food, while nitrogen values are mainly associated with dietary protein intake and generally indicate proportions of plant- vs animal-based diets.[107,108] We collected dietary stable isotopes δ^{13}C and δ^{15}N from published[16,34–36,42,94,96,102,104,105,107,109–112] and unpublished (n=38) reports. We excluded atomic mass spectrometer (AMS) values derived from radiocarbon dating, as they may not be comparable to isotope-ratio mass spectrometer (IRMS) measurements, as well as values from children below the age of three, due to increased nitrogen values from breastfeeding. Stable isotope values from older children were included in population-wide diet analyses as the isotope ranges fall within those of adults; however, we only included adults with estimated sexes in the sex-based diet analyses. If information on the sampled material was available, we chose values measured from rib collagen, as these samples are most plentiful, though they only reflect the last few years of the individual's life.

All previously unpublished osteological data was collected and analyzed by co-authors with permission from the necessary regulating organizations and in accordance with German laws and policies.

4.3 Statistical models

We tested the effects of PRS, femur length, and isotope data on stature using linear regression models including sex and geographic region as covariates in combination with other variables as appropriate (e.g., femur $\sim$ sex + region + PRS; δ^{15}N $\sim$ sex + region + femur). We included interaction terms to test the relationships between geographic regions and sex (e.g., femur $\sim$ region * sex) logistic regression with the same covariates to test for factors affecting presence/absence of paleopathologies. Two-tailed pairwise t-tests corrected for multiple testing using the Holm method were used to test for regional differences in the PRS and regional differences with-in sex in the femora since these tests do not include other covariates. Two-way ANOVA of the linear models were used to test for interaction effects between sex and region on δ^{15}N since the linear model was likely under powered to detect these effects. Changing the method used to correct for multiple comparisons did not alter the results which appear to be robust to the correction method. We did not test data for normality or equal variances. Natural log of Bayes Factors are reported for null results using the default priors and settings for lmBF() from the *BayesFactor* package v. 0.9.12-4.4[113] in R which implements the g-mixture prior described in Ref. 114. We carried out all other statistical tests using the base functions in R version 4.3.[115]

4.4 Data Availability

All non-genetic data and polygenic scores used in this analysis are provided in Supplementary Table 1. Original ancient DNA data files can be downloaded from the resources provided in their cited publications and from the Allen Ancient DNA Resource (AADR) https://reich.hms.harvard.edu/allen-ancient-d GWAS summary statistics can be downloaded from the Neale Lab (http://www.nealelab.is and the sibling GWAS summary statistics from OpenGWAS (https://gwas.mrcieu.ac.uk/). Previously published osteological data can be found in their cited sources which include the LiVES database (doi: 10.17171/2-12-2-1) and Dr. Christopher Ruff's public dataset (https://www.hopkinsmedicine.org/ Additional data from the PCAs, ADMIXTURE analysis, and sex-specific polygenic scores are available at https://github.com/mathilab/Neolithic_height.git.

4.5 Code Availability

R code used in this analysis is available at https://github.com/mathilab/Neolithic_height.git.

4.6 Acknowledgements

This work was supported by a grant from the National Science Foundation [BCS2123627] to IM. Metric skeletal data collection from Stuttgart-Mühlhausen was funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) [RO 4148/1-1]. The investigations of the skeletal collections from Derenburg, Karsdorf, and Halberstadt were supported by DFG grants [Al 287/7-1 and 7-3] to KWA and Me 3245/1-1 and 1-3] to HM. The content is the responsibility of the authors and does not necessarily represent the official view of the NSF or other funders. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

4.7 Author Contributions

S.L.C. designed the study, collected data, performed analysis, and wrote the manuscript; I.M. designed the study and wrote the manuscript; N.N. and K.W.A., contributed data and archaeological background; E.R. contributed data and performed analysis; M.F., J.W., H.M., and W.H. contributed data. All authors edited and approved the final version.

4.8 Competing Interests

The authors declare no competing interests.

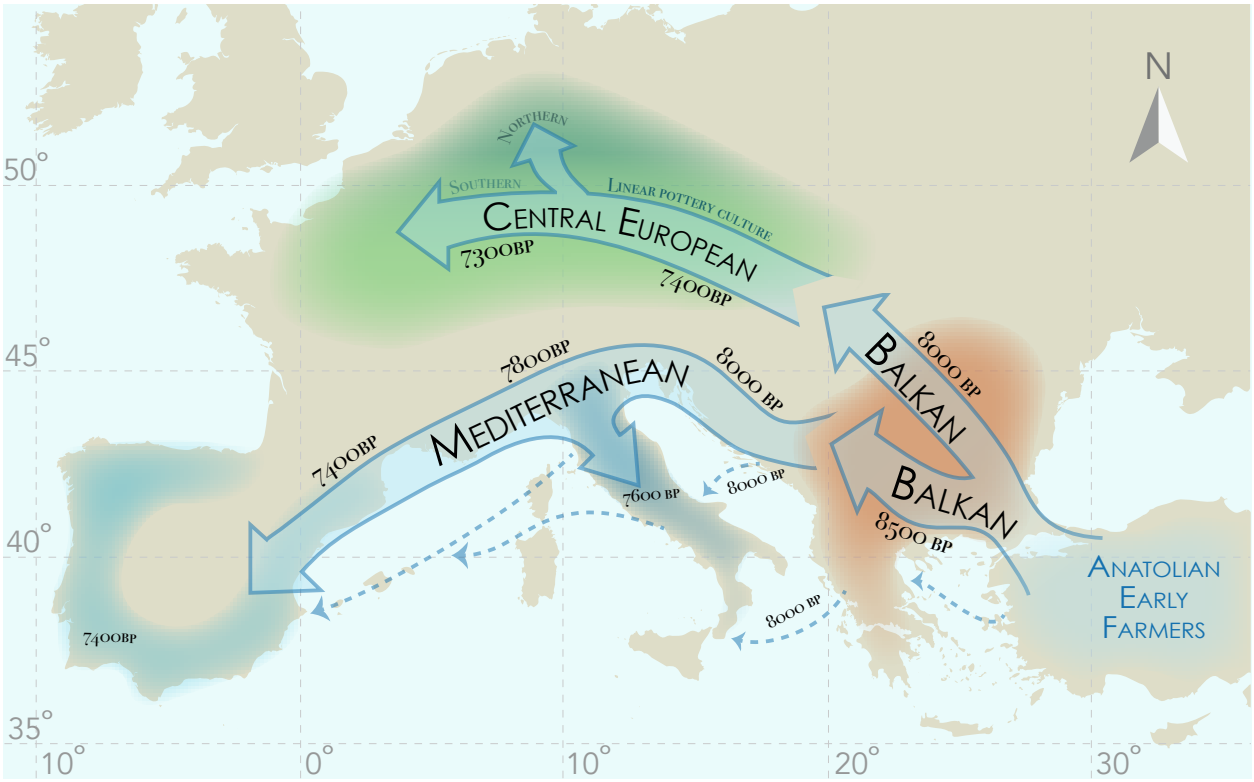


Figure 1: *Migrations of Early Farmers into Europe* This schematic illustration highlights the two main migration routes from Anatolia to Europe during the Early Neolithic period.[116] Populations followed two routes: southern, along the Mediterranean coast (including sea routes, generalized here by dashed blue lines) where they admixed with existing hunter-gatherer populations; or northern, through the Balkans and into Central Europe, with only limited hunter-gatherer admixture. We analyze patterns within the *Linearbandkeramik* culture, dividing it into Northern and Southern Central European groups. Vector map of the European continent from `stock.adobe.com`.

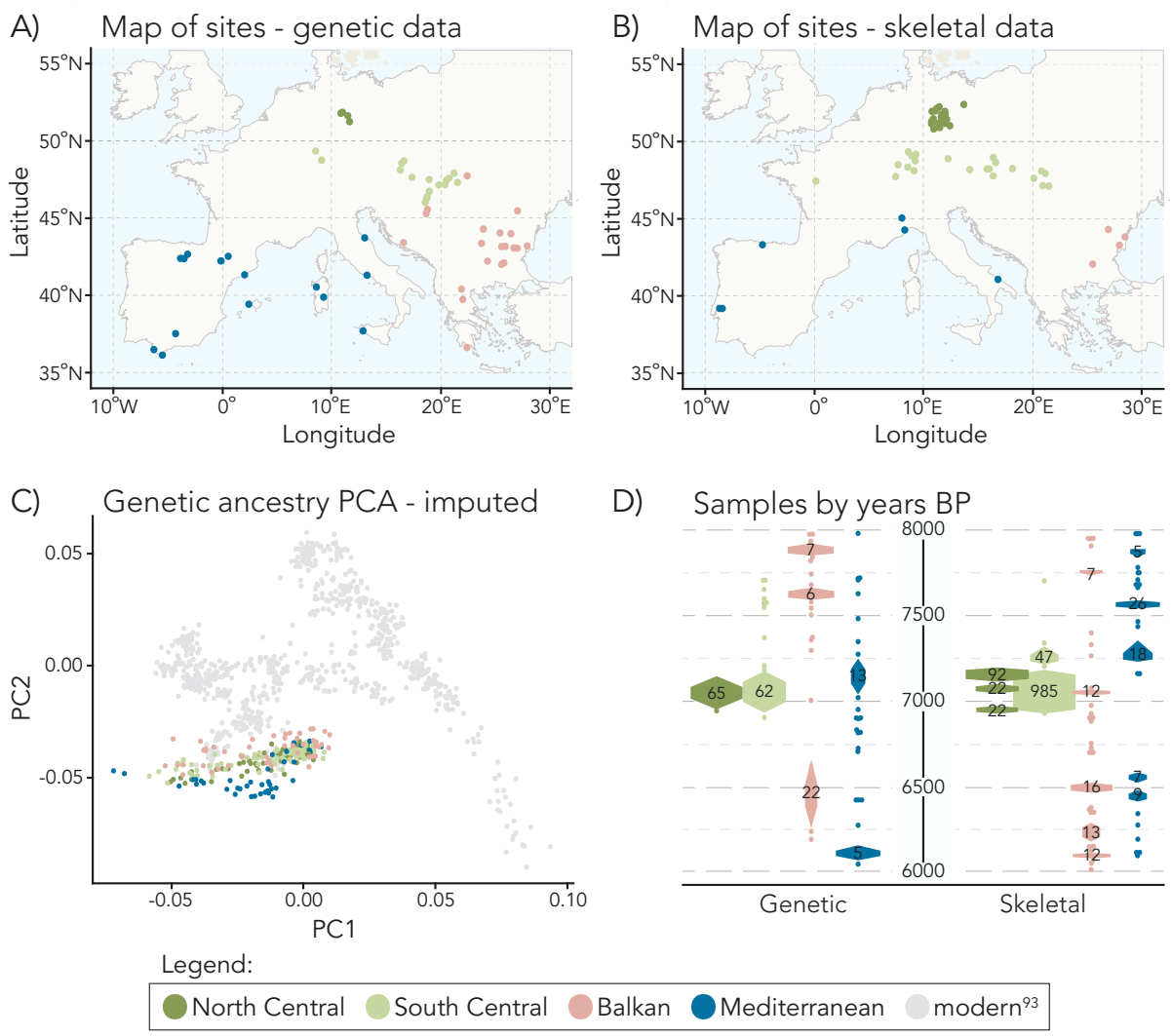


Figure 2: *Sample properties and distributions* Upper row: Sites used for genetic (a) and skeletal (b) data in the analysis. The Central European population is split into Northern and Southern groups at 50°N latitude (emphasized). Vector maps of the European continent from `stock.adobe.com`. Lower row: (c) imputed genetic data projected into the PCA space of 777 modern Eurasian individuals (grey points). (d) sample numbers by years before present (years BP) for skeletal (right) and genetic (left) data.

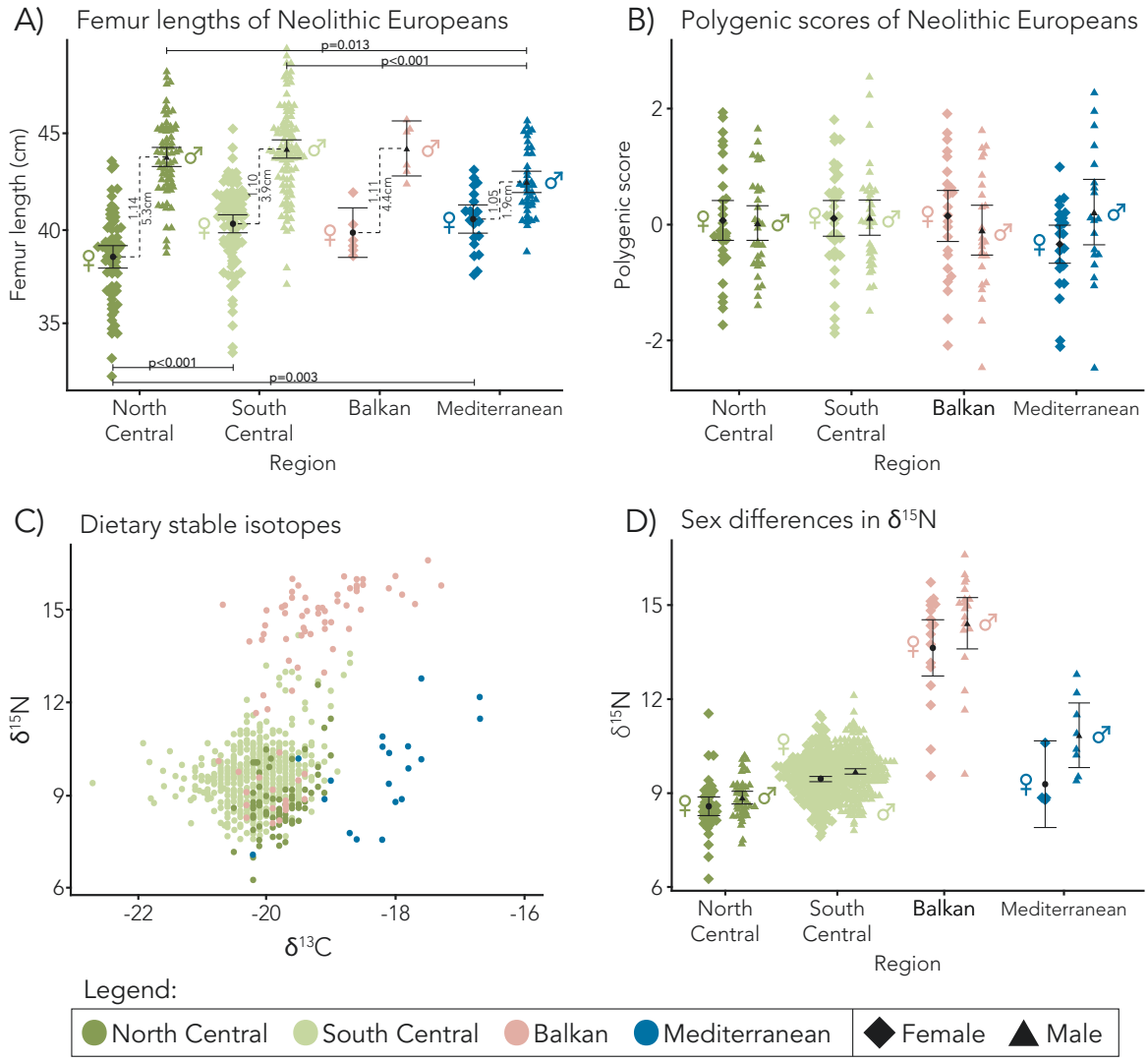


Figure 3: *Distributions of statures, polygenic scores, and isotopes* Solid bars across the tops of plots indicate significant differences between male populations by two-sided pairwise t-test adjusted for multiple comparisons; solid bars below plots indicate significant differences between female populations by two-way pairwise t-test; black points indicate the mean of each group; and vertical bars show 95% confidence intervals. a) Femur length in the four populations: values to the right of the vertical dotted line are the difference between the mean male and female femora; values to the left are the sexual dimorphism ratios of male/female femur lengths for each population. Sample sizes, left to right, males: 68, 100, 6, 40; females: 65, 88, 6, 20. a) Polygenic scores for the four populations show similar scores for individuals across all regions. Differences between male and female PRS are not significant. Sample sizes, left to right, males: 32, 35, 25, 19; females: 32, 37, 24, 23. a) Plot of δ^{13}C (x-axis) and δ^{15}N (y-axis) dietary stable isotopes for the four populations: individuals from the Balkans are distinguished by high nitrogen values, while those in the Mediterranean generally have higher carbon. North Central n=100, South Central n=666, Balkan n=78, Mediterranean n=19.d) Sex differences in δ^{15}N values by sex for each population: δ^{15}N values are slightly higher for males in all populations, but this difference is only significant in the Mediterranean (p=0.035). Sample sizes, left to right, males: 42, 254, 19, 8; females: 37, 249, 17, 4.

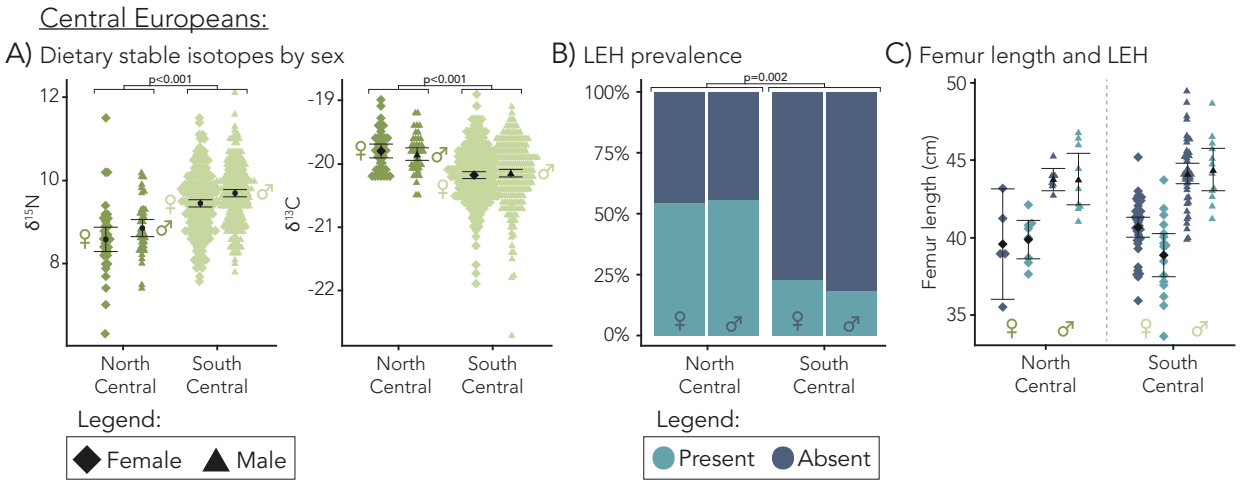


Figure 4: *Evidence of environmental stress in Northern Central Europe.* a) Differences in δ^{13}C (right) and δ^{15}N (left) values. Overall, the South has higher nitrogen values than the North and lower carbon; within each population, the difference in isotopes between sexes is not significant based on linear regression. Sample sizes, left to right, δ^{15}N: 37, 42, 249, 254; δ^{13}C: 37, 42, 250, 254. b) Proportion of linear enamel hypoplasias. The South has significantly less than the North (linear regression, p=0.002; beta=-1.386, se=0.444) and differences between sexes are not significant. c) Presence of linear enamel hypoplasias is significantly associated with shorter femora (linear regression, p=0.02, beta=-1.334, se=0.570); differences in prevalence between sexes are not significant. Sample sizes, left to right, North Central: 5, 8, 7, 9; South Central: 35, 16, 44, 12

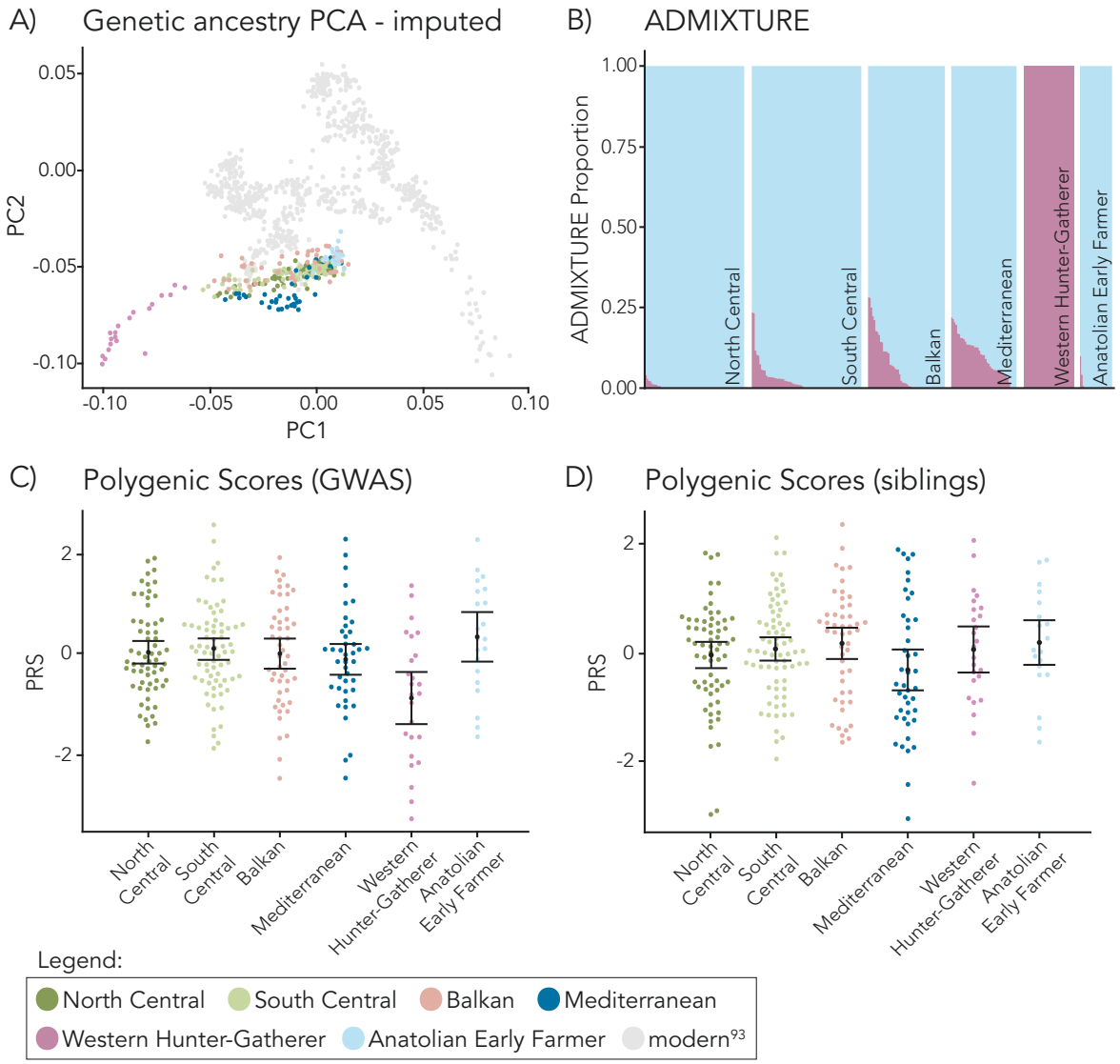


Figure 5: *Comparison with genetic source populations* a) Ancient individuals projected into modern PCA space, including those of Mesolithic Western Hunter-Gatherer (WHG) and Anatolian Early Farmer ancestry. b) ADMIXTURE plot of K=2 ancestry groups showing the increased proportion of WHG ancestry in Mediterranean individuals. c) Polygenic scores for each region including Mesolithic Hunter-Gatherers and Anatolian Early Farmers. Sample sizes left to right: 64, 72, 49, 42, 25, 21. d) Polygenic scores calculated from between-sibling summary statistics. Sample sizes left to right: 64, 72, 49, 42, 25, 21.

References

1. Harpak, A. & Przeworski, M. The evolution of group differences in changing environments. *PLoS Biol* **19**, e3001072. doi: 10.1371/journal.pbio.3001072 (2021).

2. Martin, A. R. *et al.* Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet* **51**, 584–591. doi: 10.1038/s41588–019–0379–x (2019).

3. Roseman, C. C. Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. *Proceedings of the National Academy of Sciences* **101**, 12824–12829. doi: 10.1073/pnas.0402637101 (2004).

4. Whitlock, M. C. Evolutionary inference from QST. *Mol Ecol* **17**, 1885–96. doi: 10.1111/j.1365–294X.2008.03712.x (2008).

5. Savell, K. R., Auerbach, B. M. & Roseman, C. C. Constraint, natural selection, and the evolution of human body form. *Proc Natl Acad Sci U S A* **113**, 9492–7. doi: 10.1073/pnas.1603632113 (2016).

6. Berg, J. J. & Coop, G. A population genetic signal of polygenic adaptation. *PLoS Genet* **10**, e1004412. doi: 10.1371/journal.pgen.1004412 (2014).

7. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503. doi: 10.1038/nature16152 (2015).

8. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in approximately 700000 individuals of European ancestry. *Hum Mol Genet* **27**, 3641–3649. doi: 10.1093/hmg/ddy271 (2018).

9. Cox, S. L. *et al.* Predicting skeletal stature using ancient DNA. *American Journal of Biological Anthropology* **177**, 162–174 (2022).

10. Marciniak, S. *et al.* An integrative skeletal and paleogenomic analysis of stature variation suggests relatively reduced health for early European farmers. *Proceedings of the National Academy of Sciences* **119**, e2106743119. doi: 10.1073/pnas.2106743119 (2022).

11. Cox, S. L., Ruff, C. B., Maier, R. M. & Mathieson, I. Genetic contributions to variation in human stature in prehistoric Europe. *Proceedings of the National Academy of Sciences* **116**, 21484–21492. doi: 10.1073/pnas.1910606116 (2019).

12. Bickle, P. & Whittle, A. (eds.) *The First Farmers of Central Europe* (Oxbow, Oxford, UK, 2013).

13. Bakels, C. C. *The Western European loess belt: agrarian history, 5300 BC-AD 1000* (Springer Science & Business Media, 2009).

14. Bentley, R. A. *et al.* Community differentiation and kinship among Europe's first farmers. *Proceedings of the National Academy of Sciences* **109**, 9326–9330. doi: 10.1073/pnas.1113710109 (2012).

15. Nicklisch, N., Oelze, V. M., Schierz, O., Meller, H. & Alt, K. W. A healthier smile in the past? dental caries and diet in early neolithic farming communities from central germany. *Nutrients* **14**, 183. doi: 10.3390/nu14091831 (2022).

16. Knipper, C. Kohlenstoff- und Stickstoff-Isotopenanalysen an bandkeramischen Bestattungen vom Viesenhäuser Hof bei Stuttgart-Mühlhausen: Implikationen zur Ernährungsrekonstruktion, Geschlechtsspezifik und Siedlungsdynamik. In Meyer, C., Held, P., Knipper, C. & Nicklisch, N. (eds.) *Der Zahn der Zeit – Mensch und Kultur im Spiegel interdisziplinärer Forschung. Festschrift für Kurt W. Alt* (Veröffentlichungen des LDA und Archäologie Sachsen-Anhalt, Halle(Saale), 2020).

17. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–11. doi: 10.1038/nature14317 (2015).

18. Lipson, M. *et al.* Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372. doi: 10.1038/nature24476 (2017).

19. Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197–203. doi: 10.1038/nature25778 (2018).

20. Antonio, M. L. *et al.* Ancient Rome: A genetic crossroads of Europe and the Mediterranean. *Science* **366**, 708–714. doi: 10.1126/science.aay6826 (2019).

21. Martiniano, R. *et al.* The population genomics of archaeological transition in west iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet* **13**, e1006852. doi: 10.1371/journal.pgen.1006852 (2017).

22. Howe, L. J. *et al.* Within-sibship genome-wide association analyses decrease bias in estimates of direct genetic effects. *Nature genetics* **54**, 581–592 (2022).

23. Sohail, M. *et al.* Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *Elife* **8**, e39702. doi: 10.7554/eLife.39702 (2019).

24. Berg, J. J. *et al.* Reduced signal for polygenic adaptation of height in UK Biobank. *Elife* **8**, e39725. doi: 10.7554/eLife.39725 (2019).

25. Vercellotti, G. *et al.* Exploring the multidimensionality of stature variation in the past through comparisons of archaeological and living populations. *American Journal of Physical Anthropology* **155**, 229–242 (2014).

26. Roseman, C. C. & Auerbach, B. M. Ecogeography, genetics, and the evolution of human body form. *J Hum Evol* **78**, 80–90. doi: 10.1016/j.jhevol.2014.07.006 (2015).

27. Stock, J. T. *et al.* Long-term trends in human body size track regional variation in subsistence transitions and growth acceleration linked to dairying. *Proceedings of the National Academy of Sciences* **120**, e2209482119 (2023).

28. Randall, J. C. *et al.* Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLOS Genetics* **9**, 1–19. doi: 10.1371/journal.pgen.1003500 (2013).

29. Bernabeu, E., Canela-Xandri, O., Rawlik, K. & et al. Sex differences in genetic architecture in the UK Biobank. *Nature Genetics* **53**, 1283–1289. doi: 10.1038/s41588–021–00912–0 (2021).

30. Zhu, C., Ming, M. J., Cole, J. M., Kirkpatrick, M. & Harpak, A. Amplification is the primary mode of gene-by-sex interaction in complex human traits. *bioRxiv* (2022). doi:10.1101/2022.05.06.490973.

31. Bogaard, A. *Neolithic farming in central Europe: an archaeobotanical study of crop husbandry practices* (Routledge, 2004).

32. Schier, W. Modes and Models of Neolithization in Europe: Comments to an Ongoing Debate. In Biehl, P. & Rosenstock, E. (eds.) *6000 BC: Transformation and Change in the Near East and Europe*, 372–392 (Cambridge University Press, Cambridge, UK, 2022). doi: 10.1017/9781107337640.026.

33. Betti, L. *et al.* Climate shaped how Neolithic farmers and European hunter-gatherers interacted after a major slowdown from 6,100 BCE to 4,500 BCE. *Nature Human Behaviour* **4**, 1004–1010 (2020).

34. Dürrwächter, C., Craig, O. E., Collins, M. J., Burger, J. & Alt, K. W. Beyond the grave: variability in Neolithic diets in Southern Germany. *Journal of Archaeological Science* **33**, 39–48 (2006).

35. Oelze, V. M. *et al.* Early Neolithic diet and animal husbandry: stable isotope evidence from three Linearbandkeramik (LBK) sites in Central Germany. *Journal of Archaeological Science* **38**, 270–279 (2011).

36. Bentley, R. A. *et al.* Baden-Württemberg. In Bickle, P. & Whittle, A. (eds.) *The First Farmers in Central Europe: Diversity in LBK Lifeways* (Oxbow Books, Oxford, UK, 2013).

37. Denaire, A. *et al.* The cultural project: Formal chronological modelling of the early and middle neolithic sequence in lower alsace. *Journal of Archaeological Method and Theory* **24**, 1072–1149 (2017).

38. Gillis, R. E. *et al.* Stable isotopic insights into crop cultivation, animal husbandry, and land use at the Linearbandkeramik site of Vráble-Vel'ké Lehemby (Slovakia). *Archaeological and Anthropological Sciences* **12**, 1–15 (2020).

39. Nicklisch, N. *et al.* Holes in teeth–dental caries in Neolithic and early bronze age populations in central Germany. *Annals of Anatomy-Anatomischer Anzeiger* **203**, 90–99 (2016).

40. Gillis, R. E. *et al.* The evolution of dual meat and milk cattle husbandry in Linearbandkeramik societies. *Proceedings of the Royal Society B: Biological Sciences* **284**, 20170905 (2017).

41. Salque, M. *et al.* Earliest evidence for cheese making in the sixth millennium BC in northern Europe. *Nature* **493**, 522–525 (2013).

42. Münster, A. *et al.* 4000 years of human dietary evolution in central Germany, from the first farmers to the first elites. *PloS one* **13**, e0194862 (2018).

43. Ghosh, S. Protein quality in the first thousand days of life. *Food and Nutrition Bulletin* **37**, S14–S21 (2016).

44. Gray, J. P. & Wolfe, L. D. Height and sexual dimorphism of stature among human societies. *American Journal of Physical Anthropology* **53**, 441–456 (1980).

45. Guatelli-Steinberg, D. & Lukacs, J. R. Interpreting sex differences in enamel hypoplasia in human and non-human primates: Developmental, environmental, and cultural considerations. *American journal of physical anthropology* **110**, 73–126 (1999).

46. Lubell, D., Jackes, M. & Meiklejohn, C. Archaeology and human biology of the Mesolithic-Neolithic transition in southern Portugal: a preliminary report. In *The Mesolithic in Europe. International Symposium. 3*, 632–640 (1990).

47. McFadden, C. & Oxenham, M. F. A paleoepidemiological approach to the osteological paradox: Investigating stress, frailty and resilience through cribra orbitalia. *American Journal of Physical Anthropology* **173**, 205–217 (2020).

48. German, A. & Hochberg, Z. Sexual dimorphism of size ontogeny and life history. *Frontiers in Pediatrics* **8** (2020).

49. Stini, W. Growth rates and sexual dimorphism in evolutionary perspective. In Stini, W. (ed.) *The Analysis of Prehistoric Diets* (Academic Press, Orlando, FL, 1985).

50. Brauer, G. Size sexual dimorphism and secular trend: indicators of subclinical malnutrition. In Hall, R. (ed.) *Sexual Dimorphism in Homo sapiens: A Question of Size* (Praeger, New York, NY, 1982).

51. Vlassoff, C. Gender differences in determinants and consequences of health and illness. *Journal of Health, Population, and Nutrition* **25**, 47–61 (2007).

52. Dehingia, N. & Raj, A. Sex differences in covid-19 case fatality: do we know enough? *The Lancet Global Health* **9**, e14–e15 (2021).

53. Takahashi, T. *et al.* Sex differences in immune responses that underlie covid-19 disease outcomes. *Nature* **588**, 315–320 (2020).

54. Frayer, D. W. & Wolpoff, M. H. Sexual dimorphism. *Annual Review of Anthropology* **14**, 429–473 (1985).

55. Gaulin, S. J. & Boster, J. S. Human marriage systems and sexual dimorphism in stature. *American journal of physical anthropology* **89**, 467–475 (1992).

56. Ruff, C. B. (ed.) *Skeletal variation and adaptation in Europeans: Upper Paleolithic to the twentieth century* (John Wiley & Sons, Hoboken, NJ, 2018). (data available at https://www.hopkinsmedicine.org/fae/CBR.html).

57. Kanwar, S. C., Sabharwal, A., Bhadra, K. & Narang, A. Nationwide reference data for height, weight and body mass index of Indian schoolchildren. *Natlonal Medical Journal of India* **24**, 269–77 (2011).

58. Abdulrazzaq, Y., Moussa, M. & Nagelkerke, N. National growth charts for the United Arab Emirates. *Journal of Epidemiology* **18**, 295–303 (2008).

59. Moradi, A. & Guntupalli, A. M. What does gender dimorphism in stature tell us about discrimination in rural India, 1930-1975. *Gender Bias: Health, Nutrition and Work* (2009).

60. Kanazawa, S. & Novak, D. L. Human sexual dimorphism in size may be triggered by environmental cues. *Journal of Biosocial Science* **37**, 657–665 (2005).

61. Baten, J. & Murray, J. E. Heights of men and women in 19th-century Bavaria: economic, nutritional, and disease influences. *Explorations in Economic History* **37**, 351–369 (2000).

62. Castellucci, H. *et al.* Gender inequality and sexual height dimorphism in Chile. *Journal of Biosocial Science* **53**, 38–54 (2021).

63. Jayachandran, S. & Pande, R. Why are indian children so short? the role of birth order and son preference. *American Economic Review* **107**, 2600–2629 (2017).

64. Bentley, R. A. *et al.* Prehistoric migration in Europe: strontium isotope analysis of early Neolithic skeletons. *Current Anthropology* **43**, 799–804 (2002).

65. Bickle, P. Thinking gender differently: New approaches to identity difference in the central european neolithic. *Cambridge Archaeological Journal* **30**, 201–218 (2020).

66. Garvin, H. M. *The effects of living conditions on human cranial and postcranial sexual dimorphism*. Ph.D. thesis, The Johns Hopkins University (2012).

67. Papathanasiou, A. Health, diet and social implications in Neolithic Greece from the study of human osteological material. *Human bioarchaeology of the transition to agriculture* 87–106 (2011).

68. Papathanasiou, A. Health status of the Neolithic population of Alepotrypa Cave, Greece. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists* **126**, 377–390 (2005).

69. Silva, A. M. & Cunha, E. Paleopathological study of the community exhumed from the Hipogeu of Monte Canelas I (Alcalar, Portugal). In *Actas del V Congreso Nacional de Paleopatologia*, 353–356 (2001).

70. Cucina, A. Brief communication: diachronic investigation of linear enamel hypoplasia in prehistoric skeletal samples from Trentino, Italy. *American Journal of Physical Anthropology* **119**, 283–287 (2002).

71. Brunel, S. *et al.* Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proceedings of the National Academy of Sciences* **117**, 12791–12798 (2020).

72. Childebayeva, A. *et al.* Population genetics and signatures of selection in Early Neolithic European farmers. *Molecular Biology and Evolution* **39**, msac108 (2022).

73. Fernandes, D. M. *et al.* The spread of steppe and Iranian-related ancestry in the islands of the western Mediterranean. *Nature Ecology & Evolution* **4**, 334–345 (2020).

74. Fregel, R. *et al.* Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proceedings of the National Academy of Sciences* **115**, 6774–6779 (2018).

75. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205. doi: 10.1038/nature17993 (2016).

76. Gamba, C. *et al.* Genome flux and stasis in a five millennium transect of European prehistory. *Nature communications* **5**, 1–9 (2014).

77. González-Fortes, G. *et al.* Paleogenomic evidence for multi-generational mixing between Neolithic farmers and Mesolithic hunter-gatherers in the Lower Danube Basin. *Current Biology* **27**, 1801–1810 (2017).

78. Hofmanová, Z. *et al.* Early farmers from across Europe directly descended from Neolithic Aegeans. *Proceedings of the National Academy of Sciences* **113**, 6886–6891 (2016).

79. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature Communications* **6**, 1–8 (2015).

80. Lazaridis, I. *et al.* Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218 (2017).

81. Marcus, J. H. *et al.* Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nature Communications* **11**, 1–14 (2020).

82. Nikitin, A., Stadler, P. & Kotova, e. a., N. Interactions between earliest Linearbandkeramik farmers and central European hunter gatherers at the dawn of European Neolithization. *Scientific Reports* **9**, 19544. doi: 10.1038/s41598–019–56029–2 (2019).

83. Olalde, I. *et al.* A common genetic origin for early farmers from Mediterranean Cardial and Central European LBK cultures. *Molecular Biology and Evolution* **32**, 3132–3142 (2015).

84. Olalde, I. *et al.* The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).

85. Rivollat, M. *et al.* Ancient genome-wide DNA from France highlights the complexity of interactions between Mesolithic hunter-gatherers and Neolithic farmers. *Science Advances* **6**, eaaz5344 (2020).

86. Valdiosera, C. *et al.* Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proceedings of the National Academy of Sciences* **115**, 3428–3433 (2018).

87. Villalba-Mouco, V. *et al.* Survival of Late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. *Current Biology* **29**, 1169–1177 (2019).

88. Neale Lab. UK Biobank GWAS Results. http://www.nealelab.is/uk-biobank (2018). Accessed: 2021-10-20.

89. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7. doi: 10.1186/s13742–015–0047–8 (2015).

90. Privé, F., Arbel, J. & Vilhjálmsson, B. J. LDpred2: better, faster, stronger. *Bioinformatics* **36**, 5424–5431 (2020).

91. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genetics* **2**, e190. doi: 10.1371/journal.pgen.0020190 (2006).

92. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day europeans. *Nature* **513**, 409–13. doi: 10.1038/nature13673 (2014).

93. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research* **19**, 1655–64. doi: 10.1101/gr.094052.109 (2009).

94. Hujić, A. *Das Kind in uns unter der Lupe der Isotopie, Allometrie und Pathologie. Zusammenhang zwischen $\delta^{15}N$ und $\delta^{13}C$ als Eiweißproxy und dem Knochenwachstum bei linienbandkeramischen Individuen aus Stuttgart-Mühlhausen, "Viesenhäuser Hof" und Schwetzingen unter Berücksichtigung verschiedener Indikatoren für Nährstoffversorgung.* Ph.D. thesis, Freie Universität (2016). doi: 10.17169/refubium-5564.

95. Nicklisch, N. *Spurensuche am Skelett: Paläodemografische und epidemiologische Untersuchungen an neolithischen und frühbronzezeitlichen Bestattungen aus dem Mittelelbe-Saale-Gebiet im Kontext populationsdynamischer Prozesse* (Landesamt für Denkmalpflege und Archäologie Sachsen-Anhalt, Landesmuseum, 2017).

96. Meyer, C. *et al.* Early Neolithic executions indicated by clustered cranial trauma in the mass grave of Halberstadt. *Nature Communications* **9**, 1–11 (2018).

97. Rosenstock, E., Ebert, J., Martin, R. & Groß, M. LiVES COstA Digest: checked and updated entries from the LiVES collection of Osteological Anthropometry. doi: 10.17171/2-12-2-1 (2019).

98. Rosenstock, E. *et al.* Human stature in the near east and europe ca. 10,000–1000 bc: its spatiotemporal development in a bayesian errors-in-variables model. *Archaeological and Anthropological Sciences* **11**, 5657–5690. doi: 10.1007/s12520–019–00850–3 (2019).

99. Ruff, C. B. *et al.* Stature and body mass estimation from skeletal remains in the european holocene. *American journal of physical anthropology* **148**, 601–617 (2012).

100. Acsádi, G., Nemeskéri, J. & Balás, K. *History of human life span and mortality* (Akademiai kiado Budapest, 1970).

101. Krishan, K. *et al.* A review of sex estimation techniques during examination of skeletal remains in forensic anthropology casework. *Forensic Science International* **261**, 165–e1 (2016).

102. Ash, A. *et al.* Regional differences in health, diet and weaning patterns amongst the first Neolithic farmers of central Europe. *Scientific Reports* **6**, 1–10 (2016).

103. Lillie, M. Vedrovice: demography and palaeopathology in an early farming population. *Anthropologie* **46**, 135–152 (2008).

104. Whittle, A. *et al.* Hungary. In Bickle, P. & Whittle, A. (eds.) *The First Farmers in Central Europe: Diversity in LBK Lifeways* (Oxbow Books, Oxford, UK, 2013).

105. Whittle, A. *et al.* Moravia and western Slovakia. In Bickle, P. & Whittle, A. (eds.) *The First Farmers in Central Europe: Diversity in LBK Lifeways* (Oxbow Books, Oxford, UK, 2013).

106. Brickley, M. B. Cribra orbitalia and porotic hyperostosis: A biological approach to diagnosis. *American Journal of Physical Anthropology* **167**, 896–902 (2018).

107. Scheibner, A. *Prähistorische Ernährung in Vorderasien und Europa. Eine kulturgeschichtliche Synthese auf der Basis ausgewählter Quellen. Schriften zum Lebensstandard in der Vorgeschichte 1* (Berliner Archäologische Forschungen, Leidorf: Rahden, 2016).

108. O'Brien, D. M. Stable Isotope Ratios as Biomarkers of Diet for Health Research. *Annu Rev Nutr* **35**, 565–594 (2015).

109. Richards, M. P., Montgomery, J., Nehlich, O. & Grimes, V. Isotopic analysis of humans and animals from Vedrovice. *Anthropologie* **46**, 185–194 (2008).

110. Bickle, P. *et al.* Austria. In Bickle, P. & Whittle, A. (eds.) *The First Farmers in Central Europe: Diversity in LBK Lifeways* (Oxbow Books, Oxford, UK, 2013).

111. Bickle, P. *et al.* Alsace. In Bickle, P. & Whittle, A. (eds.) *The First Farmers in Central Europe: Diversity in LBK Lifeways* (Oxbow Books, Oxford, UK, 2013).

112. Hofmann, D. *et al.* Southern Bavaria. In Bickle, P. & Whittle, A. (eds.) *The First Farmers in Central Europe: Diversity in LBK Lifeways* (Oxbow Books, Oxford, UK, 2013).

113. Morey, R. D. & Rouder, J. N. *BayesFactor: Computation of Bayes Factors for Common Designs* (2022). R package version 0.9.12-4.4.

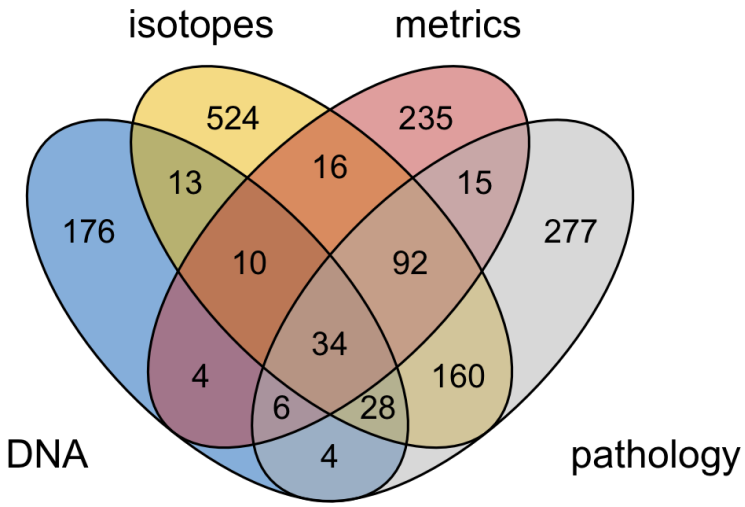
114. Rouder, J. N. & Morey, R. D. Default bayes factors for model selection in regression. *Multivariate Behavioral Research* **47**, 877–903 (2012).

115. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2021). https://www.R-project.org/.

116. Gronenborn, D., Horejs, B., Börner & Ober. Map: Expansion of farming in western Eurasia, 9600 - 4000 cal BC (update vers. 2021.2) (RGZM/ÖAI). online (2021). `https://www.academia.edu/9424525/Map_Expansion_of_farming_in_western_Eurasia_9600_4000_cal_BC_update_vers_2021_2_`.

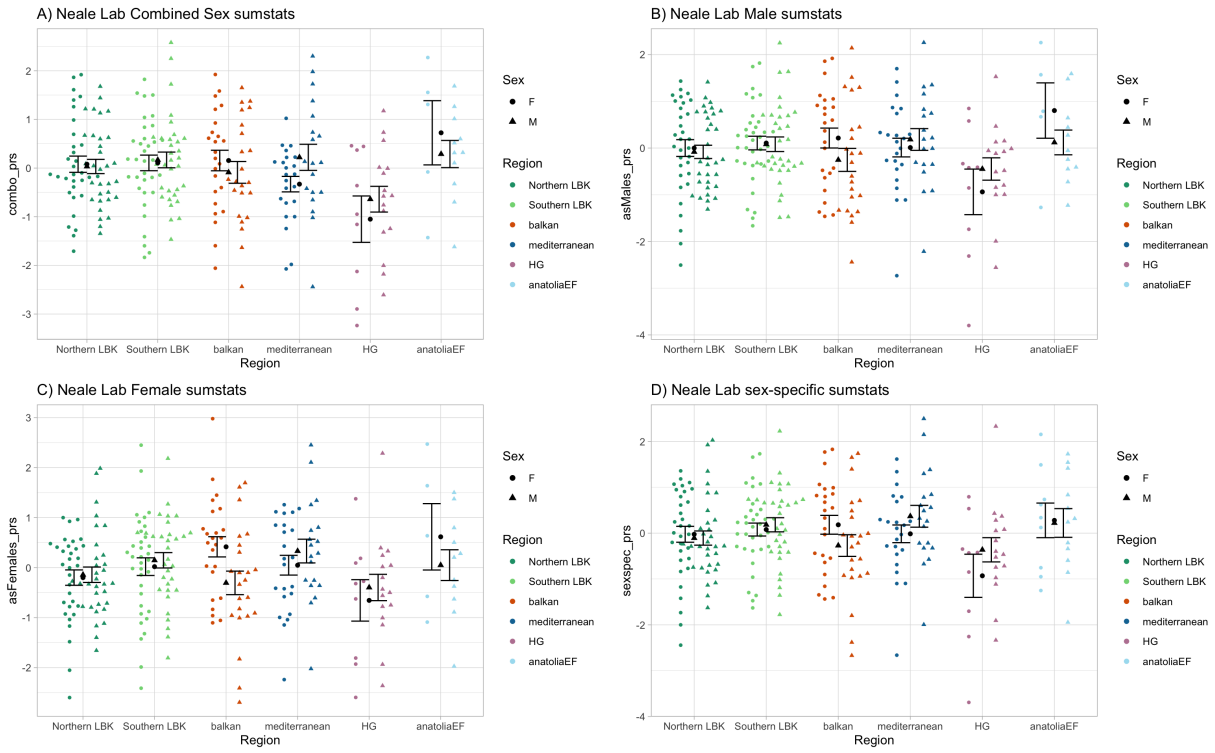
List of Supplementary Figures

Supp.Fig. 1: Data set counts



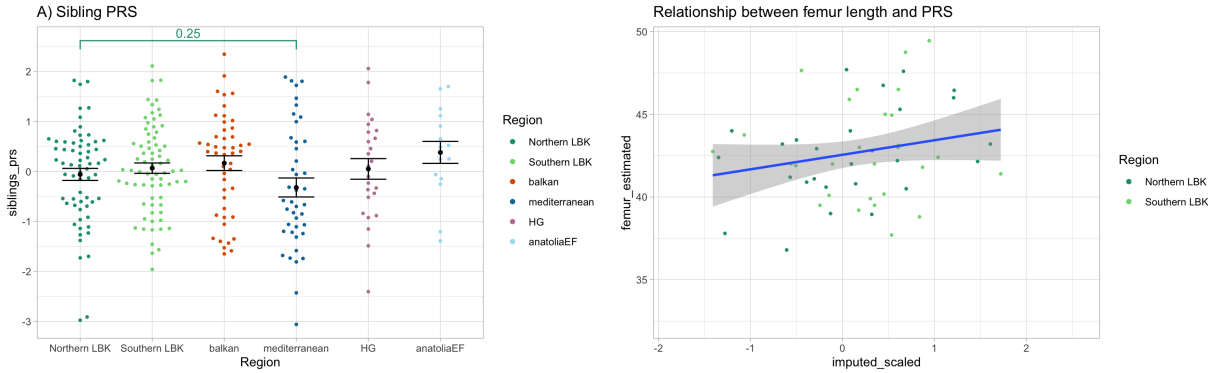
Supplementary Figure 1: Samples by data type

Venn diagram illustrating the overlaps in data. All individuals have at least one type of data, but only subsets have more than one type.



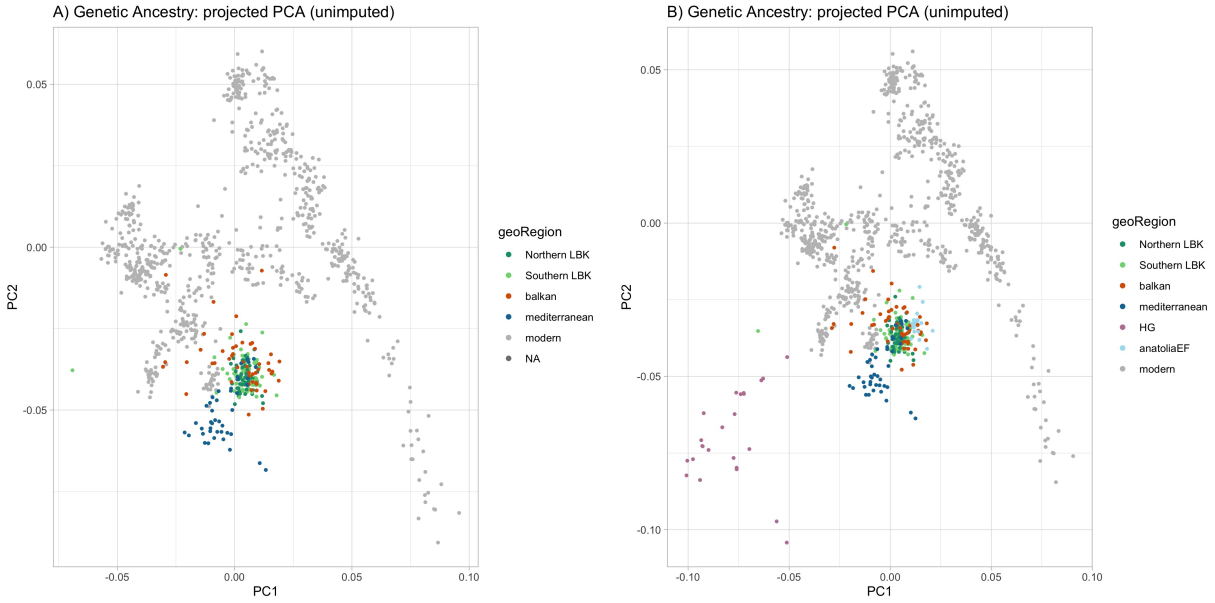
Supplementary Figure 2: Sex-specific clumping/thresholding PRS.

In each plot, black points are the group mean and black bars are the 95% confidence intervals. Differences between groups was assessed based on two-tailed pairwise t-tests adjusted for multiple comparisons; sample sizes are the same for all plots, from left to right, females: 32, 37, 24, 23, 9, 9; males: 32, 35, 25, 19, 16, 12. A) PRS calculated using the combined sex summary statistics from the Neale Lab. There are no significant differences. B) PRS calculated for all individuals using the male-specific summary statistics from the Neale Lab. There are no significant differences. C) PRS calculated for all individuals using the female-specific summary statistics from the Neale Lab. There are no significant differences. D) PRS calculated for males and females separately using their respective summary statistics. There are no significant differences. (Summary stats: Neale Lab (2018))



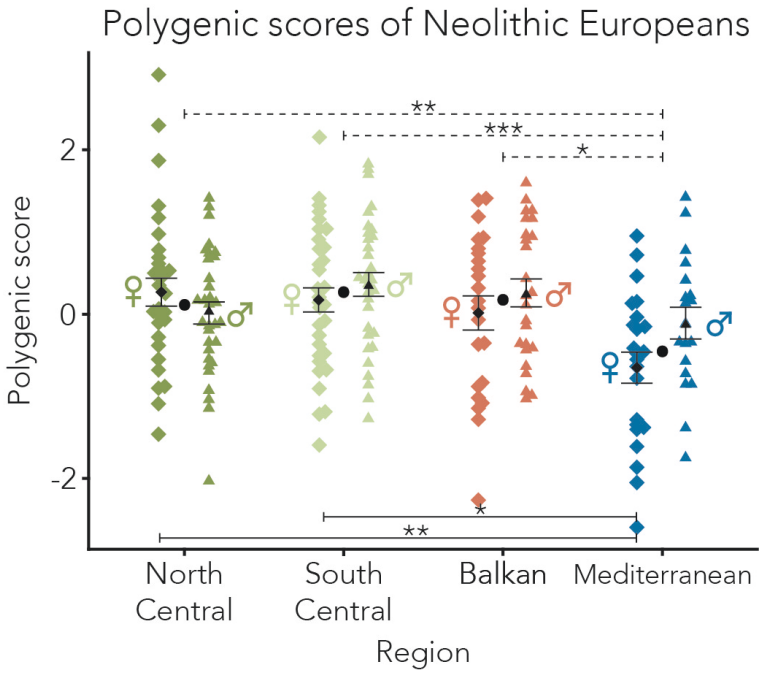
Supplementary Figure 3: Sibling-based PRS and femoral lengths

A) PRS calculated with between-sibling effect sizes using the LDpred2 model. Based on two-tailed pairwise t-tests adjusted for multiple comparisons, there are no differences between groups. Black points indicate group means. Black bars show the 95% confidence intervals. left to right, n=64, 72, 49, 42, 25, 21. B) Femur length (y-axis) increases with clumping/thresholding PRS (x-axis) though the effect is not significant likely due to the small sample size (n=55). Blue line is the regression line and grey area is the 95% confidence interval.



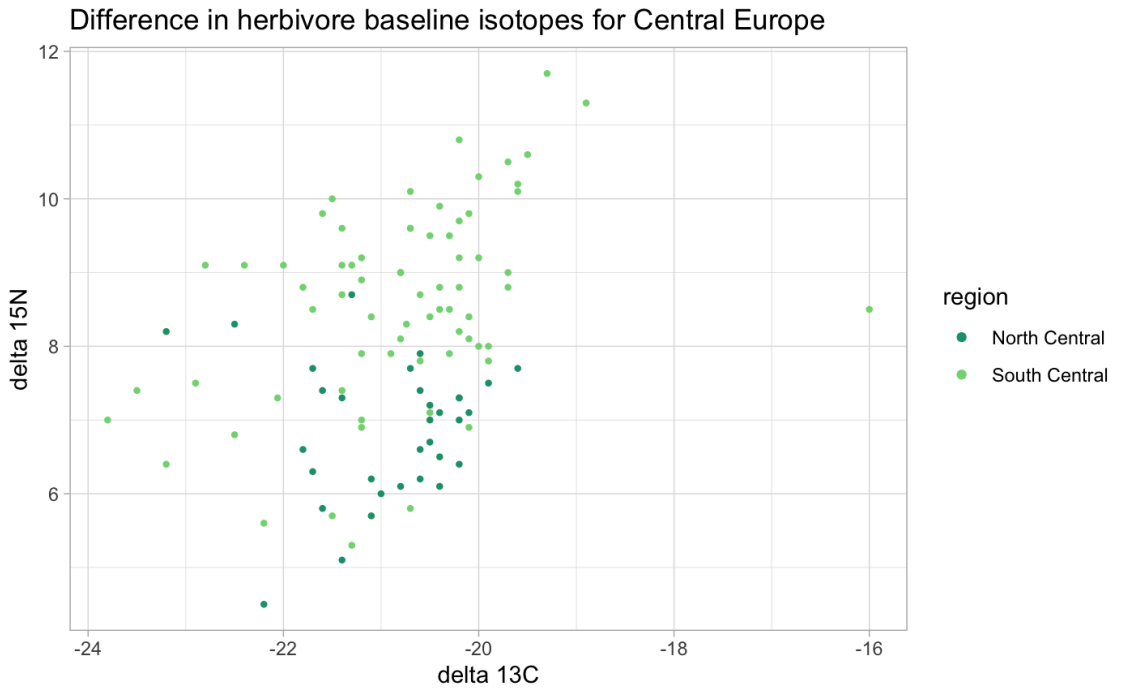
Supplementary Figure 4: PCA for unimputed DNA.

A) Unimputed PCA for Neolithic Europe, equivalent to Figure 2C in the main text. B) Unimputed PCA for all samples, equivalent to Figure 5A in the main text.

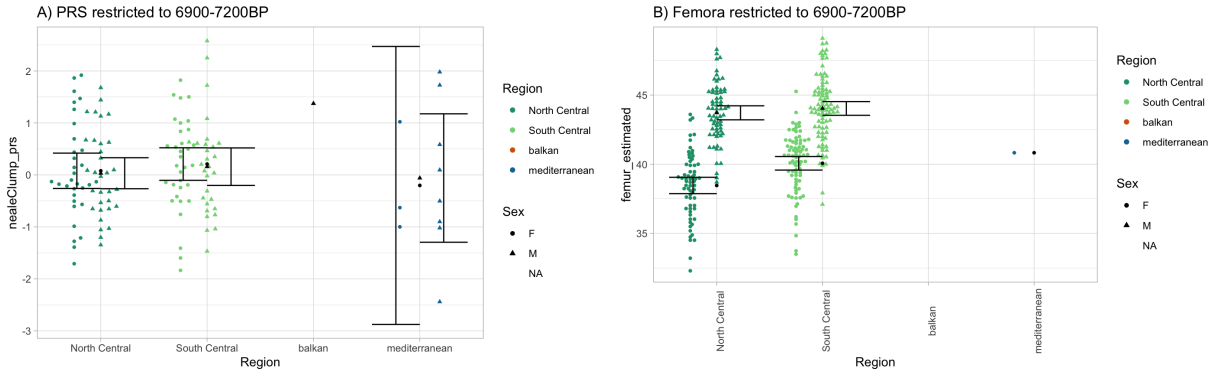


Supplementary Figure 5: LDpred2 PRS.

Polygenic scores calculated using infinitessimal LDpred 2 model (Privé, 2020). There is no difference between North Central, South Central or Balkans populations. Mediterraneans PRS are significantly lower than the other groups (linear regression, p=0.002). Similarly, the difference between males and females is only significant in the Mediterranean (two-sided pairwise t-test, adjusted for multiple comparisons, p=0.04). Sample sizes, left to right, males: 32, 35, 25, 19; females: 32, 37, 24, 23. Black points are group means, black bars represent 95% confidence intervals.



Supplementary Figure 6: Herbivore baseline.
Isotopic data from herbivores to establish a baseline for interpretation of human isotopic results. See citation in main text for data sources.



Supplementary Figure 7: Date restricted plots.
Black points are group means, black bars represent 95% confidence intervals. A) Plot of PRS data restricted to the same narrow date range used in the Central European population samples and illustrating the subsequent loss of power to detect effects. Sample sizes left to right, females: 32, 37, 0, 3; males: 32, 35, 1, 8. B) Plot of femora data restricted to the same date range as used in the Central European sample and illustrating the subsequent loss of power to detect effects. Sample sizes, left to right, females: 65, 88, 0, 1; males: 68, 100, 0, 0.