The evolution of skin pigmentation-associated variation in West Eurasia

Dan Ju*a,1 and Iain Mathieson*a,1

*Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

Edited by Anne C. Stone, Arizona State University, Tempe, AZ, and approved November 4, 2020 (received for review May 11, 2020)

Skin pigmentation is a classic example of a polygenic trait that has experienced directional selection in humans. Genome-wide association studies have identified well over a hundred skin pigmentation-associated loci, and genomic scans in present-day and ancient populations have identified selective sweeps for a small number of light pigmentation-associated alleles in Europeans. It is unclear whether selection has operated on all of the genetic variation associated with skin pigmentation as opposed to just a small number of large-effect variants. Here, we address this question using ancient DNA from 1,158 individuals from West Eurasia covering a period of 40,000 y combined with genome-wide association summary statistics from the UK Biobank. We find a robust signal of directional selection in ancient West Eurasians on 170 skin pigmentation-associated variants ascertained in the UK Biobank. However, we also show that this signal is driven by a limited number of large-effect variants. Consistent with this observation, we find that a polygenic selection test in present-day populations fails to detect selection with the full set of variants. Our data allow us to disentangle the effects of admixture and selection. Most notably, a large-effect variant at SLC24A5 was introduced to Western Europe by migrations of Neolithic farming populations but continued to be under selection post-admixture. This study shows that the response to selection for light skin pigmentation in West Eurasia was driven by a relatively small proportion of the variants that are associated with present-day phenotypic variation.

skin pigmentation | polygenic selection | complex traits | evolution | ancient DNA

Skin pigmentation exhibits a gradient of variation across human populations that tracks with latitude (1). This gradient is thought to reflect selection for lighter skin pigmentation at higher latitudes, as lower UVB exposure reduces vitamin D biosynthesis, which affects calcium homeostasis and immunity (2–4). Studies of present-day and ancient populations have revealed signatures of selection at skin pigmentation loci (5–9), and single-nucleotide polymorphisms (SNPs) associated with light skin pigmentation at some of these genes exhibit a signal of polygenic selection in Western Eurasians (10). However, this observation, the only documented signal of polygenic selection for skin pigmentation, is based on just four loci (SLC24A5, SLC45A2, TYR, and APBA2/OCA2).

Therefore, while the existence of selective sweeps at a handful of skin pigmentation loci is well established, the evidence for polygenic selection—a coordinated shift in allele frequencies across many trait-associated variants (11)—is less clear. Recently, genome-wide association studies (GWAS) of larger samples and more diverse populations (12–15) have emphasized the polygenic architecture of skin pigmentation. This raises the question of whether selection on skin pigmentation acted on all of this variation or was indeed driven by selective sweeps at a relatively small number of loci (11).

The impact of demographic transitions on the evolution of skin pigmentation also remains an open question. The Holocene history (~12,000 y before present [BP]) of Europe was marked by waves of migration and admixture between three highly diverged populations: hunter-gatherers, Early Farmers, and Steppe ancestry populations (16, 17). Skin pigmentation-associated loci may have been selected independently or in parallel in one or more of these source populations or may instead only have been selected after admixture. The impact of ancient shifts in ancestry is difficult to resolve using present-day data, but using ancient DNA, we can separate the effects of ancestry and selection and identify which loci were selected in which populations.

Here, we use ancient DNA to track the evolution of loci that are associated with skin pigmentation in present-day Europeans. Although we cannot make predictions about the phenotypes of ancient individuals or populations, we can assess the extent to which they carried the same light pigmentation alleles that are present today. This allows us to identify which pigmentation-associated variants have changed in frequency due to positive selection and the timing of these selective events. We present a systematic survey of the evolution of European skin pigmentation-associated variation, tracking over a hundred loci over 40,000 y of human history—almost the entire range of modern human occupation of Europe.

Results

Skin Pigmentation-Associated SNPs Show a Signal of Positive Selection in Ancient West Eurasians. We obtained skin pigmentation-associated SNPs from the UK Biobank GWAS for skin color released by the Neale Lab (14). We analyzed the evolution of these variants using two datasets of publicly available ancient DNA (Fig. 1 A–C and SI Appendix, Fig. S1). The first (“capture-shotgun”) consists of data from 1,158 individuals dating from 45,000–715 y BP genotyped at ~1.2 million variants (7, 16, 18–52). The second (“shotgun”) is a digital supplementary material.

Significance

Some of the genes responsible for the evolution of light skin pigmentation in Europeans show signals of positive selection in present-day populations. Recently, genome-wide association studies have highlighted the highly polygenic nature of skin pigmentation. It is unclear whether selection has operated on all of these genetic variants or just a subset. By studying variation in a thousand ancient genomes from West Eurasia covering 40,000 y, we are able to study both the aggregate behavior of pigmentation-associated variants and the evolutionary history of individual variants. We find that the evolution of light skin pigmentation in Europeans was driven by frequency changes in a relatively small fraction of the genetic variants that are associated with variation in the trait today.

Author contributions: D.J. and I.M. designed research; D.J. and I.M. performed research; I.M. contributed new reagents/analytic tools; D.J. analyzed data; and D.J. and I.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under the Creative Commons Attribution License 4.0 (CC BY).

1To whom correspondence may be addressed. Email: danju@pennmedicine.upenn.edu or mathi@pennmedicine.upenn.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2009227118/-/DCSupplemental.

Published December 21, 2020.

PNAS 2021 Vol. 118 No. 1 e2009227118
https://doi.org/10.1073/pnas.2009227118
subset of 248 individuals with genome-wide shotgun sequence data (16, 18–25, 28, 32, 35, 38, 40–44, 47, 49, 51, 53–60). This smaller dataset allows us to capture variants that may not be well tagged by the genotyped SNPs in the capture-shotgun dataset. Weighting variants by their GWAS-estimated effect sizes, we show that the polygenic score—the weighted proportion of dark pigmentation alleles—decreased significantly over the past 40,000 y (Fig. 1A and B, \( P < 1 \times 10^{-10} \) based on a genomic null distribution and accounting for changes in ancestry; SI Appendix, Fig. S2). We recapitulate this significant decrease using a different set of UK Biobank GWAS summary statistics calculated using linear mixed models (61) (SI Appendix, Fig. S3) and identifying independent signals based on predefined approximately independent linkage disequilibrium (LD) blocks (62) (SI Appendix, Fig. S4), rather than LD clumping. The genetic scores of the oldest individuals in the dataset fall within the range of present-day West African populations, showing that Early Upper Paleolithic (~50–20,000 y BP) individuals, such as Ust Ishim, carried fewer of the light skin pigmentation alleles that are common in present-day Europe.

Over the past 15,000 y, which covers most of the data, the polygenic score decreased (\( P < 1 \times 10^{-4} \)) at a similar rate to the estimated rate over the 40,000-y period (5.27 \( \times 10^{-5} \) per year vs. 3.47 \( \times 10^{-5} \) per year; \( P = 0.04 \)). Stratifying individuals by ancestry (hunter-gatherer, Early Farmer, and Steppe) using ADMIXTURE revealed baseline differences in genetic score across groups (Fig. 1C). Mesolithic hunter-gatherers carried fewer light pigmentation alleles than Early Farmer or Steppe ancestry populations. This difference (0.091 score units) was similar to the difference between Mesolithic hunter-gatherers and Early Upper Paleolithic populations (0.097 score units). We tested for evidence of change in score over time within groups while controlling for changes in ancestry, finding no significant change within hunter-gatherer or Early Farmer populations (\( P = 0.32; P = 0.29 \)). Steppe ancestry populations, however, exhibited a significant decrease (\( P = 0.007 \)), with a steeper slope than that of Early Farmer or hunter-gatherer (\( P = 0.004 \) and \( P = 0.08 \), respectively).

While the UK Biobank is well powered to detect variation that is segregating in present-day Europeans, it may not detect variants that are no longer segregating. We therefore manually curated a separate list of 18 SNPs identified as associated with skin pigmentation from seven studies of populations representing diverse ancestries (Dataset S1E). Using this set of SNPs to generate unweighted scores, we observed similar results as with the UK Biobank ascertained SNPs (Fig. 1D–F). We find a significant decrease in genetic score over both 40,000 and 15,000 y in the capture-shotgun dataset (\( P < 1 \times 10^{-4} \) for both), although the decrease in the shotgun dataset was not significant (\( P = 0.17 \)), probably reflecting a lack of power from the small number of ancient samples and SNPs.
Ancestry and Selection Have Different Effects on the Histories of Each Variant. Next, we investigated the evolution of each variant independently. We separated the effects of ancestry and selection in the data by regressing presence or absence of the alternate allele for each pigmentation-associated SNP on both date and ancestry, inferred using principal component analysis (PCA) (63). Significant effects of ancestry on frequency can be interpreted as changes in allele frequency that can be explained by changes in ancestry. Significant changes in frequency over time, after accounting for ancestry, can be interpreted as evidence for selection occurring during the analyzed time period.

The SNP rs16891982 at the SLC45A2 locus shows the strongest evidence for selection across analyses (Fig. 2), but we note this does not necessarily imply positive selection over the entire time transect. We observed little evidence for demographic transitions in driving the increase in light allele frequency over time for this SNP. In addition, we found robust signals of selection for the light allele of rs1126809 near TYR, rs12913832 and rs1635168 near HERC2, and rs7109255 near GRM5 (Fig. 2 B and C). In some cases—for example at rs6120849 near EDEM2 and rs7870409 near RALGPS1—light alleles decreased in frequency more than can be explained by changes in ancestry (Fig. 2 B and C). The dark allele at rs6120849 (EDEM2) is also significantly associated with several other traits including increased sitting and standing height in the UK Biobank, increased blood creatinine, and increased heel bone mineral density (64). The dark allele of rs7870409 (RALGPS1) is associated with increased arm and trunk fat percentage (65). These observations raise the possibility of pleotropic effects constraining the evolution of these loci or of spurious pleiotropy due to uncorrected population stratification in the GWAS.

On the other hand, the changes in frequencies of several variants appear to have been driven largely by changes in ancestry. For example, rs4778123 near OCA2, rs2153271 near BNC2, and rs3758833 near CTSC show evidence that their changes in frequency were consistent with changes in genome-wide ancestry (Fig. 2 B and C). At SLC24A4, rs2675345 shows evidence of change with both ancestry and time, suggesting that even after the spread of the light allele from Anatolia into Western Europe in the Neolithic (7) selection continued to occur post-admixture. Again, not all of these cases involve an increase in the frequency of the light pigmentation allele over time. The light allele of rs4778123 (OCA2) was at high frequency in hunter-gatherers but lower in later populations (SI Appendix, Fig. S6B). From the
manually curated set of SNPs (Fig. 2 D–F), rs12203592 near IFR4 also displays a marked effect of ancestry with higher light allele frequency in hunter-gatherers (SI Appendix, Fig. S6). While rs12203592 was not present in the UK Biobank summary statistics, another SNP at the IFR4 locus, rs3778607, was present but with a smaller ancestry effect (SI Appendix, Fig. S7).

Lack of PBS Selection Signal across Pigmentation SNPs in Source Populations of Europeans. As we described in the previous section, the frequencies of several skin pigmentation SNPs appear to have been driven by changes in ancestry. This observation suggests that their frequencies have diverged between the European source populations and then subsequently been driven by admixture. Frequency differences across the source populations might reflect the action of either genetic drift or selection. To examine this question, we looked for evidence of selection at skin pigmentation SNPs in each ancestral population using the population branch statistic (PBS) test (66). We ran ADMIXTURE with K = 3 on the capture shotgun dataset (SI Appendix, Fig. S8) and used the inferred source population allele frequencies to calculate PBS for each source in 20-SNP nonoverlapping windows, with the 1000 Genomes CHB (Han Chinese) and YRI (Yoruba) populations as outgroups.

Few of the skin pigmentation loci show extreme PBS values (Fig. 3), and even fewer show evidence of divergent frequencies across ancient groups (SI Appendix, Fig. S10). On the Early Farmer and Steppe ancestry branches, but not the hunter-gatherer branch, the SLC24A5 locus exhibited the strongest signal (Fig. 3 B and C), indicating selection at the locus in both Early Farmer and Steppe source populations. The RABGAP1 locus also exhibited an elevated signal of selection in Early Farmer and Steppe, which remained when using a 40-SNP window size (SI Appendix, Fig. S11 B and C). In all three ancestry groups, we observed an elevated PBS at the OCA2 locus, with the most extreme value on the hunter-gatherer branch. However, unlike the SLC24A5 variant, neither of these SNPs (rs644490 and rs9920172) showed a substantial effect of ancestry from the individual SNP regression analysis. As a whole, the PBS values of the 170 skin pigmentation SNPs do not significantly deviate from the genome-wide distribution of PBS for any of the ancient source populations (SI Appendix, Fig. S12).

Signals of Selection Are Restricted to the SNPs with the Largest Effect Sizes. Since our results suggest that pigmentation-associated variants exhibit different changes in allele frequency, we further examined the signal of selection found in Fig. 1. In general, SNPs with larger GWAS effect sizes changed more in frequency over time than those with small effect sizes ($P = 2.8 \times 10^{-18}$; adjusted $R^2 = 0.36$ (SI Appendix, Fig. S13A)). To understand the contribution of these large GWAS effect size SNPs, we iteratively removed SNPs from the polygenic score calculations (Fig. 4A). Removing the SNP rs16891982 at SLC45A2, we still found a significant decrease of score over time ($P < 1 \times 10^{-5}$), but the estimated rate of change ($\beta$) decreased by 39%. Removing the top two SNPs at the SLC45A2 and SLC24A5 locus further attenuated the signal with a decrease in $\beta$ of 58% ($P = 3.96 \times 10^{-6}$). We removed the five SNPs with the largest GWAS effect sizes and found the effect of time attenuated ($P = 0.026$) with a decrease in $\beta$ of 78%. Removing the top 10 SNPs attenuated the signal even more ($P = 0.050$) with an 81% decrease and removal of 15 or 20 SNPs abolished the signal ($P = 0.147, P = 0.52$).

Using present-day populations from the 1000 Genomes Project (67), we tested for polygenic selection with the $Q_x$ statistic (10), which measures overdispersion of trait-associated SNPs relative to the genome-wide expectation. We performed the test with different numbers of SNPs ordered by largest to smallest absolute GWAS effect size (Fig. 4B). Using all 170 SNPs in the test, we failed to observe a signal of polygenic selection ($P = 0.27$). However, when we restricted to the top 5 largest GWAS effect size SNPs, we found a significant signal ($P < 2.5 \times 10^{-5}$). We observed significant $Q_x$ statistics at $P < 0.05$ up to the top 30 SNPs, but it appears this signal is driven by the top SNPs of largest effect. After removing the top 5 SNPs and testing the remaining 25 top SNPs, the signal of polygenic selection disappeared ($P = 0.59$). Examining the predicted variance explained per SNP (Fig. 4C and SI Appendix, Fig. S14), we predict that, of the UK Biobank SNPs, two SNPs at SLC45A2 and SLC24A5 make the largest contribution to the difference between West African and European populations. Most UK Biobank skin pigmentation-associated SNPs are predicted to contribute relatively little to the between-population variance.

Discussion

Large whole-genome ancient DNA datasets have, in the past few years, allowed us to track the evolution of variation associated with both simple and complex traits (68, 69). The majority of samples are from Western Eurasia, and we focus on that region, noting that a parallel process of selection for light skin pigmentation has acted in East Asian populations (9, 70). In West Eurasia, selection for skin pigmentation appears to have been dominated by a small number of selective sweeps at large-effect variants. There are a number of possible explanations for this. Many of the variants detected by the UK Biobank GWAS may

![Fig. 3](https://doi.org/10.1073/pnas.2009227118)
have such small effects that they were effectively neutral or experienced such weak selection that we cannot detect it. Other variants may not have responded to selection because of pleiotropic constraint. The five SNPs with the largest effect sizes in UK Biobank do have some have small but significant associations with anthropometric phenotypes (for example, rs1805007 at MC1R with standing height), but by far their most significant associations are with hair color, facial aging, tanning, melanoma risk, and other phenotypes that are likely related to pigmentation (64). Finally, GWAS effect size estimates in European populations today may not reflect the impact these variants had in the past due to epistatic or gene–environment interactions. This would not affect the individual SNP time series and PBS analyses since they do not rely on effect sizes as weights but would reduce power for the weighted polygenic score and Qx analyses.

Our analyses centered on 170 skin pigmentation-associated SNPs that are present on the 1240K capture array. To check for possible bias introduced by the capture array, we checked how many of the 242 SNPs used in the shotgun analysis were well tagged by the capture array. Of the 242 SNPs, 33 are shared, a further 67 are tagged at $r^2 \geq 0.8$, and a further 19 SNPs are tagged at $r^2 \geq 0.5$. However, while the 142 SNPs that are not tagged at $r^2 \geq 0.8$ by the capture array do contain some moderate effect size SNPs, they do not contribute to the selection signal ($P = 0.31$) (SI Appendix, Fig. S15D). Therefore, although the capture array does not capture all of the variation contributing to present-day skin pigmentation variation, it does capture the vast majority of the variation that contributed to the evolution of the phenotype, justifying the use of the capture shotgun dataset for detailed investigation of the evolutionary trends.

We find little evidence of parallel selection on independent haplotypes at skin pigmentation loci, suggesting that differences in allele frequency across ancestry groups were mostly due to genetic drift. One exception is that the light allele at SLC24A5 was nearly fixed in both Early Farmer and Steppe ancestry populations due to selection. However, even for this variant we observe a signal of ongoing selection in our data even after admixture with hunter-gatherers, indicating continuing selection after admixture. This is analogous to the rapid selection at the same locus for the light allele introduced via admixture into the Khoesan, who now occupy southern Africa (12, 71).

We are also able to test previous claims about selection on particular pigmentation genes. We find no evidence of positive selection in Europeans at the MC1R locus in contrast to previous reports (72, 73). Among UK Biobank SNPs, rs1805007 near MC1R explains a relatively large amount of variation within the UK but is predicted to explain relatively little of the variation between Europe and West Africa (Fig. 4C). The TYR1 locus has been previously identified as a target of selection in Europeans (5, 8, 9, 74, 75), although some studies (76, 77) have questioned this finding. Our analysis shows some support for selection at this locus, with a 20-SNP window centered around rs1325132 being in the top 1.1% of genome-wide PBS windows on the hunter-gatherer lineage. Some studies (76, 78) detect a signal of recent selection in Europeans at the KitTLG SNP rs12821256 that is functionally associated with blond hair color (79). We find no evidence of selection on the West Eurasian lineage at this SNP in our time series analysis, but this may reflect a lack of power to detect a relatively small change in frequency. Finally, at the OCA2 locus, we recapitulate an observation of independent selection in Europeans and East Asians (70, 80) (SI Appendix, Fig. S9). In the ancient populations, we found an elevated PBS signal around rs9920172 that was most prominent in hunter-gatherers but also elevated in Early Farmer and Steppe ancestry populations (Fig. 3), suggesting a relatively early episode of selection in West Eurasians.

Relatively dark skin pigmentation in Early Upper Paleolithic Europe would be consistent with those populations being relatively poorly adapted to high-latitude conditions as a result of having recently migrated from lower latitudes. On the other hand, although we have shown that these populations carried few of the light pigmentation alleles that are segregating in present-day Europe, they may have carried different alleles that we cannot now detect. As an extreme example, Neanderthals and the Altai Denisovan individuals show genetic scores that are in a similar range to Early Upper Paleolithic individuals (SI Appendix, Table S1), but it is highly plausible that these populations, who lived at high latitudes for hundreds of thousands of years, would have adapted independently to low UV levels. For this reason, we cannot confidently make statements about the skin pigmentation of ancient populations.

Our study focused, for reasons of data availability, on the history of skin pigmentation evolution in West Eurasia. However, there is strong evidence that a parallel trend of adaptation to low UVB conditions occurred in East Asia (15, 70, 81, 82). Less is known about the loci that have been under selection in East Asia, aside from some variants at OCA2 (80–82). Similarly, multiple studies have documented selection for both lighter and darker skin pigmentation in parts of Africa (12, 13, 83). Future work should test whether the process of adaptation in other parts of the world was similar to that in Europe. The lack of known

![Figure 4](image-url)

**Fig. 4.** (A) Regressions of genetic score based on UK Biobank SNPs using the capture-shotgun dataset over date, with scores using all SNPs and iteratively removing top GWAS effect size SNPs. (B) Qx, empirical $-\log_{10}(P$ values) using UK Biobank skin pigmentation-associated variants with all 1000 Genomes populations. Different numbers of SNPs were used to calculate Qx, which were ordered by GWAS-estimated effect size. (C) Variance explained by UK Biobank SNPs with nearest gene labeled between UK and YRI populations ($x$-axis) and within the UK population ($x$-axis).
skin pigmentation loci in scans of positive selection in East Asian populations (84, 85) raises the possibility that selection may have been more polygenic in East Asians than in Europeans. Finally, the evidence for polygenic directional selection on other complex traits in humans is inconclusive (86, 87). We suggest that detailed studies of other phenotypes using ancient DNA can be helpful at more generally identifying the types of processes that are important in human evolution.

**Methods**

**Capture-Shotgun Ancient DNA Dataset.** We downloaded version 37.2 of the publicly available datasets of 2,107 ancient and 5,637 present-day (Human Origins dataset) samples from the Reich Lab website: https://reich.hms.harvard.edu/downloadable-genotypes-present-day-and-ancient-dna-data-compiled-published-papers. Ancient individuals were treated as pseudohaploid (i.e., carry only one allele) because most samples are low coverage. For many of these samples, 1,233,013 sites were genotyped using an in-solution capture method (88), while those with whole-genome shotgun sequence data were genotyped at the same set of sites. For our regression models, we included ancient individuals from West Eurasia, defined here as the region west of the Urals (longitude, ~60°E) and north of 35°N (SI Appendix, Fig. S1). We excluded samples with less than 0.1x coverage. We removed duplicate and closely related samples (e.g., first-degree relatives), by calculating pairwise identity by state (IBS) between ancient individuals. For each individual, we identified a corresponding individual with the highest IBS. Based on the distribution of highest IBS values, we identified the pairs with IBS > 7 as duplicates and retained the individual with fewest missing sites. We also manually removed related samples that were explicitly annotated, selecting the report as representative. Finally, we incorporated additional ancient individuals from the Iberian Peninsula from two recent papers (36, 45), applying the same criteria for coverage and relatedness. Ultimately, we compiled a list of 1,158 ancient pseudohaploid individuals in our analyses covering a time span of the last 40,000 y. Only 11 individuals were dated to >15,000 y BP, and 1,147 individuals lived during the last 15,000 y.

**Shotgun Ancient DNA Dataset.** We organized a dataset of samples that were shotgun sequenced without enrichment using the capture array (23, 25–32, 35, 39, 42, 45, 47–51, 54, 56, 58, 60–67). To generate the first 10 PCs from sites on the 1240K capture array, we combined shotgun samples with available genotype data from v37.2 of the Reich laboratory dataset and 15 individuals that we ourselves pulled down from BAM files. We manually removed duplicate samples in the shotgun dataset, preserving the sample with higher coverage. We checked for first-degree relatives in the dataset by removing related samples that were explicitly annotated, selecting the report as representative. Finally, we incorporated additional ancient individuals from the Iberian Peninsula from two recent papers (36, 45), applying the same criteria for coverage and relatedness. Ultimately, we compiled a list of 1,158 ancient pseudohaploid individuals in our analyses covering a time span of the last 40,000 y. Only 11 individuals were dated to >15,000 y BP, and 1,147 individuals lived during the last 15,000 y.

**Skin Pigmentation SNP Curation.** We obtained summary statistics for UK Biobank GWAS for skin color (data field 1717) from the publicly available release by the Neale Lab (version 3, Manifest Release 20180731) (14). The GWAS included genotyped skin color variants (four-dimensional variable: very fair, light, olive, dark olive, brown, black). To identify genome-wide significant and independent SNPs, we performed clumping using PLINK v1.0 b6 (89) with 1000 Genomes GBR as an LD reference panel (−clump-p1 5 x 10−8 −clump-r2 0.05 −clump-kb 250), and followed up with clumping based on physical distance to exclude SNPs within 100 kb of each other. We made two separate lists of UK Biobank SNPs for the shotgun and capture-shotgun datasets because the capture-shotgun dataset was restricted to the 1240K array sites. For the capture-shotgun dataset, we intersected all UK Biobank SNPs for the shotgun and capture-shotgun datasets. We included ancestry as a covariate in the model by including the first 10 PCs. We did not include SNPs from all of the considered papers for our regression models, we included ancient individuals from West Eurasia, defined here as the region west of the Urals (longitude, ~60°E) and north of 35°N (SI Appendix, Fig. S1). We excluded samples with less than 0.1x coverage. We removed duplicate and closely related samples (e.g., first-degree relatives), by calculating pairwise identity by state (IBS) between ancient individuals. For each individual, we identified a corresponding individual with the highest IBS. Based on the distribution of highest IBS values, we identified the pairs with IBS > 7 as duplicates and retained the individual with fewest missing sites. We also manually removed related samples that were explicitly annotated, selecting the report as representative. Finally, we incorporated additional ancient individuals from the Iberian Peninsula from two recent papers (36, 45), applying the same criteria for coverage and relatedness. Ultimately, we compiled a list of 1,158 ancient pseudohaploid individuals in our analyses covering a time span of the last 40,000 y. Only 11 individuals were dated to >15,000 y BP, and 1,147 individuals lived during the last 15,000 y.

**ADMIXTURE Analysis on Capture-Shotgun Data.** We performed unsupervised ADMIXTURE (98) on the dataset of ancient individuals with K = 3, which we found to produce the lowest cross-validation error for 2 < K < 3. The three identified clusters could easily be identified as corresponding to hunter-gatherer, Early Farmer, and Steppe (also referred to as Yamnaya ancestry) (SI Appendix, Fig. S8).

**Time Series Analysis of Genetic Scores.** We calculated genetic scores for each individual from present-day 10,000 Genomes populations (67) and ancient individuals. We computed the scores in two ways. First, we weighted by the GWAS-estimated effect sizes. Because of variable coverage across ancient samples, not all SNPs were present in a given sample. To account for missing information in the creation of weighted scores, we devised a weighted proportion in which we divided the realized score over the maximum possible score given the SNPs present in the sample: Score weighted = \( s_i / \sum_i s_i \), where \( m \) is the number of skin pigmentation SNPs genotyped in the individual, \( d_i \) is the presence of the dark allele at the \( i \)th SNP, and \( \beta_i \) is the GWAS-estimated effect size (Fig. 1 A–C). For the manually curated list of SNPs where we did not have comparable effect size estimates, we computed an unweighted score, effectively assuming that all variants had the same effect size. This score is the proportion of dark alleles an individual carries out of the SNPs used in the construction of the score: Score unweighted = \( (1/m) \sum_{i=1}^m d_i \), where \( m \) is the number of SNPs genotyped in the individual and \( d_i \) is the presence of the dark allele at the \( i \)th SNP (Fig. 1 D–F).

We used logistic regression to examine the association between weighted and unweighted genetic score and time for all ancient samples. We fitted separate models for ancient samples in the shotgun and capture-shotgun datasets. We included ancestry as a covariate in the model by including the first 10 PCs. PCA was performed using smartpca v16000 (63) to generate PCs from present day populations and we projected ancient individuals onto these axes of variation. For the shotgun dataset and the capture-shotgun dataset that included all samples (40,000 y BP), we used 1000 Genomes samples. For the capture-shotgun dataset that restricted to samples later than 15,000 y BP, we used West Eurasians from the Human Origins dataset (16).

We model the score of each individual as the proportion of successes in a binomial sample of size \( m \). That is, \( m \beta_i \sim \text{Binomial}(m_i, \beta_i) \), where, for individual \( i \), \( m_i \) is the number of SNPs genotyped, \( \beta_i \) is the (weighted or unweighted) score, the probability of success \( p_i \) is given by the following:

\[
\log \left( \frac{p_i}{1-p_i} \right) = \alpha + \beta_{\text{date}} \text{date}_i + \beta_{\text{lat}} \text{lat}_i + \beta_{\text{lon}} \text{lon}_i + \beta_{\text{PC1}} PC_{1,i} + \cdots + \beta_{\text{PC10}} PC_{10,i},
\]

and \( \text{date}_i, \text{lat}_i, \) and \( \text{lon}_i \) are the dates, latitude, and longitude of each individual. For the weighted score, \( m_i \beta_i \) is noninteger, but the likelihood can be computed in the same way.

For the stratified analysis, we divided the capture-shotgun dataset into mutually exclusive ancestry groups. We categorized individuals for this analysis as hunter-gatherer if they carried over 60% of the ADMIXTURE-estimated hunter-gatherer component. For placement into the Early Farmer group, we required that an individual carry over 60% of that component. Individuals we categorized as Steppe had over 30% of the ADMIXTURE component and were dated to less than 5,000 y BP. Individuals that are more than recent than 5,000 y BP and have at least 30% of the Steppe ADMIXTURE component were classified as Steppe even if they had more than 60% of the Early Farmer component. Based on these cutoffs, there were 102 hunter-gatherer, 499 Early Farmer, and 478 Steppe ancestry
assigned individuals. We used the same logistic regression model as above for the stratified analysis to control for ancestry. Because the fitted model parameter estimates are overdispersed (SI Appendix, Fig S2), we computed $P$ values from a genome-wide empirical null distribution for $\beta_{\text{date}}$. We made this null distribution by running the regression model described above on scores from sets of random, frequency-matched ($\pm 1\%$) SNPs across the genome, maintaining the same effect sizes for the weighted score. We matched the derived allele frequency based on the EUR super-population frequencies from 1000 Genomes. We reported $P$ values based on 10,000 random samples.

**Time Series Analysis of Individual SNP Allele Frequencies.** We performed logistic regression for each individual SNP using date of the sample and ancestry as covariates. That is, the probability that the haplotype sampled from individual carries the derived allele is $p_d$, where

$$\log \left( \frac{p_d}{1 - p_d} \right) = \beta_{\text{date}} + \beta_{\text{PC1}} + \ldots + \beta_{\text{PC10}} \cdot \text{PC10}.$$

We compared the full model above to a nested model with no principal components by performing a likelihood ratio test (in R we use anova() on nested models, full model, test = ‘Chisq’) to obtain a $P$ value for the ancestry term which encompasses $\beta_{\text{PC1}} - \beta_{\text{PC10}}$. To obtain a $P$ value for $\beta_{\text{date}}$ we compare a nested model without $\beta_{\text{date}}$ to the full model.

**Q, Polygenic Selection Test.** We used the Q test for directional, polygenic selection on skin pigmentation (10). For this test, we used 170 skin pigmentation SNPs obtained from the UK Biobank in all 1000 Genomes populations. We restricted sites to those on the 1240K and constructed a covariance matrix from a total of one million SNPs. To calculate empirical $P$ values, we sampled a total of 500,000 null genetic values matching skin selection on skin pigmentation (10). For this test, we used 170 skin pigmentation SNPs across the genome, maintaining the same effect sizes for the weighted score. We identified nearby genes by examining the National Center for Biotechnology Information Refseq track for assembly GRCh37/hg19 (downloaded from https://genome.ucsc.edu/cgi-bin/hgTables) for genes that overlap with windows with high $P$ values. We note that the gene names are used simply as labels for each locus and do not necessarily represent the causal genes.

**Code Availability.** Scripts used to generate the main results and figures of this paper are available at GitHub, https://github.com/matlibSkinPigmentationCode.

**ACKNOWLEDGMENTS.** We thank Sarah Tishkoff, Alexander Platt, Bábara Bitarello, Samantha Cox, Arslan Zaidi, and Andrew Chen for helpful comments on earlier versions of this manuscript. This research was supported by a Research Fellowship from the Alfred P. Sloan Foundation (FG-2018-10647), a New Investigator Research Grant from the Charles E. Kaufman Foundation (KA2018-98539), National Institute of General Medical Sciences Award R35GM133708, and National Research Service Award T32GM008216. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.
25. A. Keller et al., New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. Nature 552, 698–702 (2018).
26. M. Krzewinkowa et al., Ancient genomes suggest the eastern Pontic-Caspian steppe as the source of western Iron Age nomads. Sci. Adv. 4, eaat4457 (2018).
27. M. Krzewinska et al., Genomic and strontium isotope variation reveal immigration patterns in a Viking Age town. Curr. Biol. 28, 2730–2738.e10 (2018).
28. M. S. Allen et al., Population genomics of Bronze Age Eurasia. Nature 522, 167–172 (2015).
29. I. Lázáridis et al., Genetic origins of the Minoans and Mycenaeans. Nature 548, 214–218 (2017).
30. I. Lázáridis et al., Genomic insights into the origin of farming in the ancient Near East. Nature 536, 419–424 (2016).
31. M. Lipson et al., Parallel palearctic transects reveal complex genetic history of early European farmers. Nature 551, 368–372 (2017).
32. R. Martiniello et al., Genomic signals of migration and continuity in Britain before the Anglo-Saxons. Nat. Commun. 7, 10326 (2016).
33. R. Martinoia 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).
34. I. Mathieson et al., The genomic history of southeastern Europe. Nature 555, 197–203 (2018).
35. I. Olalde et al., Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature 507, 225–228 (2014).
36. I. Olalde et al., The genomic history of the Iberian Peninsula over the past 8000 years. Science 363, 1230–1234 (2019).
37. I. Olalde et al., The Beaker phenomenon and the genomic transformation of north-western Europe. Nature 509, 190–196 (2014).
38. 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).
39. C. E. G. Omrani et al., Understanding 6th-century barbarian social organization and migration through paleogenomics. Nat. Commun. 9, 3547 (2018).
40. A. O'Meara et al., Evidence establishing Anatolia as the source of the European Neolithic gene pool. Curr. Biol. 26, 270–275 (2016).
41. L. Saag et al., Extensive farming in Estonia started through a sex-biased migration from the steppe. Curr. Biol. 27, 2185–2193.e6 (2017).
42. S. Schiffels et al., Iron age and Anglo-Saxon genomes from East England reveal British migration history. Nat. Commun. 7, 10408 (2016).
43. M. Sikora et al., Ancient genomes. PLoS Genet. 6, e1001082 (2010).
44. P. Skoglund et al., Genomic diversity and admixture differ for Stone-Age Scandinavians and farmers in central and elderly Europe. Science 358, 659–662 (2017).
45. V. Villalob-Mouco et al., Survival of late Pleistocene hunter-gatherer ancestry in the Iberian Peninsula. Curr. Biol. 29, 1169–1177.e7 (2019).
46. M. Unterländer et al., Ancestry and demography and descendants of Iron Age no-

47. 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).
48. P. B. Damgaard et al., 137 ancient human genomes from across the Eurasian steppes. Nature 557, 369–374 (2018).
49. P. de Barros Damgaard et al., The first horse herders and the impact of early Bronze Age steppe expansions into Asia. Science 360, eaar7711 (2018).
50. D. M. Fernandez et al., A genomic Neolithic time transect of hunter-farmer admixture in central Poland. Science 348, 14879 (2015).
51. Q. Fu et al., The genetic history of Ice Age Europe. Nature 534, 200–205 (2016).
52. Q. Fu et al., An early modern human from Romania with a recent Neanderthal an-

53. Q. Fu et al., Genome sequence of a 45,000-year-old modern human from western Siberia. Nature 514, 445–449 (2014).
54. T. Güntner et al., Population genomics of Mesolithic Scandinavia: Investigating early postglacial migration routes and high-latitude adaptation. PLoS Biol. 16, e2003703 (2018).
55. S. Brace et al., Ancient genomes indicate population replacement in Early Neolithic Britain. Nat. Evol. 3, 765–771 (2019).
56. A. Seguin-Orlando et al., Genomic structure in Europeans dating back at least 36,200 years. Science 346, 1113–1118 (2014).
57. K. R. Veeramah et al., Population genomic analysis of elongated skulls reveals ex-

58. M. A. Yang et al., 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. Curr. Biol. 27, 3202–3208.e9 (2017).
59. M. Rasmussen et al., The genome of a Late Pleistocene human from a Clovis burial site in western Montana. Nature 506, 225–229 (2014).
60. G. McInnes et al., Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent. Science 350, 820–822 (2015).
61. L. Ji et al., A resource-efficient tool for mixed model association analysis of large-scale data. Nat. Genet. 51, 1749–1755 (2019).
62. L. Beria, J. K. Pickrell, Approximately independent linkage disequilibrium blocks in human populations. Bioinformatics 23, 283–285 (2006).