Genetic contributions to variation in human stature in prehistoric Europe

Samantha L. Cox[a], Christopher B. Ruff[b], Robert M. Maier[c,d,e], and Iain Mathieson[a,1]

[a]Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104; [b]Center for Functional Anatomy and Evolution, Johns Hopkins University School of Medicine, Baltimore, MD 21205; [c]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA 02142; [d]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142; and [e]Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA 02114

The relative contributions of genetics and environment to temporal and geographic variation in human height remain largely unknown. Ancient DNA has identified changes in genetic ancestry over time, but it is not clear whether those changes in ancestry are associated with changes in height. Here, we directly test whether changes over the past 38,000 y in European height predicted using DNA from 1,071 ancient individuals are consistent with changes observed in 1,159 skeletal remains from comparable populations. We show that the observed decrease in height between the Early Upper Paleolithic and the Mesolithic is qualitatively predicted by genetics. Similarly, both skeletal and genetic height remained constant between the Mesolithic and Neolithic and increased between the Neolithic and Bronze Age. Sitting height changes much less than standing height— consistent with genetic predictions—although genetics predicts a small post-Neolithic increase that is not observed in skeletal remains. Geographic variation in stature is also qualitatively consistent with genetic predictions, particularly with respect to latitude. Finally, we hypothesize that an observed decrease in genetic heel bone mineral density in the Neolithic reflects adaptation to the decreased mobility indicated by decreased femoral bending strength. This study provides a model for interpreting phenotypic changes predicted from ancient DNA and demonstrates how they can be combined with phenotypic measurements to understand the relative contribution of genetic and developmentally plastic responses to environmental change.

stature | height | ancient DNA | evolution

Stature, or standing height, is one of the most heavily studied human phenotypes. It is easy to measure in living individuals and relatively straightforward to estimate from skeletal remains. As a consequence, geographic variation and temporal changes in stature are well documented (1–3), particularly in western Europe, where there is a comprehensive record of prehistoric changes (4). The earliest anatomically modern humans in Europe, present by 42,000 to 45,000 y before present (BP) (5, 6), were relatively tall (mean adult male height in the Early Upper Paleolithic was ∼174 cm). Mean male stature then declined from the Paleolithic to the Mesolithic (∼164 cm) before increasing to ∼167 cm by the Bronze Age (4, 7). Height can respond rapidly in a developmentally plastic manner to changes in environment, as demonstrated by large increases in Europe, and worldwide, during the secular trends of the 19th and 20th centuries (1, 4). In European countries today, mean adult male height is ∼170 to 180 cm (1). It is broadly agreed that prehistoric changes were likely to have been driven by a combination of environmental (e.g., climate or diet) and genetic factors including drift, admixture, and selection (4, 7–9), although the effects of these variables cannot be separated based on skeletal data alone. In this study, by combining the results of genome-wide association studies (GWAS) with ancient DNA (aDNA), we directly estimate the genetic component of stature and test whether population-level skeletal changes between ∼35,000 and 1,000 BP are consistent with those predicted by genetics.

Height is highly heritable (10–14) and therefore amenable to genetic analysis by GWAS. With sample sizes of hundreds of thousands of individuals, GWAS have identified thousands of genomic variants that are significantly associated with the phenotype (15–17). Although the individual effect of each of these variants is tiny [on the order of ±1 to 2 mm per variant (18)], their combination can be highly predictive. Polygenic risk scores (PRS) constructed by summing together the effects of all height-associated variants carried by an individual can now explain upwards of 30% of the phenotypic variance in populations of European ancestry (16). In effect, the PRS can be thought of as an estimate of "genetic height" that predicts phenotypic height, at least in populations closely related to those in which the GWAS was performed. One major caveat is that the predictive power of PRS is much lower in other populations (19). The extent to which differences in PRS between populations are predictive of population-level differences in phenotype is currently unclear (20). Recent studies have demonstrated that such differences may partly be artifacts of correlation between environmental and genetic structure in the original GWAS (21, 22). These studies also suggested best practices for PRS comparisons, including the use of GWAS summary statistics from large homogenous studies (instead of metaanalyses), and replication of results using summary statistics derived from within-family analyses that are robust to population stratification.

Bearing these caveats in mind, PRS can be applied to ancient populations thanks to recent technological developments that have dramatically increased aDNA sample sizes. These have provided remarkable insights into the demographic and evolutionary history of both modern and archaic humans across the world (23–25),

Significance

Measurements of prehistoric human skeletal remains provide a record of changes in height and other anthropometric traits over time. Often, these changes are interpreted in terms of plastic developmental response to shifts in diet, climate, or other environmental factors. These changes can also be genetic in origin, but, until recently, it has been impossible to separate the effects of genetics and environment. Here, we use ancient DNA to directly estimate genetic changes in phenotypes and to identify changes driven not by genetics, but by environment. We show that changes over the past 35,000 y are largely predicted by genetics but also identify specific shifts that are more likely to be environmentally driven.

particularly in Europe, and allow us to track the evolution of variants underlying phenotypes ranging from pigmentation to diet (26–29). In principle, PRS applied to ancient populations could similarly allow us to make inferences about the evolution of complex traits. A few studies have used PRS to make predictions about the relative statures of ancient populations (29–31) but looked at only a few hundred individuals in total and did not compare their predictions with stature measured from skeletons. Here, we compare measured skeletal data to genetic predictions and directly investigate the genetic contribution to height independent of environmental effects acting during development.

Results

PRS and Skeletal Measurements. We collected published aDNA data from 1,071 ancient individuals from Western Eurasia (west of 50° E), dated to between 38,000 and 1100 BP (27, 29, 30, 32–57). Using GWAS summary statistics for height from the UK Biobank (generated and made available by the Neale Laboratory: http://www.nealelab.is/), we computed height PRS for each individual, using a P value cutoff of 10^{-6}, clumping variants in 250-kb windows, and replacing missing genotypes with the mean across individuals (*Methods*). We refer to this as PRS(GWAS). Because of concerns about GWAS effect sizes being inflated by residual population stratification, we also computed a PRS where we used GWAS P values to select single-nucleotide polymorphisms (SNPs), but computed the PRS using effect sizes estimated from a within-family test of ~17,000 sibling pairs from UK Biobank (*Methods*), which we refer to as PRS(GWAS/Sibs) and which should be unaffected by stratification. We also obtained stature estimates from 1,159 individuals dating to between 33,700 and 1100 BP taken from a larger dataset of 2,177 individuals with stature and body proportion estimates from substantially complete skeletons (4, 58). There is limited overlap in these datasets (12 individuals), but they cover the same time periods and broadly the same geographic locations (*SI Appendix*, Fig. S1), although the genetic data contain more individuals from further east (30 to 50° E) compared to the skeletal data. We divided these individuals into 5 groups based on date:

Early Upper Paleolithic (>25,000 BP) (EUP), Late Upper Paleolithic (25,000 to 11,000 BP) (LUP), Mesolithic (11,000 to 5500 BP), Neolithic (8500 to 3900 BP), and post-Neolithic (5000 to 1100 BP, including the Copper and Bronze Ages, plus later periods), resolving individuals in the overlapping periods using either archaeological or genetic context (*Methods*). These groups broadly correspond to transitions in both archaeological culture and genetic ancestry (33, 38, 59) (*SI Appendix*, Fig. S1 *C* and *D* and Table S1).

Trends in PRS for Height Are Largely Consistent with Trends in Skeletal Stature. Both PRS and skeletal stature decreased from the EUP to Mesolithic periods and increased between the Neolithic and post-Neolithic (*SI Appendix*, Fig. S2). Fitting group (time period) as a covariate, we found a significant effect on PRS(GWAS) (ANOVA $P = 1.9 \times 10^{-9}$), PRS(GWAS/Sibs) ($P = 0.045$), and skeletal stature ($P = 2.8 \times 10^{-11}$). There was no evidence of difference between LUP, Mesolithic, and Neolithic groups (*SI Appendix*, Fig. S3 *A* and *B*), so we merged these 3 groups (we refer to the merged group as LUP-Neolithic). We find that PRS(GWAS) in the LUP-Neolithic period is 0.47 SD lower than in the EUP ($P = 0.002$), and 0.40 SD lower ($P = 8.7 \times 10^{-11}$) than in the post-Neolithic period (Fig. 1A). PRS(GWAS/Sib) shows a very similar pattern (Fig. 1B), demonstrating that this is not a result of differential relatedness of the ancient individuals to the structured present-day GWAS populations. Skeletal stature shows a qualitatively similar pattern to the genetic predictions (Fig. 1C), with a 1.5 SD (9.6 cm; $P = 2.9 \times 10^{-7}$) difference between EUP and LUP-Neolithic and a 0.27 SD (1.8 cm; $P = 3.6 \times 10^{-5}$) difference between LUP-Neolithic and post-Neolithic. Broad patterns of change in stature over time are therefore consistent with genetic predictions.

Additionally, we fit a piecewise linear model allowing PRS to decrease from the EUP to the Neolithic and then increase and change slope in the post-Neolithic (Fig. 1 *D–F*). In this model, PRS(GWAS) decreases by about 1.8×10^{-5} SD/y ($P = 0.014$) from EUP to Neolithic, and increases by 2.0×10^{-4} SD/y ($P = 0.001$) post-Neolithic (Fig. 1D). PRS(GWAS/sib) decreases by about

Fig. 1. Changes in height PRS and stature through time. Each point is an ancient individual, white lines show fitted values, gray area is the 95% confidence interval, and boxes show parameter estimates and P values for difference in means (δ) and slopes (β). (*A–C*) PRS(GWAS) (*A*), PRS(GWAS/Sibs) (*B*), and skeletal stature (*C*) with constant values in the EUP, LUP-Neolithic, and post-Neolithic. (*D–F*) PRS(GWAS) (*D*), PRS(GWAS/Sibs) (*E*), and skeletal stature (*F*) showing a linear trend between EUP and Neolithic and a different trend in the post-Neolithic.

1.6×10^{-5} SD/y ($P = 0.037$) from EUP to Neolithic, and then increases by 1.6×10^{-4} SD/y throughout the post-Neolithic ($P = 0.011$; Fig. 1E). Again, these changes are qualitatively consistent with changes in stature (Fig. 1F), with a 4.7×10^{-5} SD/y (3.3×10^{-4} cm/y; $P = 2.4 \times 10^{-8}$) decrease from EUP to Mesolithic, and an increase of ~0.5 SD into the Neolithic. However, in this model, stature, unlike PRS, actually decreases during the post-Neolithic period (7.5×10^{-4} cm/y; $P = 2.0 \times 10^{-4}$).

To further explore these trends, we fitted a broader range of piecewise linear models to both datasets (*Methods* and *SI Appendix*, Figs. S4–S6). In the most general model, we allowed both the mean and the slope of PRS or stature with respect to time to vary between groups. More constrained models fix some of these parameters to zero—eliminating change over time—or merging 2 adjacent groups. We compared the fit of these nested models using Akaike's information criterion (*SI Appendix*, Table S2). The linear model in Fig. 1 *D–F* is one of the best models in this analysis. In general, all of the best-fitting models support the pattern—for both PRS and measured stature—of a decrease between the EUP and Mesolithic and an increase between the Neolithic and post-Neolithic (*SI Appendix*, Figs. S4–S6). Some models suggest that the increase in stature—but not PRS—may have started during the Neolithic (*SI Appendix*, Fig. S6 *A–C*). Finally, we confirmed that these results were robust to different constructions of the PRS—using 100- and 500-kb clustering windows rather than 250 kb (*SI Appendix*, Figs. S7 and S8).

Sitting Height PRS Is Partially Consistent with Trends in Body Proportions. Standing height is made up of 2 components: leg length and sitting height (made up of the length of the trunk, neck, and head), with a partially overlapping genetic basis (60). Throughout European prehistory, changes in leg length tended to be larger than changes in sitting height (4). We constructed PRS(GWAS) and PRS(GWAS/Sibs) for sitting height and analyzed them in the same way as standing height (Fig. 2). In contrast to standing height, we find no evidence of change between the EUP and Neolithic. Both PRS(GWAS) and PRS(GWAS/Sibs)

do increase, either between the Neolithic and post-Neolithic, or during the post-Neolithic period (Fig. 2 *A, B, D,* and *E*). On the other hand, using only skeletons with complete torsos to estimate sitting height, we find no evidence of change in any period. Thus, the skeletal data are consistent with the genetic data for the EUP-Neolithic period, but inconsistent in the post-Neolithic period, where PRS predicts an increase that is not reflected in the skeletons. This could be because of more limited skeletal measurements (only 236 out of 1,159 skeletons are sufficiently complete to estimate sitting height directly), because the change in PRS is artifactual, it is being buffered by nongenetic effects, or by opposing genetic effects that we do not capture. Overall, we find mixed consistency between PRS and skeletal measurements. The decrease in standing but not sitting height between the EUP and Neolithic is consistent in both, as is the increase in standing height between the Neolithic and post-Neolithic. However, PRS predicts a continued increase in stature through the post-Neolithic period that is not seen in skeletal remains.

Geographic Variation in Standing Height. As well as varying through time, human stature is stratified by geography, with trends related to both longitude and latitude (61). North-to-south trends following Allen's (62) and Bergmann's (63) rules are most often interpreted as environmental adaptations to the polar-to-equatorial climate gradient. Today, Northern Europeans are generally taller than Southern Europeans (1), a pattern that emerged between the Mesolithic and post-Neolithic (4, 7). Longitudinal variation within Europe is present during the Mesolithic (64), although these trends are difficult to interpret due to sampling bias across the time period (4). We therefore tested whether geographic variation in PRS could explain these geographic trends, as it partially explains temporal trends.

We regressed the residuals from our fitted linear height model (the model shown in Fig. 1 *D–F*) on longitude and latitude. Stature increases significantly with latitude ($P = 1.2 \times 10^{-10}$) in the post-Neolithic period. PRS(GWAS) increases in the post-Neolithic ($P = 0.006$), although this is not replicated by PRS(GWAS/Sibs)
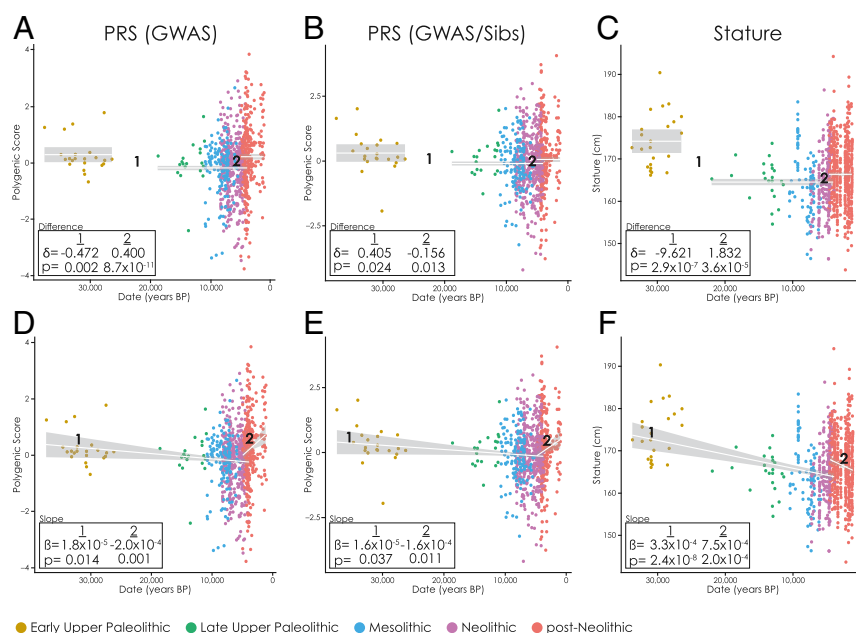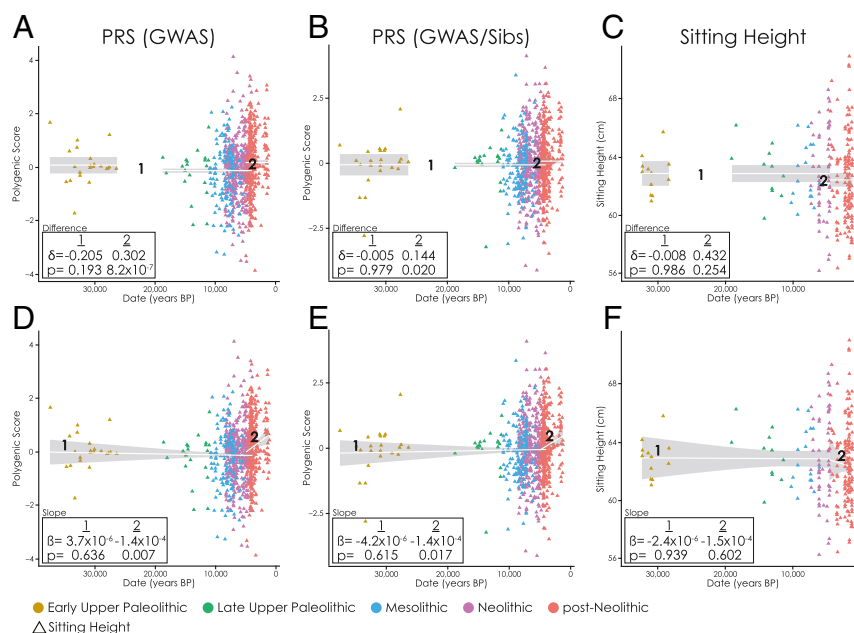

Fig. 2. Changes in sitting-height PRS and sitting height through time. Each point is an ancient individual, lines show fitted values, gray area is the 95% confidence interval, and boxes show parameter estimates and *P* values for difference in means (δ) and slopes (β). (*A–C*) PRS(GWAS) (*A*), PRS(GWAS/Sibs) (*B*), and skeletal sitting height (*C*), with constant values in the EUP, LUP-Neolithic, and post-Neolithic. (*D–F*) PRS(GWAS) (*D*), PRS(GWAS/Sibs) (*E*), and skeletal sitting height (*F*) showing a linear trend between EUP and Neolithic and a different trend in the post-Neolithic.
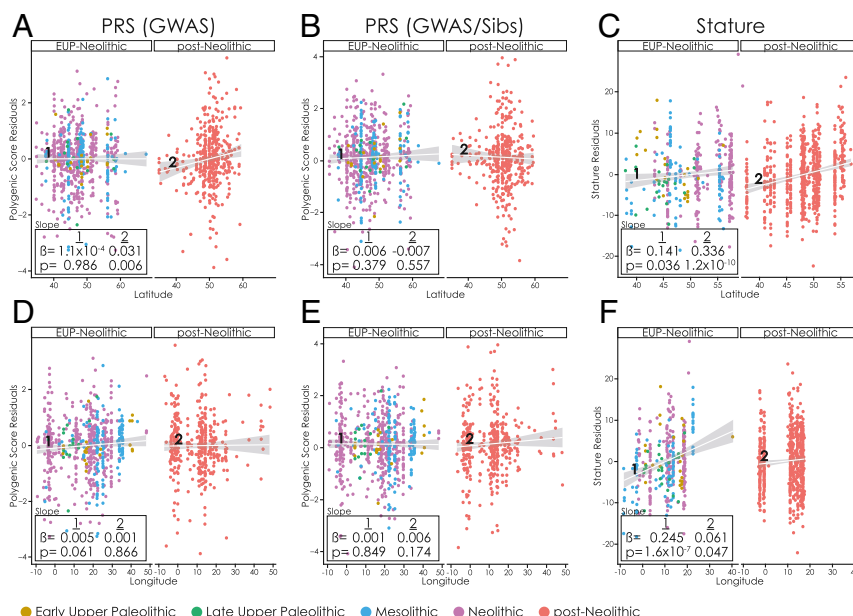
Fig. 3. Geographic variation in PRS and skeletal standing height. Residuals for the linear height model (Fig. 1 *D–F*) against (*A–C*) latitude and (*D–F*) longitude. Each point is an ancient individual, lines show fitted values, gray area is the 95% confidence interval, and boxes show parameter estimates (β) and *P* values for slopes.

($P = 0.557$). PRS does not increase significantly with latitude in the EUP-Neolithic period. There is some evidence of a modest trend in stature in the EUP-Neolithic period (Fig. 3*C*). However, there is only evidence for this in the Neolithic, not in the EUP-Mesolithic (*SI Appendix,* Fig. S9*A*). Furthermore, because time and geography are correlated in our Neolithic sample, this can also be explained by a temporal increase during the Neolithic, in which case there is no geographic trend (*SI Appendix,* Fig. S9*B*).

In contrast to latitude, there is a significant increasing trend of stature with longitude before but not during the Neolithic (0.36 cm/degree, $P = 1.6 \times 10^{-7}$; Fig. 3 and *SI Appendix,* Fig. S9*C*). This is partly driven by a small number of samples from a single site, but still persists if these samples are removed (0.20 cm/degree, $P = 0.004$; *SI Appendix,* Fig. S9*D*). There is little or no trend (0.06 cm/degree; $P = 0.047$) in the post-Neolithic period (Fig. 3*F*). We find no evidence for longitudinal clines in PRS. In summary, we find that stature increases with latitude in the post-Neolithic, possibly in the Neolithic, but not before. This cline may have a genetic basis. Stature also increases with longitude, particularly in the Mesolithic, but this cline is not predicted by genetics.

Correlated Changes in Bone Density PRS and Femoral Bending Strength. Beyond stature, we wanted to investigate the utility of using PRS to interpret other measurable phenotypes in ancient individuals. Decreased mobility through time, associated with large-scale lifestyle transitions between hunting-gathering, agriculture, and ultimately modern industrialism, is well documented through declines in lower limb bone diaphyseal strength and trabecular density (4, 65, 66). Today, heel bone mineral density (hBMD) is often used as an indicator of general activity levels in younger people (67) and of osteoporosis in older individuals (68, 69); UK Biobank has GWAS data for this trait, indirectly estimated by ultrasound. However, evaluating differences in BMD in archaeological and paleontological specimens can be problematic. In the short term, soil leaches bone minerals, while later the bone begins to fossilize, leading to unpredictable patterns of density in ancient remains (70) and requiring special processing methods (65) that are difficult to apply to large samples. However, femoral diaphyseal bending strength can be calculated from bone

cross-sectional geometric measurements that are not as affected by bone preservation (71). Here, we focus on anteroposterior bending strength (section modulus) of the midshaft femur (FZx), which has been linked specifically to mobility (72). Since both trabecular density and diaphyseal strength should respond to mobility and activity levels, we reasoned that they would be likely to show correlated patterns of temporal change. Following established protocols (71), we standardized FZx first by sex, then the product of estimated body mass and femoral length (4).

Qualitatively, PRS(GWAS) and FZx show similar patterns, decreasing through time (Fig. 4 and *SI Appendix,* Figs. S2 and S3). There is a significant drop in FZx (Fig. 4*C*) from the Mesolithic to Neolithic ($P = 1.2 \times 10^{-8}$), and again from the Neolithic to post-Neolithic ($P = 1.5 \times 10^{-13}$). PRS(GWAS) for hBMD decreases significantly from the Mesolithic to Neolithic (Fig. 4*A*; $P = 5.5 \times 10^{-12}$), which is replicated in PRS(GWAS/Sibs) ($P = 7.2 \times 10^{-10}$; Fig. 4*B*); neither PRS shows evidence of decrease between the Neolithic and post-Neolithic. We hypothesize that both FZx and hBMD responded to the reduction in mobility that accompanied the adoption of agriculture (72). In particular, the lower genetic hBMD and skeletal FZx of Neolithic compared to Mesolithic populations may represent adaptation to the same change in environment, although we do not know the extent to which the change in FZx was driven by genetic or plastic developmental response to environmental change. On the other hand, FZx continues to decrease between the Neolithic and post-Neolithic (Fig. 4 *C* and *F*)—which is not reflected in the hBMD PRS (Fig. 4 *A*, *B*, *D*, and *E*). One possibility is that the 2 phenotypes responded differently to the post-Neolithic intensification of agriculture. Another is that the nongenetic component of hBMD, which we do not capture here, also continued to decrease.

Are Changes in PRS Driven by Selection or Genetic Drift? The Q_x statistic (73) can be used to test for polygenic selection. We computed it for increasing numbers of SNPs from each PRS (Fig. 5 *A–C*), between each pair of adjacent time periods and over all time periods. We estimated empirical *P* values by replacing allele frequencies with random derived allele frequency-matched SNPs

ANTHROPOLOGY

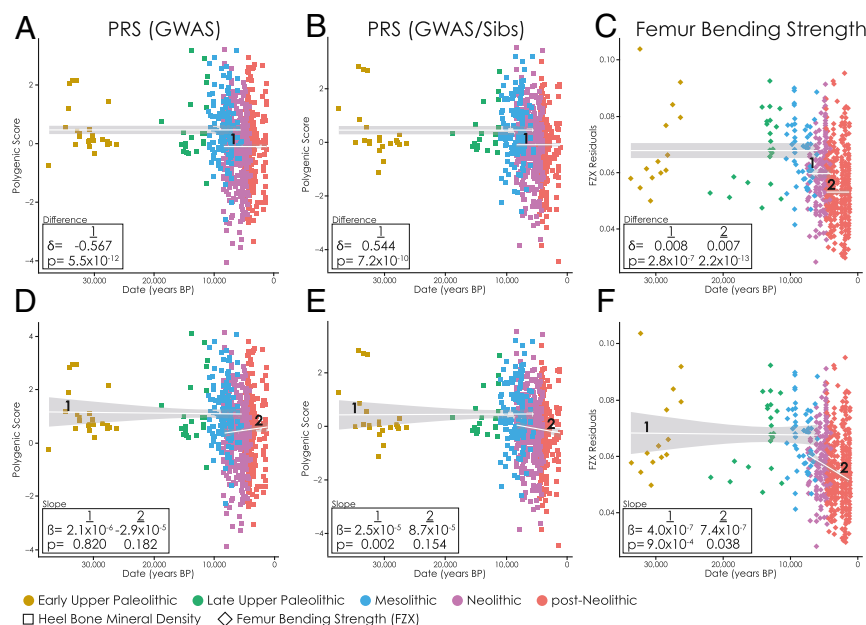**Fig. 4.** Changes in heel bone mineral density (hBMD) PRS and femur bending strength (FZx) through time. Each point is an ancient individual, lines show fitted values, gray area is the 95% confidence interval, and boxes show parameter estimates and P values for difference in means (δ) and slopes (β). (*A* and *B*) PRS(GWAS) (*A*) and PRS(GWAS/Sibs) (*B*) for hBMD, with constant values in the EUP-Mesolithic and Neolithic–post-Neolithic. (*C*) FZx constant in the EUP-Mesolithic, Neolithic, and post-Neolithic. (*D* and *E*) PRS(GWAS) (*D*) and PRS(GWAS/Sibs) (*E*) for hBMD showing a linear trend between EUP and Mesolithic and a different trend in the Neolithic–post-Neolithic. (*F*) FZx with a linear trend between EUP and Mesolithic and a different trend in the Neolithic–post-Neolithic.

from across the genome, while keeping the same effect sizes. To check these Q_x results, we simulated a GWAS from the UK Biobank dataset (*Methods*), and then used these effect sizes to compute simulated Q_x statistics. The Q_x test suggests selection between the Neolithic and Post-Neolithic for stature ($P < 1 \times 10^{-4}$; Fig. 5*A*), which replicates using effect sizes estimated within siblings ($10^{-4} < P < 10^{-2}$; *SI Appendix*, Fig. S10). The reduction in the sibling effect compared to the GWAS effect sizes is consistent with the reduction expected from the lower sample size (*SI Appendix*, Fig. S10). However, several (11/200) simulated datasets produce higher Q_x values than observed in the real data (Fig. 5*D*). This suggests that reestimating effect sizes between siblings may not fully control for the effect of population structure and ascertainment bias on the Q_x test. The question of whether selection contributes to the observed differences in height PRS remains unresolved.

For sitting height, we find little evidence of selection in any time period ($P > 10^{-2}$). We conclude that there was most likely selection for increased standing but not sitting height in the Steppe ancestors of Bronze Age European populations, as previously proposed (29). One potential caveat is that, although we reestimated effect sizes within siblings, we still used the GWAS results to identify SNPs to include. This may introduce some subtle confounding, which remains a question for future investigation. Finally, using GWAS effect sizes, we identify some evidence of selection on hBMD when comparing Mesolithic and Neolithic populations ($10^{-3} < P < 10^{-2}$; Fig. 5*C*). However, this signal is relatively weak when using within-sibling effect sizes and disappears when we include more than about 2,000 SNPs.

Discussion

We showed that the well-documented temporal and geographic trends in stature in Europe between the EUP and the post-Neolithic period are broadly consistent with those that would be predicted by PRS computed using present-day GWAS results combined with aDNA. However, because of the limited predictive power of current PRS, we cannot provide a quantitative estimate

of how much of the variation in phenotype between populations might be explained by variation in PRS. Similarly, we cannot say whether the changes were continuous, reflecting evolution through time, or discrete, reflecting changes associated with known episodes of replacement or admixture of populations that have diverged genetically over time. Finally, we find cases where predicted genetic changes are discordant with observed phenotypic changes—emphasizing the role of developmental plasticity in response to environmental change and the difficulty in interpreting differences in PRS in the absence of phenotypic data.

Our results indicate 2 major episodes of change in genetic height. First, there was a reduction in standing-height PRS—but not sitting-height PRS—between the EUP and LUP, coinciding with a substantial population replacement (33). These genetic changes are consistent with the decrease in stature—driven by leg length—observed in skeletons during this time period (4, 64, 74, 75). One possibility is that the stature decrease in the ancestors of the LUP populations could have been adaptive, driven by changes in resource availability (76) or to a colder climate (61). Comparison between patterns of phenotypic and genetic variation suggest that, on a broad scale, variation in body proportions among present-day people reflects adaptation to environment largely along latitudinal gradients (77, 78). EUP populations in Europe would have migrated relatively recently from more southern latitudes and had body proportions that are typical of present-day tropical populations (75). The populations that replaced them would have had more time to adapt to the colder climate of northern latitudes. On the other hand, we do not find genetic evidence for selection on stature during this time period—suggesting that the changes could have been neutral and not adaptive.

The second episode of change in height is either between the Neolithic and post-Neolithic, or during the post-Neolithic period. This period is characterized by the eastward movement of substantial amounts of "Steppe ancestry" into Central and Western Europe (27, 30, 38, 50). Our results are thus consistent with previous results that migration and admixture from Bronze Age

Fig. 5. Signals of selection on standing height, sitting height, and bone mineral density. (*A*–*C*) –$\log_{10}$ bootstrap P values for the Q_x statistics (*y* axis, capped at 4) for GWAS signals. We tested each pair of adjacent populations, and the combination of all of them ("All"). We ordered PRS SNPs by increasing P value and tested the significance of Q_x for increasing numbers of SNPs (*x* axis). (*D*) Distribution of Q_x statistics in simulated data (*Methods*). Observed height values for 6,800 SNPs shown by vertical lines.

populations of the Eurasian steppe increased genetic height in Europe (29, 30). Whether this increase was driven by selection in the ancestors of these populations remains unresolved. The geographic gradient of increasing skeletal stature is unclear in the Paleolithic, largely west to east in the Mesolithic (7, 64) and largely south to north by the Bronze Age (4, 7, 9). Latitudinal, but not longitudinal, patterns are qualitatively consistent with geographic patterns in PRS, suggesting that, like temporal variation, both genetics and environment contribute to geographic variation.

A major confounding factor in analysis of temporal and geographic variation in PRS, particularly in the Bronze Age, is genetic population structure. Present-day European population structure is correlated with geography and largely driven by variation in Steppe ancestry proportion, with more Steppe ancestry in Northern Europe and less in Southern Europe (38). Suppose that environmental variation in stature is also correlated with geography, and that Northern Europeans are taller than Southern Europeans for entirely nongenetic reasons. Then, GWAS that do not completely correct for stratification will find that genetic variants that are more common in Steppe populations than Neolithic populations are associated with increased height. When these GWAS results are used to compute PRS for ancient populations, they will predict that Steppe ancestry populations were genetically taller simply because they are more closely related to present-day Northern Europeans (21, 22). In this study, we attempted to avoid this confounding in 2 ways: first, by using GWAS effect sizes from the UK Biobank—a homogenous dataset that should be well controlled for population stratification, and second, by replicating our results after reestimating the effect sizes within siblings, which should be robust to population stratification, although less precise. However, we cannot exclude the possibility that some confounding remains, for example, because although we re-estimated effect sizes using the within-siblings design, we still ascertained loci using the GWAS results. A related concern is that, even in the absence of bias, variance explained by the PRS is likely to decrease as we move back in time and ancient populations become less closely

related to present-day populations (19). However, our results indicate that the PRS still captures enough of the genetic variance to predict changes in phenotype even though, presumably, we would gain even more resolution using a PRS that captured more of the genetic variation in the phenotype.

As well as genetic contributions to phenotype, our results shed light on possible environmental contributions. In some cases, we can make hypotheses about the relationship between environmental or lifestyle changes, and genetic change. For example, if we interpret change in femur bending strength as reflecting a decrease in mobility, the coincident Mesolithic/Neolithic change in hBMD PRS can be seen as a genetic response to this change. However, in the Neolithic/post-Neolithic periods, the 2 observations are decoupled. This emphasizes the role of developmental plasticity in response to changes in environment, and of joint interpretation of phenotypic and genetic variables. Even when looking at the same phenotype, we find cases where genetic predictions and phenotypic data are discordant—for example, in post-Neolithic sitting height. We must therefore be cautious in the interpretation of predicted genetic patterns where phenotypes cannot be directly measured, even if it is possible to control stratification. Predicted genetic changes should be used as a baseline, against which nongenetic effects can be measured and tested.

Methods

aDNA and Polygenic Risk Score Construction. We collected published aDNA data from 1,071 ancient individuals, taken from 29 publications. The majority of these individuals had been genotyped using an in-solution capture reagent ("1240k") that targets 1.24 million SNPs across the genome. Because of the low coverage of most of these samples, the genotype data are pseudohaploid. That is, there is only a single allele present for each individual at each site, but alleles at adjacent sites may come from either of the 2 chromosomes of the individual. For individuals with shotgun sequence data, we selected a single read at each 1240k site. We obtained the date of each individual from the original publication. Most of the samples have been directly radiocarbon dated, or else are securely dated by context. We summarized the genetic relationships between ancient and present-day groups by computing F_{ST}

ANTHROPOLOGY

using smartpca v16000 (79) (*SI Appendix*, Table S1) and multidimensional scaling using pairwise distances computed using plink v1.90b5.3 (options—distance flat-missing 1-ibs) (80) (*SI Appendix*, Fig. S1*C*) and unsupervised ADMXITURE (81) (*SI Appendix*, Fig. S1*D*).

We obtained GWAS results from the Neale Laboratory UK Biobank page (http://www.nealelab.is/uk-biobank/; round 1, accessed February and April 2018). To compute PRS, we first took the intersection of the 1240k sites and the association summary statistics. We then selected a list of SNPs to use in the PRS by selecting the SNP with the lowest P value, removing all SNPs within 250 kb, and repeating until there were no SNPs remaining with P value less than 10^{-6}. We then computed PRS for each individual by taking the sum of genotype multiplied by effect size for all included SNPs. Where an individual was missing data at a particular SNP, we replaced the SNP with the average frequency of the SNP across the whole dataset. This has the effect of shrinking the PRS toward the mean and should be conservative for the identification of differences in PRS. We confirmed that there was no correlation between missingness and PRS, to make sure that missing data did not bias the results (correlation between missingness and PRS, $\rho = 0.02$; $P = 0.44$, *SI Appendix*, Fig. S11). Finally, we normalized the PRS across individuals to have mean 0 and SD 1.

We estimated within-family effect sizes from 17,358 sibling pairs in the UK Biobank to obtain effect estimates that are unaffected by stratification. Pairs of individuals were identified as siblings if estimates of IBS0 were greater than 0.0018 and kinship coefficients were greater than 0.185. Of those pairs, we only retained those where both siblings were classified by UK Biobank as "white British," and randomly picked 2 individuals from families with more than 2 siblings. We used Hail (82) to estimate within-sibling pair effect sizes for 1,284,881 SNPs by regressing pairwise phenotypic differences between siblings against the difference in genotype. We included pairwise differences of sex (coded as 0/1) and age as covariates, and inverse-rank–normalized the phenotype before taking the differences between siblings. To combine the GWAS and sibling results, we first restricted the GWAS results to sites where we had estimated a sibling effect size and replaced the GWAS effect sizes by the sibling effects. We then restricted to 1240k sites and constructed PRS in the same way as for the GWAS results.

To test whether the differences in the GWAS and GWAS/Sibs PRS results can be explained by differences in power, we created subsampled GWAS estimates that matched the sibling in the expected SEs, by determining the equivalent sample size necessary and randomly sampling N_{sub} individuals. $N_{sub} = N_{sib}/(2var(\delta_{sib}))$, where δ_{sib} is the difference in normalized phenotype between siblings after accounting for the covariates age and sex.

Stature Data. We obtained stature data from Ruff (2018) (4) (data file and notes available at https://www.hopkinsmedicine.org/fae/CBR.html), which also includes estimated body mass, femoral midshaft anteroposterior strength (FZx), and other osteometric dimensions. Statures and body masses were calculated from linear skeletal measurements using anatomical reconstruction or sample-specific regression formulae (4, 58). We calculated sitting height as basion–bregma (cranial) height plus vertebral column length. Analysis was restricted to 1,159 individuals dated earlier than 1165 BP (651 males and 508 females), of which 1,130 had estimates for stature, 1,014 for FZx, and 236 for sitting height. Sitting and standing height were standardized for sex by adding the mean difference between male and female estimates to all of the female values. Sex differences in stature remain relatively constant over time (4), making it reasonable to adjust all female heights by the same mean value. Male/female counts in each period were as follows: EUP, 13/10; LUP, 15/8; Mesolithic, 56/37; Neolithic, 130/91; and post-Neolithic, 437/362. For FZx, we first standardized for sex as we did for stature, then divided each by estimated body mass multiplied by biomechanical femur length (4).

Grouping. We grouped individuals into broad categories based on date and, in some cases, archeological and genetic context. All individuals were assigned to one time period group, based on median age estimates of the sample obtained from the original publications. Date ranges for each time period are based on a combination of historical, climatic, and archaeological factors. The EUP comprises all samples older than 25,000 BP, which roughly coincides with the end of the Last Glacial Maximum. The LUP begins when the European glaciers are beginning to recede (25,000 BP) and extends until 11,000 BP and a shift in lithic technology that is traditionally used to delineate the beginning of the Mesolithic period. Transitions between the Mesolithic, Neolithic, and Bronze Age are staggered throughout Europe, so creating universally applicable date ranges is not possible. We instead defined overlapping transition periods between the Mesolithic and Neolithic periods (8500 to 5500 BP) and between the Neolithic and post-Neolithic (5000 to

3900 BP). For the genetic data, samples in the overlapping periods were assigned based on genetic population affiliation, inferred using supervised ADMIXTURE (81), which, in most of Western Europe, corresponds closely to archaeological context (38, 48). In particular, the Mesolithic/Neolithic overlap was resolved based on whether each individual had more (Neolithic) or less (Mesolithic) than 50% ancestry related to northwest Anatolian Neolithic farmers. The Neolithic/post-Neolithic overlap was resolved based on whether individuals had more than 25% ancestry related to Bronze Age Steppe populations ("Steppe ancestry"; see ref. 83 for details). For the skeletal data, group assignment in the overlapping periods was determined by the archaeology of each site. Broadly, sites belonging to the Neolithic have transitioned to agricultural subsistence. Similarly, post-Neolithic populations are broadly defined by evidence of metal working (Copper, Bronze, and Iron Ages, and later periods). In particular, we included Late Eneolithic (Copper Age) sites associated with Corded Ware and Bell Beaker material culture in the post-Neolithic category, but for consistency with the genetic classifications, we included 8 Early Eneolithic (before 4500 BP) individuals in the Neolithic category, since this precedes the appearance of Steppe ancestry in Western Europe. We excluded samples more recent than 1165 BP.

Linear Models. We fitted a series of linear models to changes in both PRS and stature data with time. In the most general model, we allow both the intercept and slope to vary between groups. We then either force some of the slopes to be zero, or some of the adjacent groups to have identical parameters. We describe the models using underscores to indicate changes in parameters, lowercase to indicate slopes (change with respect to time) fixed to zero, and uppercase to indicate free slopes (i.e., linear trends with time). For example, "E_L_M_N_B" is the most general model, "elmnb" indicates that all groups have the same mean and there is no change with time, and "ELMN_B" indicates that the first 4 groups share the same parameters, and the post-Neolithic has different parameters. The models shown in Figs. 1 and 2 are "e_lmn_b" (*A* and *B*), "e_lm_nb" (*C*), "ELMN_B" (*D* and *E*), and "ELM_NB" (*F*). To analyze geographic variation, we used the residuals of the "ELMN_B" model for the PRS and "ELM_NB" for skeletal stature, and fitted regressions against latitude and longitude.

Polygenic Selection Test. We computed bootstrap P values for the Q_x statistic (73) by recomputing the statistic for random sets of SNPs in matched 5% derived allele frequency bins (polarized using the chimpanzee reference gnome panTro2). For each bootstrap replicate, we keep the original effect sizes but replace the frequencies of each SNP with one randomly sampled from the same bin. Unlike the PRS calculations, we ignored missing data, since the Q_x statistic uses only the population-level estimated allele frequencies and not individual-level data. We tested a series of nested sets of SNPs (x axis in Fig. 5), adding SNPs in 100 SNP batches, ordered by increasing P value, down to a P value of 0.1.

Simulated GWAS Data. We simulated GWAS, generating causal effects at a subset of around 159,385 SNPs in the intersection of SNPs, which passed QC in the UK Biobank GWAS, are part of the 1240 k capture, and are in the POBI dataset (84). We assumed that the variance of the effect size of an allele of frequency f was proportional to $[f(1 - f)]^{\alpha}$, where the parameter α measures the relationship between frequency and effect size (85). We performed 100 simulations with $\alpha = -1$ (the most commonly used model, where each SNP explains the same proportion of phenotypic variance) and 100 with $\alpha = -0.45$ as estimated for height (85). We then added an equal amount of random noise to the simulated genetic values, so that the SNP heritability equaled 0.5. We tested for association between these SNPs and the simulated phenotypes. Using these results as summary statistics, we computed PRS and Q_x tests using the pipeline described above.

1. NCD Risk Factor Collaboration (NCD-RisC), A century of trends in adult human height. *eLife* **5**, e13410 (2016).
2. T. J. Hatton, B. E. Bray, Long run trends in the heights of European men, 19th–20th centuries. *Econ. Hum. Biol.* **8**, 405–413 (2010).
3. E. Arcaleni, Secular trend and regional differences in the stature of Italians, 1854–1980. *Econ. Hum. Biol.* **4**, 24–38 (2006).
4. C. B. Ruff Ed., *Skeletal Variation and Adaptation in Europeans: Upper Paleolithic to the Twentieth Century* (John Wiley and Sons, Hoboken, NJ, 2018).
5. S. Benazzi *et al.*, Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature* **479**, 525–528 (2011).
6. T. Higham *et al.*, The earliest evidence for anatomically modern humans in north-western Europe. *Nature* **479**, 521–524 (2011).
7. E. Rosenstock *et al.*, Human stature in the Near East and Europe ca. 10,000–1000 BC: Its spatio-temporal development in a Bayesian errors-in-variables model. *Archaeol. Anthrop. Sci.*, 10.1007/s12520-12019-00850-12523 (2019).
8. J. M. Perkins, S. V. Subramanian, G. Davey Smith, E. Özaltin, Adult height, nutrition, and population health. *Nutr. Rev.* **74**, 149–165 (2016).
9. E. Rosenstock, M. Groß, A. Hujić, A. Scheibner, "Back to good shape: Biological stan-dard of living in the Copper and Bronze ages and the possible role of food" in *The Third Food Revolution? Setting the Bronze Age Table: Common Trends in Economic and Subsistence Strategies in Bronze Age Europe*, J. Kneisel *et al.*, Eds. (Dr. Rudolph Habelt GmbH, Bonn, Germany, 2015), pp. 121–152.
10. F. Galton, Regression towards mediocrity in hereditary stature. *J. Anthropol. Inst. G. B. Irel.* **15**, 246–263 (1886).
11. G. Lettre, Recent progress in the study of the genetics of height. *Hum. Genet.* **129**, 465–472 (2011).
12. R. A. Fisher, The correlation between relatives on the supposition of Mendelian in-heritance. *Trans. R. Soc. Edinb.* **52**, 399–433 (1918).
13. K. Silventoinen *et al.*, Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Res.* **6**, 399–408 (2003).
14. P. M. Visscher, W. G. Hill, N. R. Wray, Heritability in the genomics era—concepts and misconceptions. *Nat. Rev. Genet.* **9**, 255–266 (2008).
15. A. R. Wood *et al.*; Electronic Medical Records and Genomics (eMEMERGEGE) Con-sortium; MIGen Consortium; PAGEGE Consortium; LifeLines Cohort Study, Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).
16. L. Yengo *et al.*; GIANT Consortium, Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
17. H. Lango Allen *et al.*, Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
18. E. Marouli *et al.*; EPIC-InterAct Consortium; CHD Exome+ Consortium; ExomeBP Consortium; T2D-Genes Consortium; GoT2D Genes Consortium; Global Lipids Genetics Consortium; ReproGen Consortium; MAGIC Investigators, Rare and low-frequency coding variants alter human adult height. *Nature* **542**, 186–190 (2017).
19. A. R. Martin *et al.*, Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
20. A. R. Martin *et al.*, Human demographic history impacts genetic risk prediction across diverse populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
21. J. J. Berg *et al.*, Reduced signal for polygenic adaptation of height in UK Biobank. *eLife* **8**, e39725 (2019).
22. M. Sohail *et al.*, Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* **8**, e39702 (2019).
23. P. Skoglund, I. Mathieson, Ancient genomics of modern humans: The first decade. *Annu. Rev. Genomics Hum. Genet.* **19**, 381–404 (2018).
24. S. Marciniak, G. H. Perry, Harnessing ancient genomes to study the history of human adaptation. *Nat. Rev. Genet.* **18**, 659–674 (2017).
25. M. Slatkin, F. Racimo, Ancient DNA and human history. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6380–6387 (2016).
26. S. Wilde *et al.*, Direct evidence for positive selection of skin, hair, and eye pigmen-tation in Europeans during the last 5,000 y. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 4832–4837 (2014).
27. M. E. Allentoft *et al.*, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
28. J. Burger, M. Kirchner, B. Bramanti, W. Haak, M. G. Thomas, Absence of the lactase-persistence-associated allele in early Neolithic Europeans. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3736–3741 (2007).
29. I. Mathieson *et al.*, Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
30. R. Martiniano *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 (2017).
31. J. J. Berg, X. Zhang, G. Coop, Polygenic adaptation has impacted multiple anthro-pometric traits. bioRxiv:10.1101/167551v4 (18 April 2019).
32. L. M. Cassidy *et al.*, Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 368–373 (2016).
33. Q. Fu *et al.*, The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
34. C. Gamba *et al.*, Genome flux and stasis in a five millennium transect of European prehistory. *Nat. Commun.* **5**, 5257 (2014).
35. G. González-Fortes *et al.*, Paleogenomic evidence for multi-generational mixing be-tween Neolithic farmers and Mesolithic hunter-gatherers in the Lower Danube Basin. *Curr. Biol.* **27**, 1801–1810.e10 (2017).
36. T. Günther *et al.*, Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. *PLoS Biol.* **16**, e2003703 (2018).
37. T. Günther *et al.*, Ancient genomes link early farmers from Atapuerca in Spain to modern-day Basques. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11917–11922 (2015).
38. W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
39. Z. Hofmanová *et al.*, Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6886–6891 (2016).
40. E. R. Jones *et al.*, Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, 8912 (2015).
41. E. R. Jones *et al.*, The Neolithic transition in the Baltic was not driven by admixture with early European farmers. *Curr. Biol.* **27**, 576–582 (2017).
42. A. Keller *et al.*, New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat. Commun.* **3**, 698 (2012).
43. G. M. Kılınç *et al.*, The demographic development of the first farmers in Anatolia. *Curr. Biol.* **26**, 2659–2666 (2016).
44. I. Lazaridis *et al.*, Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218 (2017).
45. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
46. M. Lipson *et al.*, Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
47. R. Martiniano *et al.*, Genomic signals of migration and continuity in Britain before the Anglo-Saxons. *Nat. Commun.* **7**, 10326 (2016).
48. I. Mathieson *et al.*, The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
49. I. Olalde *et al.*, Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. *Nature* **507**, 225–228 (2014).
50. I. Olalde *et al.*, The Beaker phenomenon and the genomic transformation of north-west Europe. *Nature* **555**, 190–196 (2018).
51. I. Olalde *et al.*, A common genetic origin for early farmers from Mediterranean Cardial and Central European LBK cultures. *Mol. Biol. Evol.* **32**, 3132–3142 (2015).
52. A. Omrak *et al.*, Genomic evidence establishes Anatolia as the source of the European Neolithic Gene Pool. *Curr. Biol.* **26**, 270–275 (2016).
53. L. Saag *et al.*, Extensive farming in Estonia started through a sex-biased migration from the Steppe. *Curr. Biol.* **27**, 2185–2193.e6 (2017).
54. S. Schiffels *et al.*, Iron Age and Anglo-Saxon genomes from East England reveal British migration history. *Nat. Commun.* **7**, 10408 (2016).
55. M. Sikora *et al.*, Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).
56. P. Skoglund *et al.*, Genomic diversity and admixture differs for Stone-Age Scandina-vian foragers and farmers. *Science* **344**, 747–750 (2014).
57. K. R. Veeramah *et al.*, Population genomic analysis of elongated skulls reveals ex-tensive female-biased immigration in Early Medieval Bavaria. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3494–3499 (2018).
58. C. B. Ruff *et al.*, Stature and body mass estimation from skeletal remains in the Eu-ropean Holocene. *Am. J. Phys. Anthropol.* **148**, 601–617 (2012).
59. P. Skoglund *et al.*, Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466–469 (2012).
60. Y. Chan *et al.*; GIANT Consortium, Genome-wide analysis of body proportion classifies height-associated variants by mechanism of action and implicates genes important for skeletal development. *Am. J. Hum. Genet.* **96**, 695–708 (2015).
61. C. B. Ruff, Variation in human body size and shape. *Annu. Rev. Anthropol.* **31**, 211–232 (2002).
62. J. A. Allen, The influence of physical conditions in the genesis of species. *Radical Rev.* **1**, 108–140 (1887).
63. C. Bergman, Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien* **3**, 595–708 (1847).
64. V. Formicola, M. Giannecchini, Evolutionary trends of stature in upper Paleolithic and Mesolithic Europe. *J. Hum. Evol.* **36**, 319–333 (1999).
65. H. Chirchir *et al.*, Recent origin of low trabecular bone density in modern humans. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 366–371 (2015).
66. H. Chirchir, C. B. Ruff, J. A. Junno, R. Potts, Low trabecular bone density in recent sedentary modern humans. *Am. J. Phys. Anthropol.* **162**, 550–560 (2017).
67. U. Pettersson, M. Nilsson, V. Sundh, D. Mellström, M. Lorentzon, Physical activity is the strongest predictor of calcaneal peak bone mass in young Swedish men. *Osteoporos. Int.* **21**, 447–455 (2010).
68. D. C. Bauer *et al.*; Osteoporotic Fractures in Men (MrOS) Research Group, Quantitative ultrasound predicts hip and non-spine fracture in men: The MrOS study. *Osteoporos. Int.* **18**, 771–777 (2007).
69. D. C. Bauer *et al.*; Study of Osteoporotic Fractures Research Group, Broadband ul-trasound attenuation predicts fractures strongly and independently of densitometry in older women. A prospective study. *Arch. Intern. Med.* **157**, 629–634 (1997).
70. M. Kneissel *et al.*, Age- and sex-dependent cancellous bone changes in a 4000y BP population. *Bone* **15**, 539–545 (1994).
71. C. B. Ruff, "Biomechanical analyses of archaeological human skeletal samples" in *Biological Anthropology of the Human Skeleton*, M. A. Katzenburg, A. L. Grauer, Eds. (John Wiley and Sons, New York, ed. 3, 2019), pp. 189–224.
72. C. B. Ruff *et al.*, Gradual decline in mobility with the adoption of food production in Europe. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7147–7152 (2015).
73. J. J. Berg, G. Coop, A population genetic signal of polygenic adaptation. *PLoS Genet.* **10**, e1004412 (2014).
74. C. B. Ruff, E. Trinkaus, T. W. Holliday, Body mass and encephalization in Pleistocene *Homo*. *Nature* **387**, 173–176 (1997).
75. T. W. Holliday, Body proportions in Late Pleistocene Europe and modern human or-igins. *J. Hum. Evol.* **32**, 423–448 (1997).

ANTHROPOLOGY

76. V. Formicola, B. M. Holt, Resource availability and stature decrease in Upper Palaeolithic Europe. *J. Anthropol. Sci.* **85**, 147–155 (2007).

77. C. C. Roseman, B. M. Auerbach, Ecogeography, genetics, and the evolution of human body form. *J. Hum. Evol.* **78**, 80–90 (2015).

78. K. R. Savell, B. M. Auerbach, C. C. Roseman, Constraint, natural selection, and the evolution of human body form. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 9492–9497 (2016).

79. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).

80. C. C. Chang *et al.*, Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).

81. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

82. Hail Team, Hail. https://github.com/hail-is/hail/releases/tag/0.2.13. Accessed 24 September 2019 (2019).

83. S. Mathieson, I. Mathieson, FADS1 and the timing of human adaptation to agriculture. *Mol. Biol. Evol.* **35**, 2957–2970 (2018).

84. S. Leslie *et al.*; Wellcome Trust Case Control Consortium 2; International Multiple Sclerosis Genetics Consortium, The fine-scale genetic structure of the British population. *Nature* **519**, 309–314 (2015).

85. A. P. Schoech *et al.*, Quantification of frequency-dependent genetic architectures in 25 UK Biobank traits reveals action of negative selection. *Nat. Commun.* **10**, 790 (2019).