Genetic and climatic factors in the dispersal of Anatomically Modern Humans Out of Africa

Raymond Tobler[1,9,*], Yassine Souilmi[1,2,*], Christian D. Huber[1,10,*], Nigel Bean[3,4], Chris S.M. Turney[5], Alan Cooper[6,†], Shane T. Grey[7,8,†]

Corresponding authors: R.T. (raymond.tobler@adelaide.edu.au), Y.S. (yassine.souilmi@adelaide.edu.au), A.C. (alanjcooper42@gmail.com), S.T.G. (s.grey@garvan.org.au)

This PDF file includes:

Supplementary Information

Extended Data Figures 1-12

Captions for Extended Data Tables 1-8

Captions for Supplementary Data Figures 1-56

Other Supplementary Information for this manuscript include the following:

Supplementary Data Figures 1-56

Supplementary Information

1. Sweep Detection

1.1 Population designation

The assignment of individuals to 18 distinct populations was governed by minimizing temporal and spatial variability amongst samples previously defined in the ancient DNA literature as sharing genetic affinities (Fig. 1, Extended Data Fig. 1, Extended Data Table 1), while also preserving the archaeological context and retaining enough individuals to ensure that our analyses were sufficiently powerful to detect selection[7]. The population assignments in this study produce reasonably distinct genetic clusters in principal component analysis (PCA) space (Extended Data Fig. 2) that align with previously reported PCA results[4,50,51]. We use a population nomenclature that follows the guidelines recently proposed for ancient DNA research[52].

1.2 Selection terminology and evidence of selection in human history

The classical definition of a hard sweep is based on the fixation of a new beneficial allele and linked neutral variants (*sensu* ref. [53]). In contrast, soft sweeps capture scenarios where the beneficial allele is present prior to the onset of selection, such that the allele is present on multiple haplotypes which can potentially be picked up by selection. In the current study, however, our primary focus is the characteristic genetic signatures left by the fixation of a single sweeping haplotype – which is possible when the beneficial variant is rare – rather than the initial frequency of the beneficial allele *per se* (i.e. a *de novo* mutation vs standing variation). Hence, a hard sweep in this study refers to positive selection driving a single haplotype to sufficiently high frequencies to result in the characteristic distortion of the site frequency spectrum (i.e. creating the genomic signature that can then be detected by SweepFinder2) – a definition that also encompasses selection from rare standing variation[54].

Another form of selection, known as polygenic selection, involves simultaneous changes across multiple genomic loci that all contribute to the beneficial trait. This mode of selection tends to result in numerous subtle allele frequency shifts that are subsequently unlikely to produce distinctive sweep signals at individual beneficial loci[55,56] and be detected in our analyses.

1.3 Data collection and processing

To produce a robust dataset and avoid potential bioinformatic batch effects, the raw sequence read data for 1,162 ancient genomic datasets (Extended Data Table 1) was retrieved from the European Nucleotide Archive and processed through the following standardized pipeline. To minimize the risk of modern contamination, the forward and reverse reads of the paired-end reads were merged (collapsed) using fastp[57], and only merged reads were retained (modern data is more likely to comprise large DNA fragments that do not collapse). All collapsed reads were filtered for potential residual adaptor sequences and chimaeras using Poly-X with fastp[57]. The retained filtered set of sequence reads were aligned to the human reference genome (h37d) using the Burrows-Wheeler Aligner v0.7.15[58]. All mapped reads were sorted using SAMtools v1.3[59] and then realigned around insertions and deletions and potential PCR duplicate reads marked and removed using the Genome Analysis ToolKit (GATK) v3.5[60].

Prior to variant calling, all remaining aligned reads were screened and base-calls recalibrated for aDNA postmortem damage using mapDamage2[61]. To further limit the impact of postmortem damage on variant calling[62], bamUtil[63] was used to trim 3 base-pairs from each of the 5' and 3' ends of each mapped read. From the resulting set of reads, pseudohaploid variants were called at the set of 1240k capture SNPs[50] across the 22 autosomes, using a combination of SAMtools mpileup[64] and sequenceTools (https://github.com/stschiff/sequenceTools). The 1240k capture was developed to minimize ascertainment in non-African populations and was used to generate data for most samples used in the study, whereby concentrating on the 1240k variants ensured a common and robust set of variants for the subsequent analyses. Pseudohaploidization of read data is a standard strategy in aDNA analyses, whereby a single read is randomly sampled at each prespecified SNP position[50] in order to mitigate potential biases introduced by differences in coverage or post-mortem damage between samples[4]. The pseudohaploid variant calls were converted from EIGENSOFT format[65,66] to binary Plink format using EIGENSOFT. Plink v1.9[67,68] was used to assign samples to the predefined populations (Extended Data Table 1) and convert the variants to reference polarized VCF files, with correct polarization being checked using BCFtools[64]. Finally, a custom Python script was used to generate the site frequency spectrum (SFS) input files for SweepFinder2 analysis (https://gist.github.com/yassineS/fe2712ad52d76460b927e3f391ea51f6).

1.4 Sweep scans

For all 18 ancient and 5 modern human populations, we computed the SweepFinder2 CLR (composite likelihood ratio) statistic[9,69] across the genome in successive 1kb intervals. The CLR statistic evaluates evidence for hard selective sweeps in dynamically sized windows, by comparing the distribution of allele frequencies expected under a mathematical model of a hard sweep with the expectation under neutral evolution (larger CLR scores indicate more evidence for selection). The expected site frequency spectrum (SFS) under the hard selective sweep model is computed conditional on the neutral SFS, assuming a certain selection coefficient and recombination rate. The neutral SFS is based on the background SFS calculated from the whole genome, assuming that the influence of sweeps on the SFS on a genome-wide scale is negligible.

SweepFinder2 is robust to genome-wide effects such as ascertainment bias and demography[70] by allowing these processes to affect the expected SFS under neutrality (i.e., the background SFS). Further, unlike many other selection methods, the assumptions on the input data for SweepFinder2 are suitable for the low coverage and ascertained nature of ancient DNA datasets. Since it is only based on the spatial (genomic) pattern of allele frequencies but not on haplotype homozygosity or population differentiation, it is possible to detect selection without reference to a second population, calling genotypes, or phasing haplotypes. The reliance on an empirically estimated null model (i.e. the background SFS) and the model-based alternative hypothesis makes it both more powerful and more robust compared to alternate test statistics that are based on deviations of the SFS from expectations under the standard neutral model (e.g., Tajima's *D*, Fay and Wu's *H*).

Note that SweepFinder2 also has an option to detect sweeps based on local genomic reductions in diversity. However, we did not calculate this diversity-based metric since accurate and unbiased estimation of diversity requires full genome data, whereas our dataset consists of an ascertained set of SNPs.

1.5 Outlier gene detection

Human gene annotations were obtained from ENSEMBL database[71] (genome reference version GRCh37), which was accessed using the R bioMart package[72,73] (version 2.36.1). Of the 24,554 annotated 'genes' on the bioMart database, we removed any that were not annotated in the NCBI database

(ftp.ncbi.nih.gov/gene/DATA/GENE_INFO/Mammalia/Homo_sapiens.gene_info.gz) and also excluded those that lacked specific protein and RNA based annotations (in the bioMart transcript_biotype field). This resulted in a list of 19,603 genes, from which we removed 26 that did not contain any polymorphic sites in our datasets (all being situated in the most terminal areas of chromosomes), leaving 19,577 genes that were used in the subsequent analyses. For each population, all SweepFinder2 CLR scores were log_{10} transformed and assigned to each of 19,577 genes by binning the transformed scores within the genomic boundaries of each gene. The gene boundaries were extended by 50kb on either side in order to also capture *cis*-regulatory regions. Because this typically resulted in several scores per gene, we took the maximum score to represent the evidence for a sweep involving that gene. Each gene score was corrected for gene-length using a non-parametric standardization algorithm [74,75], resulting in the gene scores having an approximately standard normal distribution. Finally, *p*-values were calculated for all genes and a *q-value* correction[72] applied for each population. The *q-value* is a Bayesian posterior estimate of the *p*-value that accounts for the expected inflation of false positives due to multiple testing[72], whereby a *q-value* of 0.01 implies a false discovery rate of 1% per population.

1.6 Candidate sweep classification

Sweeps were identified by determining a set of outlier genes across all populations (above), which were classified into sweep regions according to 1) the distance between the mid-point of neighbouring pairs of outlier genes (*i.e.* inter-gene distance) and 2) overlapping sweep regions between populations. Step 1 was performed independently for each population, whereby all outlier genes with midpoints that were less than a specific distance apart from the midpoint of a neighbouring outlier gene were collapsed into a single category. After generating the collapsed categories for each population, step 2 was applied to ensure that the sweep categories sharing at least one gene across different populations were considered as a single historical sweep.

We ran our sweep quantification pipeline at three *q*-value thresholds (*i.e.* $q < 0.01$, 0.05 or 0.10; which imply false discovery rates of 1%, 5% and 10% per population, respectively) and the three different inter-gene distances (midpoint distances less than 250kb, 500kb or 1Mb). As expected, changing the *q*-value had a large impact on the number of sweeps (ranging from ~50 for $q<0.01$ to ~500 for $q<0.1$), whereas changing the inter-gene midpoint distance had comparatively little impact overall, particularly at more stringent *q*-value cutoffs (Extended

Data Fig. 8). Based on these results, we decided to use the most stringent q-value cutoff and the most liberal inter-gene distance to define a robust set of candidate sweeps for all further analysis. However, because this stringent cutoff might lead to the removal of potentially causal genes in a sweep (which could have values slightly lower than 0.01), we first defined our sweeps based on the more permissive $q<0.1$ threshold, and then removed all sweeps that did not have at least one gene with $q<0.1$.

To further improve sweep determination, we removed populations with small sample sizes from the sweep determination process (see section 1.7), as previous results of modern genomes suggest that SweepFinder2 has poor power to detect sweeps when less than 10 haploid genome copies are being analyzed[76].

Finally, to ensure that the sweeps being defined were all relevant to western Eurasian history, the two modern populations from East Asia (*CHB*) and Africa (*YRI*) were also excluded from the sweep classification process. This strategy resulted in a total of 57 candidate sweeps that were used in subsequent analyses. Plots showing the distribution of CLR scores and gene scores across all candidate sweeps are included in Supplementary Data Figures 1-56 (https://figshare.com/s/7d3711649677746bc231).

1.7 Sample size metric that accounts for missing data

Previous results suggest that SweepFinder2 has reduced power to detect sweeps in modern genomic datasets when fewer than 10 haplotypes are being analyzed[76]. Notably, the number of sweeps detected in the ancient populations in this study showed a marked decline as sample sizes decreased (Extended Data Fig. 9). To better capture the impact of sample size on ancient population genomic datasets, we introduce a new metric – the effective sample size (n_{eff}) – which is the product of the ploidy of each sample (k) with the sample size (n) and the average proportion of non-missing sites (M), that is:

$$n_{eff} = k * n * (1 - M)$$

This metric includes salient features of sample characteristics for ancient population genomic datasets that are not captured when simply computing the number of samples. For instance, an ancient population that comprises 10 pseudohaploid samples and 50% data missingness on average will has $n_{eff} = 1 * 10 * (1 - 0.5) = 5$. In comparison, a modern population containing 10 high coverage genomes has $n_{eff} = 2 * 10 * (1 - 0) = 20$, since it does not contain missing data and two alleles are identifiable at each position for each sample.

Previous work based on simulated Eurasian datasets demonstrates that SweepFinder2 has consistently high power to detect hard sweeps in Eurasian populations with reasonable sample sizes (i.e. $n_{eff} \geq 10$) and that lack admixture, but power is poor for admixed populations with low effective sample sizes [7]. These simulations also revealed that *p*-value distributions strongly deviated from the expected uniform distribution for populations with low n_{eff}, a feature that was also observed amongst the empirical populations with similarly low n_{eff} values [7]. Based on these results, we restricted our analyses of sweeps to populations with neff $\geq$ 10 in the present study.

1.8 Comparison with SweepFinder2 scans in modern African populations

The 57 sweep signals observed in ancient Eurasian populations were absent in a modern African population (e.g. Yorubans from the 1000 genome project; *YRI*), consistent with the underlying selection pressure most likely arising after humans left Africa. To explore the origin of the sweeps in more detail, we explored if any of our 57 sweep regions overlapped with positively selected regions identified in a recent study of high coverage genomes from multiple African populations[77]. We compared our 57 sweeps with the 492 candidate loci from six different African populations, identifying nine unique candidate loci that overlap four of our sweeps (namely, *DOCK3*, *GABBR1*, *PNLIPRP3*, *MIR662*). However, none of the nine overlapping candidate loci from ref. [77] were found to be enriched for SNPs with high iHS scores (i.e. the top 1% of all SNPs with iHS scores from each of the six tested African populations), suggesting that these loci might be false positives.

We further tested if the four selected regions shared by our study and ref. [77] were consistent with random expectations, which would further suggest these overlaps could be artefactual (for instance, representing separate episodes of selection or false positives in either study). To do this, we used the R package regioneR[78] to randomly sample 57 genomic windows from across the genome that matched the sizes of our sweeps, then counted the number of sweeps that overlapped at least one of the 492 candidate regions from ref. [77]. To obtain an empirical *p*-value we repeated this sampling process 10,000 times and calculated the proportion of random samples that had an equal or larger number of sweeps overlapping African candidate regions than we observe (i.e. four unique sweeps). We also repeated this process to account for the possibility that both our hard sweep signals and the reported outlier windows were more likely to be detected regions with low recombination rates (which may have more power in genome-wide selection scans), by sampling from a modified genome that had been

rescaled according to the local recombination rate. Specifically, we used the sex-averaged recombination rates measured across 10kb genomic intervals[79] to rescale the length of each interval (l_i) according to $l_i = 1/(1+r_i)$, where r_i is the recombination rate for the ith genomic interval. We also rescaled each sweep window through linear interpolation to the recombination-rescaled genome, and used regioneR[78] to randomly sample 10,000 sets of 57 genomic windows matching these lengths from the rescaled genome. The sampled windows were then converted back to a standard physical length by interpolating to the original (unscaled) genome, and the overlap with outlier regions tested in the standard way. Indeed, both sets of permutation tests were returned non-significant results (Standard test: expected number of overlapping sweeps ~2.92, $p = 0.33$; Recombination rate adjusted test: expected number of overlapping sweeps ~3.32, $p = 0.3986$), consistent with the number of overlaps between the two studies matching chance expectations.

2. Detecting and Dating Sweep Haplotypes

2.1 Determining marker alleles for sweep haplotypes

To track the frequency of the sweep haplotype through time, we developed a heuristic method to determine a set of marker SNP alleles for each of the 57 sweeps, and then used these markers to quantify the presence of the inferred sweep haplotype in modern and ancient human samples.

Our analyses suggest that the underlying selection pressure driving the 57 sweeps most likely arose after humans left Africa (see section 1.8). Given that this would result in SNPs carried on the sweep haplotype becoming highly divergent when contrasting European and Yoruban populations, we sought to identify a set of up to 30 marker SNPs in each sweep that were consistently highly divergent in pairwise comparisons between Yorubans and Eurasian populations. Specifically, we used the derivation of Hudson's F_{st} estimator from ref. [80] (equation 10) to calculate pairwise F_{st} in all Yoruban-Eurasian pairs across all ~1M SNPs in the genome. For each pairwise test, we excluded SNPs that did not have at least 10 copies in each population (this condition was always satisfied for the Yoruban population, so was dependent on the number of samples and SNP missing-ness in the relevant Eurasian population). For each sweep, we then averaged the F_{st} values after omitting values from Eurasian populations that did not show evidence for the sweep signal (i.e. populations that had no gene with $q < 0.2$ across the specific sweep region). In addition to estimating the mean F_{st} for SNPs within sweep regions, we also computed this statistic for all SNPs occurring in

regions of the genome lying outside of the 57 sweep regions. These SNPs provide a conservative 'neutral' background (conservative because it will also include SNPs from regions impacted by positive selection that are not detected by our selection scan), which we used to calculate thresholds to identify the most strongly differentiated SNPs in each sweep region. Accordingly, we excluded all SNPs in sweep regions with mean F_{st} values below the 80th percentile of 'neutral' mean F_{st} values (the threshold mean F_{st} value ranged from ~0.2 to 0.27 across all sweeps, and tended to increase as an inverse function of the number of Eurasian populations used to compute the mean). The remaining SNPs were ranked according to the mean F_{st} and up to 30 with the lowest ranks (*i.e.* highest mean F_{st}) were chosen as marker SNPs.

After identifying the set of marker SNPs for each sweep region, for each marker SNP we determined the putative allele linked to the sweep haplotype as the most common (major) allele across the same subset of European populations that were used to calculate the mean F_{st}. Note that prior to ranking the SNPs, we also excluded SNPs where the major allele was not the same when the mean frequency was estimated from raw population frequencies, and when frequencies were weighted by the number of samples in each population. This step ensured that the major allele was always unambiguous according to these two estimators. Applying this procedure to each sweep resulted in a unique set of up to 30 marker alleles. The position of the marker SNPs in the sweep is shown in Supplementary Data Figures 1-56. To ensure that sweep haplotype estimation was robust to SNP missingness, only samples with at least 50% of marker SNPs in a given sweep were evaluated. Note that one sweep (*LINC01153*) was excluded from further marker-based analyses as it only had 3 SNPs that passed all marker filtering criteria, whereas all other sweeps had ≥6 marker SNPs and the majority (46/57; ~81%) had 30 or more marker SNPs.

2.2 Sweep haplotype presence in ancient and modern samples

To establish when the selection pressures underlying each sweep may have first arisen, and their subsequent dynamics up to the present day, we quantified the presence of the sweep in a combined dataset of ancient and modern samples. We obtained recently published samples from the Eurasian Upper Paleolithic – namely, Zlaty kun[81] in Czechia and six specimens from Bacho Kiro[21] in Bulgaria – and combined these with 223 additional ancient samples curated in the Allen Ancient DNA Resource (AADR; version 42.4), a collection of ancient, archaic, and modern human genomes curated by the Reich Laboratory in Harvard University

(https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data). We limited our investigations to key ancient samples from the Eurasian Upper Paleolithic period and all Eurasian samples at least 10,000 years old where at least 80% of the ~1.1M probe markers have a pseudohaploid calls (which resulted in 181 samples), and also included historic and ancient samples from Africa (12 samples), the Americas (26 samples), and Oceania and the Andamans (5 samples) that also had ≥80% of the ~1.1M marker SNPs present (see Extended Data Tables 3, 6). For the modern populations, we took all 2,504 samples from phase 3 of the 1000 Genomes Project (1KGP)[10] also made available in the AADR v42.4, providing a broad representation of modern human populations distributed throughout Africa, Europe, East Asia, South Asia, and the Americas. To provide additional representation for the ancient migrations that brought modern humans into Island South East Asia and Sahul, we also included 28 phased Papuan samples from the recent Human Genome Diversity Panel (HGDP)[14], along with two modern Aboriginal Australian samples that are also included in the AADR v42.4.

Alignment (i.e. BAM[59]) files for the Zlaty kun and six Bacho Kiro specimens were downloaded from links provided in each paper and pseudohaploidized according to the protocols outlined in section 1.3 of the Methods. For the AADR samples, we either used the pseudohaploid calls provided or, for samples with diploid calls (i.e. all modern samples and high coverage ancient samples such as Ust'-Ishim), randomly selected a single allele at each SNP to mimic pseudohaploidy (see section 2.3 for details on the impact of pseudohaploidization on our sweep haplotype inference procedure). Two of the Bacho Kiro samples (BachoKiro_AA7-738 and BachoKiro_CC7-2289) had too few sites for robust haplotype detection and were removed from subsequent further analyses.

For each ancient and modern sample, we defined the sweep haplotype as present if a certain minimum proportion of marker alleles were observed in the sample. Specifically, to determine sweep presence, we tested different combinations of the number of marker SNP and the minimum proportion of marker alleles observed – either 90% or 95% of the top 10, 20, or 30 ranked marker alleles – creating six categories overall. Finally, we determined sweep presence based on the aggregate signal across all six categories, requiring the sweep to be present in at least three of the six categories. We chose this intermediate value as it provides a reasonable balance of the false positive and false negative rate (see section 2.3 for details). Finally, several key ancient Eurasian samples were re-dated using the recent

IntCal20 calibration curve[82] to obtain improved estimates of their antiquity (Extended Data Table 7).

There is a strong male bias in the available genomic sequences older than the Holocene that are of sufficient coverage for analysis (20 of 23 samples, 87%, Extended Data Table 7, Extended Data Fig. 10). Male sex biases have been detected in Pleistocene megafaunal remains and have been related to behavioural and ranging differences between male and female individuals, particularly young males, and resulting changes in the probabilities of preservation and discovery in the environment[83]. However, the vast majority of the 23 pre-Holocene AMH samples were found in human occupation sites (caves, rockshelters, camps) and range widely in geographic location (from western Europe to Beijing) and temporal age (across 35,000 years). This broad range helps provide a reasonable survey despite the small sample size. In contrast, there are 9 available hominin genomic datasets that pass our quality filtering criteria for sufficient genomic coverage (one Denisovan, one hybrid Denisovan-Neandertal, seven Neandertal), and of these only one Neandertal specimen is known to be male, although more are known[84]. We cannot see any reasons that a pronounced male bias in the older part of our AMH dataset should noticeably influence our analyses or conclusions.

2.3 Statistical properties of the sweep haplotype detection method

An important consequence of the pseudohaploidization process is that it will tend to underestimate the true sweep haplotype frequency in populations where the sweep is at intermediate frequencies. To quantify the impact of pseudohaploidy on our estimation procedure, we quantified the presence of each sweep using the phased 1KGP modern datasets, and also after randomly sampling an allele to mimic the pseudohaploidy of ancient samples. As expected, the frequency of the sweep haplotype was consistently underestimated for pseudohaploid samples, and underestimates were proportionate to the level of heterozygosity of the sweep haplotype in that population (underestimating the sweep frequency by as much as 20% when the true frequency is around 50%; Extended Data Fig. 11).

At the level of individual samples, pseudohaploidization will not affect sweep detection in samples that are homozygous for the sweep haplotype (i.e. sweep homozygotes), even in ancient samples where information from some of the marker SNPs may be missing. While errors in allele calling or recombination could potentially result in other alleles appearing at

the marker SNPs, the potential impact of these factors will be diminished to some extent because our detection method does not require all SNPs to be present. Thus, we expect that fixed sweeps will be consistently detected across ancient and modern samples, and the patchy patterns observed across ancient samples implies that no sweep was fixed prior to the colonisation of Eurasia ~50,000 years ago, although *DOK5* and *LIN28B* appear to have been in very high frequency in the Oceania and IUP samples.

While false negative calls in pseudohaploid samples are unlikely when the sweep is fixed in a population, ancient samples that carry two non-sweep haplotypes (i.e. non-sweep homozygotes) can still be falsely called as sweep carriers (i.e. false positives) – this will depend on the number and informativeness of missing marker SNPs and the similarity of the alleles on non-sweep and sweep haplotypes. Further, our previous results suggest that ancient samples that carry one sweep haplotype and one non-sweep haplotype (i.e. sweep heterozygotes) will tend to called as non-sweep carriers, but this will be ameliorated to some extent by the factors that impact false positive calls (i.e. the patterns of SNP missingness and allelic similarities between sweep and non-sweep haplotypes).

Accordingly, we quantified the impact of pseudohaploidization on the sweep haplotype calling rate for non-sweep homozygotes (i.e. the false positive rate; FPR) and also for sweep heterozygotes in ancient samples. To estimate the FPR, for each sweep we randomly sampled 2,000 non-sweep haplotypes from the 1KGP phased samples. We then randomly paired these haplotypes to create 1,000 diploid samples, and sampled a single allele for each marker SNP to generate pseudohaploids. Finally, we set specific marker SNPs to missing, according to the SNP missingness observed in the ancient samples and modern samples that were important for inferring the timing of the onset of the selection pressure for each sweep (see section 2.2 for details on our selection dating inference). To measure the calling rate for sweep heterozygotes, we repeated the process outlined above after randomly replacing one of the non-sweep haplotypes with a sweep haplotype.

Following the same test criteria we use to detect sweep haplotypes – the sweep is called in at least three of the six possible tested categories (i.e. combinations of 90% or 95% of the top 10, 20, or 30 marker SNPs) – the FPR is generally below 10% for any combination of sweep and sample (Extended Data Fig. 12). As expected, the sweep detection rate is systematically elevated in sweep heterozygotes (i.e. when replacing one of the non-sweep haplotype with a sweep haplotype) – and this is strongly positively correlated with the false positive rate

(Extended Data Fig. 12). This positive correlation appears to be largely the result of non-sweep haplotypes being more similar to the sweep haplotype, as this increases the probability that alleles sampled from the non-sweep haplotype will match the sweep marker allele.

Taken together, our results indicate that we will tend to underestimate the sweep presence in ancient samples. Based on the criteria used in the study (i.e. the sweep is called in at least three of the six possible tested categories), only around 26% of sweep heterozygotes will be called as sweep carriers on average (Extended Data Fig. 12). Because the FPR is comparatively lower (~6%; Extended Data Fig. 12) under the same criteria, on balance we are more likely to miss sweep haplotypes in heterozygotes rather than to falsely call them in non-sweep homozygotes. Further, while we do not know what the true sweep haplotype frequencies were historically, the patchy signals observed across ancient Eurasian samples in the current study suggest that the majority of sweep haplotypes were segregating at intermediate to high frequencies for much of the Upper Paleolithic period.

2.4 Recalibration of age of Neandertal admixture

Genomic data reveals that Neandertal admixture with AMH has been occurring for some time. For example, a limited early admixture event between Neandertals and an early lineage related to modern AMH (possibly around 370 ka, although this is unclear) appears to have resulted in the replacement of the Neandertal Y and mitochondrial genomes[84]. This admixture event may have involved the late-surviving Middle Pleistocene fossil *Homo* group recently identified in the Levant[85]. However, modern global non-African populations contain Neandertal DNA resulting from a much more recent admixture event between Main Eurasians and Neandertals around 50-60ka, thought to have occurred in the Near East immediately prior to the dispersal of AMH across Eurasia. The timing of this admixture has been studied using the size of the Neandertal genomic fragments in the Ust'-Ishim AMH specimen from western Siberia, whose age and high quality genomic data provide a good calibration point close in time to the events. The Ust'-Ishim femur has been radiocarbon dated at (41,410 +/- 960 BP), calibrated using IntCal13 at 44-45.8 (68% CI), or 43.2-46.8ka cal BP (95% CI)[20,86]. A recombination clock suggested that 232-430 generations would be required to produce the size distribution of introgressed Neandertal fragments in the Ust'-Ishim genome, under a simple model of a single pulse of Neandertal admixture. This would give an admixture date of 50-60ka[20,86]. This range was constrained by more detailed

modelling which indicated the credible range of peak Neandertal geneflow was 7,521 +/- 854y (260 +/-30 generations) earlier than Ust'-Ishim (equating to 50-55ka), again under a single pulse model[20,86]. However, two distinct size classes of Neandertal fragments were identified in the Ust'-Ishim genome, such that a model of two admixture pulses was much more strongly supported[20,86]. The two pulse model indicated the Ust'-Ishim genome Neandertal content resulted from a major early admixture event (estimated at 6,660 +/- 218y, or ~228 +/- 21 generations before Ust'-Ishim, or 50-54ka), along with a minor secondary event that occurred considerably later (1,258 +/- 113y, or ~43 +/- 4 generations before Ust'-Ishim, or 44-47ka). The latter event was potentially restricted to just the Ust'-Ishim lineage, and took place several millennia after the dispersal of AMH groups across Eurasia.

Recent modelling suggests that the admixture date estimates may be sensitive to non-homogenous recombination rates, and have difficulty discerning a single pulse of admixture from ongoing geneflow over potentially hundreds of generations[87]. As a result, the 50-54ka estimate should be regarded as a credible range for the mean time of Neandertal geneflow rather than an absolute range, as this may have started earlier.

A simple means to further refine these temporal estimates is through re-calibrating the radiocarbon date of the Ust'-Ishim specimen, using the recent IntCal20 calibration curve, which has considerably improved the calibration accuracy in this time period. The IntCal20 recalibration results in a revised calibrated date for Ust'-Ishim of 42.9-46.0ka cal BP (68% CI)[82] (Extended Data Table 7). Applying the same recombination clock values discussed above, the credible range of the mean time of Neandertal geneflow can therefore be re-calibrated at 49.5-52.7ka (~50-53ka) under the two-phase admixture model (or 49.6-53.7ka under the less supported single-phase model).

The younger bound for the timeframe of the estimated Neandertal geneflow agrees with, and is presumably also constrained by, the AMH colonisation of Australia given that indigenous Australian and New Guinea populations appear to possess the standard Main Eurasian ~2% Neandertal signal, and show no signs of genetic input after their initial colonisation of the Sahul continent (Australia/Tasmania, New Guinea and related islands)[88–90]. The colonisation of Australia has been estimated by a number of genetic and archaeological studies at being close to 50 ka or slightly earlier, with a range of dated archaeological sites rapidly appearing across Australia between 50-47ka. In contrast, dates much earlier than 52ka have been called into question[88–90].

The presence of Neandertal genomic content in Aboriginal and New Guinea populations at comparable levels to other Out of African AMH populations suggests that their ancestral lineage diverged from other Main Eurasians after the majority of Neandertal admixture had occurred. Hence, while it is possible that Neandertal admixture with the Main Eurasian lineage may have been ongoing (potentially for many thousands of years) prior to the ~50-53ka estimates of the credible range of the mean time of geneflow[87], it seems likely that the Main Eurasian dispersal (including Ust'-Ishim and Oceanians) had not occurred prior to this date. In contrast, the colonisation of Australia indicates it had definitely occurred by 50ka.

The movement from the Arabian Peninsula area to Australia/New Guinea would presumably take some time as this represents an overland distance of around 18,000 km, and required major marine voyages through eastern ISEA, even with increased land exposure due to ~50m lower sea levels[15,42]. Allowing one millenia for this movement as a relatively conservative estimate (*i.e.* an average of 18km/yr, which is lower than some suggestions for hunter-gatherer movement through new territory[91]) means that the Neandertal content of the Sahul populations probably further constrains the main admixture to older than ~51ka. As a result, we assume a date of ~51-53ka for the Main Eurasian dispersal as our best estimate and use this in Fig. 1, although a more conservative ~50-54ka date (or even 50-60ka) would not change our conclusions.

One interesting implication of an inferred dispersal of AMH populations from the Arabian Peninsula to New Guinea between ~51-53ka is the rapid nature of the movement, which has been noted previously and has raised questions about the apparent speed of progress through ISEA given that over-the-horizon marine voyaging for hundreds to thousands of individuals was required[92]. While archaeological evidence during the AS is limited, the potentially prolonged restriction (e.g. from ~80-55ka) of AS populations to coastal sites or the Arabian Gulf area has been used to suggest a potential development of nascent marine technology[93]. This could provide a potential explanation of the apparent lack of significant delay for the development of marine technology during the ISEA migration, addressing debates about the extent to which purposeful voyaging and boat construction was possible during this event[92]. Interestingly, in contrast to the rapid eastwards movement to Australia, the colonisation of Europe appears to have been comparatively delayed until ~47ka[43,94] possibly due to the cold conditions of Greenland Stadial 13/Heinrich Event 5 (Fig. 5), and the requirement for further

genetic adaptation to cold conditions, first seen as the sweeps appearing in the IUP specimens (Fig. 1, Extended Data Tables 3, 6).

2.5 Location of Neandertal admixture event and Main Eurasian Dispersal

While the geographic location of the inferred Neandertal-Main Eurasian admixture remains unclear, it is likely this took place near the Arabian Peninsula as essentially all modern and ancient non-African individuals outside of this area appear to carry introgressed DNA from this admixture. (Lineages descended from the Basal Eurasians are an exception). The apparent consistency of the amount and distribution of introgressed Neandertal DNA in modern human populations suggests that the initial admixing Main Eurasian population was both relatively small, and also stayed together for some time either during or after the admixture to allow for the homogenisation of the Neandertal genetic contribution prior to the dispersal across Eurasia. Indeed, some studies have estimated the starting Neandertal genomic content in the admixed AMH population may have been as high as 10%[95], although it has been noted that this value would be influenced by the nature and timing of the Neandertal geneflow[87].

Early AMH groups appear to have moved out of Africa (OOA) and broadly throughout the Arabian Peninsula by early Marine Isotope Stage 5, around 125ka (Figs. 1, 5)[96,97]. Initial sites tend to be more coastal, such as the Levant (Skhul and Qafzeh) and the Arabian Sea (Jebel Faya). Recent fossil discoveries in the Levant have raised the potential that 140-120ka Levant populations might have also represented a separate, late-surviving, group of Middle Pleistocene *Homo* that may have interacted with initial AMH populations moving OoA[85]. Occupation throughout the interior of the Arabian Peninsula was associated with short periods of relatively warm and moist climatic conditions around 120, 100 and 80ka (Figs. 1, 5), when rivers and lakes were apparent[35]. This has led to suggestions that lowered sea levels during glacial conditions facilitated AMH movement out of Africa (OOA) across the Red Sea but that arid and mountainous regions to the north, such as the interior of the Arabian Peninsula, prevented further colonisation until ensuing moist spells[35,96,97]. As a result, the predominantly cold arid conditions between 110-60ka (Fig. 5) could explain the limited signs of AMH dispersal much further than the Arabian Peninsula during this period, and archaeological data suggesting population isolation and primary occupation of coastal sites, potentially including the largely exposed Arabian Gulf [93]. The lower Arabian Gulf would have been largely exposed by lowered sea levels, and this potentially fertile area is likely to

have featured numerous subterranean water sources which would have been an important site for AMH populations in the cold arid conditions between 120 and 50ka, and potentially up until the Holocene[93]. The long term AMH occupation of this area has been described as the Arabian Gulf Oasis model[93].

A demographic scenario where a small AMH population remained homogenous during a period of geneflow with local Neandertals, and which had also remained genetically isolated for a long period between ~80ka (the estimated origin date for the hard sweeps, and approximate divergence time with Basal Eurasians) to ~53-51ka (the Main Eurasian dispersal) is consistent with the Arabian Gulf Oasis model, hence this is shown (in simplistic fashion) in Fig. 1. While the geographic source of the Neandertal admixture is unknown, the location depicted in Fig. 1 corresponds to the Zagros Mountains of Iran, which are a known concentration of Neandertal individuals geographically close to the Arabian Standstill population. Other potential areas include the Levant and the Caucasus regions. It is possible that the more cold-adapted Neandertal populations in these regions expanded further southwards during cold periods, and retreated northwards during warmer periods while the warm-adapted AS population potentially mirroring these movements further to the south. In this way, effective contact between the populations may have been minimised despite the close geographic proximity and availability of corridors such as the Euphrates river valley between them. The timing of peak Neandertal geneflow into the Main Eurasian population is estimated to have been around 53-51ka (see above), although it possibly had been ongoing at lower rates over a much longer time period prior to this period[87].

The migration routes indicated in Figure 1 are a synthesis of current archaeological and genetic data, but have been significantly simplified for visual clarity. This includes the initial movements Out of Africa and around the Arabian Peninsula ~120ka, and into the interior around 80ka (shown as resulting in the separation of the Basal Eurasian populations; Figs. 1, 5). Basal Eurasian genetic ancestry is present in groups such as the Bedouin, and is currently most common in northern Iran, although a movement into this area is not shown in Fig. 1 for visual simplicity. For similar reasons, the colonisation of Europe is simplified and shown as three pulses, with the Initial Upper Paleolithic (IUP) and (Proto)Aurignacian entering via the southeast near Bulgaria, and the Gravettian diverging from the Aurignacian. Tools at IUP and Proto-Aurignacian sites appear to represent dispersals from at least two cultural phases from the Levant area (Emiran and Ahmarian), consistent with the divergent genetic backgrounds including the hard sweep signals. Based on the shared hard sweep heritage, the Gravettian is

shown as diverging from the Proto-Aurignacian lineage, and crossing the exposed Black Sea shelf before appearing for the first time in the Sunghir specimens, shortly after 35ka. While the earliest Gravettian sites appear in the mid-upper Danube and Crimea, it was not possible to depict this detail.

Similarly, the movement into eastern Eurasia has been simplified to show a likely northern IUP route resulting in the Tianyuan lineage, and separate southern routes into Oceania and Southeast Asia. The proposed route through lower southeastern Asian area (Thailand, Vietnam, southern China) follows river courses through the savannah habitat predicted to exist in the area at the time[42]. Similarly, early AMH movement throughout the broader South East Asian area has been proposed to track savannah habitat which presumably was similar to the African ancestral ecosystems to which these groups had originated. Inferred areas of hominid genetic introgression[42] are depicted, but are likely considerable underestimates of the true number of events. While these movements on Fig. 1 are necessarily simplistic depictions of far more complex processes, they are designed to show the key points of the model.

2.6 Classification of sweep origins

The spatio-temporal patterns of the sweeps provide a view on the timing and location of selective pressures, and potentially also carry information about broadscale historical migration patterns during the Upper Paleolithic, which remain poorly understood. Because the combination of pseudohaploidy and missing SNP data will make detection of sweep haplotype presence in single ancient samples less reliable, we combined sweep calling results from multiple ancient and modern samples to create temporally-consistent consensus signals. To identify the earliest group of sweeps we examined isolated populations thought to be formed from the initial Main Eurasian population movement that colonised ISEA and Sahul (i.e. the former landmass comprising modern Australia and New Guinea) ~50ka (which we term the Oceanians), shortly after the main Neandertal geneflow ~53-51ka (above). If any earlier AMH groups were present in the southeast Asia area they do not appear to have contributed genetically to modern populations, as the latter all feature the standard Main Eurasian ~2% Neandertal genome content, along with additional proportions of Denisovan genomic DNA[15,42]. In this regard, recent physical dating studies have undermined many putative sites used to support the much earlier presence of AMH in Asia, as opposed to a number of archaic hominid groups such as Denisovans, and concluded that there is currently no convincing evidence of AMH presence prior to ~50ka[15,16]. We searched for the presence of

sweeps within genomic data from ancient and modern individuals of the Andaman Islands (ref. [98]; n=1), Australia (ref. [99]; n=2), Vanuatu (refs. [99,100]; n=4) and Papua New Guinea (ref. [14]; n=28; comprising individuals from the Highlands, Sepik, and Bougainville populations, using a 35% frequency of the sweep loci within any of these three groups as the cut-off for detection). While the number of individuals surveyed in several of the Oceanic groups was small, the combination of the relative genetic isolation of these groups following the initial AMH colonisation, and the four distinct sampling areas across a broad geographic coverage of Oceania, is expected to provide a reasonable estimate of the sweep haplotypes that were present at moderate frequencies the initial colonising population.

We identified the presence of 31 sweeps (Extended Data Table 4) within the Oceania/Andamanese samples, with 28 being detected in multiple groups or multiple individuals (90%), and just three detected in only one individual (1 each in Vanuatu, Australia and the Andaman Island specimen). The number of sweeps detected in Australia (21), Papua New Guinea (20), and Vanuatu (22) was relatively constant (i.e. around 68% of the total) despite the very different number of samples in each area (from two in Australia versus 28 in PNG), in contrast to 12 within the Andaman Island sample. The latter may relate to the high false negative rate for heterozygote loci in the single sample available for this area, and/or the very small long-term population size of the Andaman Islands leading to higher rates of genetic drift. The number of shared sweeps between each Oceanic area is consistent their known phylogenetic relationships, with the greatest similarity between Papua New Guinea and Vanuatu (16 shared/26 total, or 62%), followed by Australia and Vanuatu (16/27, 59%) and Australia and Papua New Guinea (13/28, 46%), with the most distant being Andaman Islands versus the others (Vanuatu 9/25, Papua New Guinea 9/23, Australia 8/25, average 35%) (Extended Data Table 4).

These patterns suggest that while relatively few of the sweeps were fixed within the initial AMH population moving into ISEA (potentially *DOK5*, *LIN28B* and *GTSE1*), many were at high frequency and as a result 23 (74%) have ended up in more than one geographic area. Due to the 31 sweeps being present immediately after the Main Eurasian dispersal, but not being of Neandertal origin (see section 3.9), we designated them as originating during the AS and therefore likely originating somewhere between ~80-50ka (Fig. 1).

The oldest group of AMH samples are those of the Eurasian Initial Upper Paleolithic (IUP) groups which were broadly distributed from eastern Europe, across eastern Eurasia to

Mongolia. The origin of these groups appears to be within the Levant Emiran culture[101]. The four available IUP genomic sequences were from Bacho Kiro, Bulgaria (*n*=3, CC7-335, BB7-240, F6-620, dated 45-43ka), and Zlaty kun, Czechia (*n*=1, with poor genomic coverage and an approximate date of 45ka). The Bacho Kiro genomic studies also demonstrated a close relationship between the eastern Chinese Tianyuan specimen (40ka) and the European IUP specimens[21], and this is consistent with the pattern of shared hard sweeps. While the three Bacho Kiro specimens shared an average of 17% of their hard sweep signals with each other (reflecting the small number of samples, and lack of fixed loci in the Main Eurasian dispersal), the Tianyuan specimen shared an average of 19% with each individual, and 11 of the 32 (34%) hard sweeps observed across all of the Bacho Kiro specimens. This pattern suggests a much more diverse origin than the Oceanian samples from one another.

While the other very old specimen, Ust'-Ishim (45ka), had no associated archaeological evidence so could not be associated with the IUP, genomic data suggests it diverged from the IUP genetic lineage or perhaps separated from the Main Eurasians shortly before the IUP group[21]. In this regard, while the Ust'-Ishim specimen only had eight hard sweeps, seven were shared with the IUP or Tianyuan specimens, and seven with the Oceanian samples. However, Ust'-Ishim possessed only one of the IUP-Tianyuan specific sweeps (*LINC00293*), suggesting a separate origin. As a result, it has been depicted with a dashed line on Fig. 1 indicating these uncertain origins. Interestingly, one hard sweep (*DOK5*) was fixed across all the Oceanian, IUP (including Tianyuan) and the Ust'-Ishim sample. This appears to be the closest to a fixed loci in the AS/Main Eurasian dispersal detected in the study, although it was subsequently absent in some of the younger European samples.

The four IUP samples at Bacho Kiro and Zlaty kun[21,81] were published after the first phase of analyses in this paper had been completed. As a result of their early age within the Main Eurasian dispersal, they effectively provide an independent test of the approach of using the Oceania samples to define the sweeps occurring within the AS. The IUP samples (including Tianyuan) contained 24 of the 30 (80%) AS hard sweeps defined by the Oceania samples, which is consistent with the small number of samples in both datasets, and elevated false negative detection rate. The remaining six AS hard sweeps absent from the IUP were detected in the next youngest Eurasian samples, suggesting they were indeed present in Europe/Eurasia but simply remained undetected in the IUP dataset.

In total, the IUP/Tianyuan/Ust'-Ishim specimens (45-40ka) contained eight hard sweeps that were not detected in the Oceanian samples (*CCDC7, DHODH, DOCK3, LINC00293, METTL6, PAX1, WWOX, ZMYM6*) and so we assume these originated on the Levant (Emiran)/West Eurasian/IUP lineage somewhere between ~51ka and 45ka (or potentially as late as 40ka for *WWOX* and *CCDC7,* which are first detected in the Tianyuan specimen). It is also possible that some of these hard sweeps may have been present in the AS but simply were not present in the founders of the Oceanian dispersal, or were lost by drift early in the process.

Following the IUP/Tianyuan/Ust'-Ishim group, a distinct group of nine new hard sweeps are observed for the first time in two slightly younger western European samples, the 38ka Kostenki14 specimen in western Russia (*ANKS6*, *BAIAP2L1*, *DNAH7*, *FANCD2*, *TBC1D7*) and 35ka Goyet 116 specimen from Belgium (*BAIAP2L1, DNAH7, ELMO1, ENAM, FANCD2, OIT3, MARS*). These individuals are associated with the Aurignacian Upper Paleolithic culture (~43-35ka[22]), one of the earliest pan-European technocomplexes which evolved out of the widely distributed Proto-Aurignacian AMH groups, who were the genetic successors to the IUP populations[21]. Both of these European Early Upper Paleolithic populations appear to result from distinct dispersal waves stemming from the Levant area. The initial IUP (ca. 47-43ka, expressed as localised cultures e.g. Bohunician in Ukraine) was apparently related to the Levant Emiran culture, while the following more widespread Proto-Aurignacian cultures (ca 43-41k) descended from a separate movement related to the Levant Ahmarian culture[101]. The initial movements into Europe start around 47ka (immediately following the extended severe cold periods of GS13/Heinrich Event 5) and appear to enter through southeastern Europe, where the earliest archaeological evidence in Bulgaria and Romania suggests a point of entry. The simplistic pattern shown in Fig. 1 depicts two separate movements descended in some fashion from the stem lineage of western Main Eurasians originating from around the Levant area.

Only two samples with sufficient genomic coverage are available for the Aurignacian culture, although the geographic sampling is broad (Belgium and western Russia – shown by the large oval in main text Fig. 1). Given the small number of samples, it is perhaps surprising to see nine sweeps appear for the first time in this group. However, we can be reasonably sure this is accurate due to the widespread sampling of the period immediately after the AS, comprising more than 41 samples (i.e. IUP/Tianyuan/Ust'-Ishim, Oceania/Andamans samples) in which these nine sweeps are not detected. The distinct pattern of Aurignacian hard sweeps is

consistent with genomic studies that indicate a complete population replacement between the IUP and Aurignacian[43]. Therefore we assume that the above nine sweeps occurred somewhere in the period ~51-38ka in the Levant (Ahmiran) or Western Eurasian Proto-Aurignacian populations during the movement into southeastern Europe. This is marked by the number 3 box in Fig. 1.

Genomic analyses suggest the European Aurignacian technocomplex was relatively rapidly replaced ~35ka by the Gravettian technocomplex (which spanned from 35ka to ~25ka[22,23,102]), and which is suggested to have originated around the Danube River Valley and initially spread across eastern Europe. There are a number of ancient AMH samples during the Gravettian with sufficient genomic coverage for analysis, notably Sunghir 1-4 (n=4) and Vestonice 16 (n=1), along with other specimens in the Gravettian time period from northern Siberia (*e.g.* Yana, n=2). The potential route for this group is shown in Fig. 1 and reflects the location of Sunghir (and Crimea) in western Russia as the earliest sampling points. While only one hard sweep (*SMCO2*) is detected for the first time in the Gravettian individuals (in the oldest Sunghir specimen, 34.6ka), the Gravettian specimens have a distinct pattern of hard sweeps that appears to be a combination of those seen earlier in the IUP and Aurignacian individuals (Extended Data Table 3). For example, four of the eight IUP-unique hard sweeps, and five of the nine Aurignacian hard sweeps appear within the Gravettians. This pattern is consistent with genomic analyses that have suggested that Sunghir individuals and Vestonice-16 share a significant amount of ancestry with the Aurignacian Kostenki14 and Goyet 116 lineages[103]. The pattern of hard sweeps is similar to that seen in the Yana individuals, although the latter have been identified as a local regional culture.

The sweep signals suggest that while the Gravettian represented a significant new genetic input, the source lineage was related to that of the Aurignacian lineage and the two had perhaps not been diverged for long as only one new hard sweep (*SMCO2*) is detected (Fig. 1, Extended Data Table 3). An alternative model is that a large degree of genetic admixture occurred between the early Gravettian and late surviving Aurignacian individuals, and that on-going selective pressure was sufficient to return the Aurignacian sweeps to a high enough frequency in the Sunghir individuals for detection in our tests. However, given the very short time period between the end of the Aurignacian (~35k) and the early Sunghir 2 and 3 specimens (~34.6ka and ~34.5ka) there does not appear to be sufficient time or generations for selection to act across all the loci as required for the latter model, and an abrupt

population replacement seems much more consistent with the results, and matches existing genomic analyses[22,23,102].

Interestingly, the Gravettian hard sweep patterns suggest that a younger specimen recovered from excavations in the 1970's at Bacho Kiro (BK-1653, 35.3-34.6ka) is also Gravettian (Extended Data Table 3). This specimen was found in layers with sparse artefacts, which were attributed to be Aurignacian[21,43]. However, the hard sweep pattern in BK-1653 shows the characteristic combination of IUP and Aurignacian hard sweeps seen in the other Gravettian samples. The age of the specimen is exactly at the transition from the Aurignacian to the Gravettian, although like other Upper Paleolithic replacements the process is spatiotemporally complex.

It is also notable that the end of the IUP and onset of the Aurignacian (or Proto-Aurignacian at some sites), occurs around the time of major environmental changes that have been associated with the Laschamps Geomagnetic Event around 42-41ka[24]. Surprisingly, the end of the Aurignacian and onset of the Gravettian (~35ka) is similarly correlated with a major geomagnetic excursion - Mono Lake at 35ka[24]. Lastly, genetic and some archaeological estimates place the end of the Gravettian close to another geomagnetic excursion, Rockall at ~26ka[104]. While the end of the Gravettian in particular is not well defined, the overall pattern of correlations raises the possibility that climatic shifts or other changes associated with major geomagnetic events have negatively impacted Upper Paleolithic European human cultures[24].

The last Upper Paleolithic sweep we detect is *FBXO15*, which appears in the Magdalenian culture El Miron specimen (19ka), towards the end of the Last Glacial Maximum (LGM). The El Miron genetic group[20] has been shown to be genetically related to the Goyet-116 lineage and indeed they share 13/34, or 38%, of their combined sweeps. The Magdalenian culture is a western European culture at the end of the LGM, which was superseded by the Azilian culture in western Europe during the Epipaleolithic period (represented by the 13.7ka Bichon specimen from Switzerland).

After the LGM (ie <18ka), three sweeps (*CCDC138*, *FBN1*, *PPARD*) appear in European populations of eastern hunter gatherers, of which *PPARD* was also detected in the older 24.3ka Mal'ta specimen (Irkutsk, eastern Russia). As a result of this observation, and known geneflow from eastern areas at this time, we have designated these sweeps as likely originating in eastern Eurasia, and largely outside of our sampling area. This makes it

difficult to ascertain the actual age of first appearance, although the oldest observation of any sweep in this group is 24.3ka with the Mal'ta specimen.

The last group of three sweeps (*GABBR1*, *MIR662*, *SLC7A1*) all appear in a 15.4ka Anatolian Hunter Gatherer specimen[105] known to have ~25% genomic ancestry from the Basal Eurasians (who diverged from the Main Eurasians prior to the main phase of Neandertal admixture). As a result, we designate these sweeps as potentially Basal Eurasian in origin, having presumably originated at some point on that lineage after the divergence from the Main Eurasians. Again it is difficult to determine the origin date for these sweeps.

2.7 Potential implications of the reconstructed model of Eurasian dispersals 51ka onwards

It is worth noting that the use of Andaman and Oceanian (Aboriginal, Papuan, Vanuatu) genomes as the closest modern proxy to the AMH population dispersing across Eurasia at ~51-53ka also illustrates that the shared physiological features distributed across these modern groups may also comprise the best approximation of the initial western European populations of the Early Upper Paleolithic, potentially as far as the Aurignacian cave artists who recorded the detailed figurative art from around 42ka[24,106]. The latter possibility is consistent with the rapid and widespread appearance of similar cave art from 45-40ka across the breadth of the area covered in the initial Eurasian dispersal - from southern Europe, to ISEA and Australia, where common artistic themes involve red ochre handprints (negatives and positives), series of red ochre dots, as well as figurative artistic depictions (eg animals). As such, the fact that the Andaman and Oceanian populations have remained isolated since the initial dispersal means that the characteristics shared across these groups may provide a useful starting point to conceptualise traits (such as genetic, physiological, and behavioural) of those EUP western Eurasian groups prior to the selection imposed by the cold western Eurasian environment.

Recent studies of early Sri Lankan sites dated as early as ~48ka[107] have suggested the use of bows and arrows, nets, and advanced symbolism such as ochre bead necklaces. It seems likely that the AMH groups from this site were descendants of the original dispersal through southeast Asia including the ancestors of the Oceanians, who clearly used advanced deep sea marine voyaging capabilities to cross to Sahul. As a result, the implied combined set of technological and symbolic skills held by the dispersing Eurasian groups at 51-50ka appears to be impressive, and perhaps much further advanced than is often conceived. As such, a re-calibration of these concepts is probably justified.

2.8 Accumulation of sweeps throughout the Upper Paleolithic

Our analyses suggest that many of the sweeps are driven by selection pressures (*e.g.* cold temperatures, see section 3) that persisted across the majority of the Upper Paleolithic in Eurasia. Accordingly, the persistence of the selection pressure may explain why sweeps tend to aggregate through time across the Upper Paleolithic before undergoing a notable decrease in the Holocene (Fig. 2) – potentially as a consequence of extensive admixture throughout this period with lineages that did not carry the sweep haplotypes, such as the Middle East particularly during the dispersal of farmers, and especially the subsequent Empire era following the Bronze Age.

To investigate this pattern further, we used the *lm* function in R v3.6.3[108] to regress the total number of sweeps as a linear function of sample date in Western Hunter Gatherers (WHG) and moderate- to high-coverage specimens from the early period of the Upper Paleolithic back to the Neandertal admixture/Main Eurasian population. To account for missing sweep information in some samples (sweep haplotypes were not estimated for samples that had <50% of the total sweep marker SNPs), we used the proportion of total sweeps as the explanatory variable, excluding sweep *LINC01153* that could not be reliably dated due to the lack of marker SNPs. We also included several Oceanic samples to model a plausible range of sweep counts occurring around the separation time of these lineages from Main Eurasian branch ~51kya. Notably, the linear function provided an excellent fit to the data, with an adjusted-R^2 of 0.7048. Under the assumption that sweeps have aggregated in an approximately linear fashion since humans first left Africa, we explored the age when this linear model suggests there would have been 0 sweeps (i.e. the predicted x-intercept in the regression model), as this provides a rough estimate for the initial isolation of the OOA population. To take into account the known uncertainties in the ages of each sample, we randomly sampled the ages from Normal distributions, using the given mean and standard deviations reported for each sample (Extended Data Table 1). For each set of sampled ages, we used the bootstrap *boot* package v1.3-24 in R[109] to resample the residuals and refit a linear model, to determine 1,000 samples for the x-intercept. We aggregated the x-intercept samples from each of the 1,000 sets of sampled dates and then determined the mean x-intercept to be 83,141 yBP and the associated empirical 95% confidence interval to be between 72,369 and 97,284 yBP.

Since the linear regression model assumes that hard sweeps will accumulate linearly through time – an assumption that may be unrealistic given the complex demographic and evolutionary history of Western Eurasians – we performed a more extensive analysis that explicitly models key aspects of Eurasian demography and evolutionary history, drawing the relevant parameters from recent population genetic studies of Eurasian history. Specifically, we simulated a demographic scenario where novel beneficial mutations arise at random in a single panmictic population over the last 200,000 years, with the mutation frequency depending on the current effective population size (i.e. beneficial mutations are more[less] likely to occur as N_e gets larger[smaller]). The population size dynamics are based on N_e estimates leading up to Early European Farmers (EEF) from ref [110], which included ancient Eurasian samples in their inference procedure. For simplicity, we ignore other historical events in our model, such as the separation of Main Eurasian and Basal Eurasian lineages and also the Neandertal admixture event. Using the coalescent simulator msms[111], we simulated 5,000 selected allele frequency trajectories under this evolutionary model. For each selected locus, we also simulated linked neutral genetic variation data across a 5Mb region for the hypothetical EEF population from ~8kya (i.e. LBK branch in ref. [110]) and ran these sequences through SweepFinder2 to filter out selected loci that did not leave a significant sweep signal in this population. To ensure that the simulated data match the statistical properties of the ancient data used in the present study, we replicated the empirical SNP ascertainment scheme by conditioning on SNP presence in a simulated African population and also generated the same number of simulated 5Mb sequences as observed for the Neolithic Anatolian population in the present study. Similarly, selection coefficients were sampled from an exponential distribution with a mean selection coefficient of s = 1%, which leads to selection coefficients that broadly match those estimated from our sweeps (s ranging between ~1 to ~10%; see ref. [7]). More details of our simulation approach, including command line parameters and estimation of selection coefficients, are available in ref. [7].

To infer the earliest time for the onset of selection that best fits the data under our evolutionary model, we removed all simulated loci with a significant SFS signal that started prior to time point x, and then averaged the frequency trajectories of remaining loci to estimate the aggregate probability of observing a sweep haplotype across time. This is equivalent to estimating the proportion of all sweeps that are observable at each time point, conditional on the sweep arising after time point x and being detected by SweepFinder2 in the EEF population 8kya. Additionally, to incorporate the statistical properties introduced by

missing data and sample pseudohaploidization (see section 2.3) in our estimation, we modeled the probability of detecting a particular sweep haplotype in an individual as $p_x = (1-f_x)^2 \times 0.061 + 2f_x(1-f_x) \times 0.256 + f_x^2$, where f_x is the frequency of the sweep at time point x. In other words, we are estimating the expected weighted frequency of the beneficial haplotype at a given time, where the genotype frequencies (assuming Hardy-Weinberg equilibrium) are weighted by the estimated probability of detecting a sweep haplotype given that genotype (see section 2.3). The subsequent estimates of p_x are then averaged over all selected loci to obtain the probability for detecting a sweep haplotype at time x.

To compare the simulated results to our empirical estimates, we used maximum likelihood (ML) to measure the fit of key ancient and modern samples to the predicted p_x value, varying the earliest onset of selection over successive 1,000 year intervals between 25ka to 150ka. Because our simulations were based on a single panmictic population, we only estimated the ML function for ancient Eurasian specimens >30ky old, as more recent specimens are likely to be affected by the substantial structuring observed in West Eurasian populations that arose during the Last Glacial Maximum ~30-15ka[20], and consequently are only expected to contain a subset of all sweeps that occurred along the ancestral lineage that leads to the EEF population. Using the same samples that were used for the linear regression analysis reported above (other than samples dated within the past 30ka), we arrived at a likely selection start between 74ka-91ka (95% CI from ML estimation, point estimate = 79ka). Thus, the linear regression and model-based approach arrive at very similar point estimates, ca. 80ka, and concordant confidence intervals for the earliest starting time of selection for our hard sweeps.

Taken together, the two approaches suggest that after an early date for the initial isolation of the OOA migrants (*e.g.* between 100-120ka based on genomic data), the Main Eurasians started to accumulate hard sweeps from around 80ka, potentially as a result of becoming genetically isolated from other populations around that time. A resulting prolonged period of genetic isolation (e.g. 80-53ka) is far more compatible with the large number of hard sweeps observed in the earliest Oceanic and Eurasian groups than the common 50-60ka estimate for the OoA movement simply based on the mean time of Neandertal geneflow into Main Eurasians. An estimated 80ka origin date also closely coincides with paleoenvironmental evidence of a moist phase ~80ka within the Arabian Peninsula (Fig. 5), when AMH populations appear to have dispersed more widely through the interior, before potentially being isolated by the subsequent return of arid conditions shortly thereafter. This population fragmentation may also have caused the separation of the Main and Basal Eurasian

populations within the Arabian Peninsula, which genetic estimates indicate occurred in this area around this time. The above events are shown in Fig. 1, as part of the Arabian Standstill model, while the relationship to the moist phases, and timing of major global cooling events following the onset of MIS4 (~71ka) and Toba eruption (~74ka) are indicated in Fig. 5.

2.9 'Out of Arabia': Climate changes in the Arabian Peninsula and wider region

In contrast to late Pleistocene climate changes recorded in Greenland[49], hydroclimate and temperatures over northeast Africa, the Levant and Arabia did not mirror the stadial-interstadial cycles of the North Atlantic (the glacial period). Instead, distinctly different regional changes appear to have been driven by orbital variability (most notably precession[35]), Eurasian snow and ice coverage[112], lowered atmospheric greenhouse gas concentrations and/or changes in monsoonal intensity and seasonality[113]. To disentangle late Pleistocene hydroclimate and temperature trends, several key reconstructions have been generated from continuous marine records straddling the Gulf of Aden, the upwelling zone of the Arabian Sea and extending towards the southwest Indian Ocean[35,114]. Leaf waxes in the Gulf of Aden core RC09-166 are of particular importance. Aerially derived from the Horn of Africa and Afar regions, leaf wax (ice-volume corrected) δD provides a measure of the isotopic composition of precipitation used by higher plants to create their lipids, a proxy for aridity [35]. These show that during Marine Isotope Stage 6 (MIS 6), while there were large swings in aridity, the region experienced relatively high sustained mean annual temperatures [35,114]. Importantly, these relatively warm temperatures persisted through the Last Interglacial (Marine Isotope Stage 5) with moist conditions during MIS 5e, 5c and 5a, consistent with relatively high precipitation and lake levels[115,116], creating 'green corridors' across the region [35]. These brief moist spells have been related to signs of AMH presence throughout the Arabian Peninsula, especially around 80 ka[19]. At the termination of MIS 5 (from 80 ka), mean annual temperatures dramatically decreased by ~4˚C, accompanied by increased aridity into the last glacial period. Potentially importantly, AMH populations effectively trapped in the Arabian Peninsula area would have likely been exposed to considerably colder temperatures than these reconstructions suggest. The quantified estimates from the region are derived for alkenones and represent mean annual temperatures[117], implying the boreal winter was considerably more frigid (*i.e.* this is a conservative estimate of the temperature decline). Under the arid conditions of MIS 4 from 80ka onwards, apart from coastal sites and the Arabian Gulf, the biodiversity-rich mountain ranges >2,000m asl across Arabia (e.g. Shada and Hajar Mountains)[118] would have provided potential food and water resources, albeit at the

cost of exposure to still colder temperatures. Population movements into these interior sites during the brief moist phase around 80ka may have resulted in population fragmentation, potentially such as that separating the Main and Basal Eurasians depicted in Fig. 1.

Importantly, the eruption of the Toba volcano in northern Sumatra around 74ka, may have amplified the impacts of these climate changes. Tephra-fall deposits from Toba, the largest explosive eruption in the Quaternary, have been identified across Indonesia, southern continental Asia, the Indian Ocean and into the Arabian Sea[119–121]. The associated injection of large quantities of sulfate into the troposphere and stratosphere has been estimated to be up to 100 times that of the 1991 Mt Pinatubo eruption, and as a result, likely had a substantial impact on incoming solar radiation[122]. With changes in surface albedo, Toba has been used to explain shifts in climate and vegetation patterns around the world at this time, including aridification and replacement of woodland by open grassland in India[123], cooling in the south Atlantic[124], a weakened Asian Monsoon[125], and even the onset of the last glacial period [36]. Analysis of the Greenland ice core records suggest amplified cooling during the first two centuries following the eruption[126] with some modelling studies suggesting global cooling on the order of 11˚C, with Africa potentially decreasing by as much as 17˚C[122]. Whilst centennial to millennial-duration environmental impacts remain uncertain[127], some modelling studies suggest the loss of vegetation may have amplified the amount of cooling (8-17˚C) on the order of decades, sufficiently long in duration and magnitude to have substantial impacts on ecosystems and human populations[128].

From 57-50ka (Marine Isotope Stage 3), reconstructions suggest a relatively ‘brief’ moist period across the region, accompanied by sustained cool conditions, providing a potential green corridor for migration northwards from the Arabian Peninsula[35], consistent with the timing of the subsequent main phase of admixture with Neandertals ~51-53 ka.

3. Evolutionary and functional properties of selected genes

3.1 Refining the targets of selection

Because 52 of the 57 detected sweeps contained two or more genes, we attempted to refine the putative target of selection to a single gene in each multigene sweep using recently developed iSAFE software[27]. The iSAFE metric quantifies the evidence for selection at each SNP within a predefined window that has been applied to well established cases of positive selection in recent human history to identify putative causal mutations. Using simulated 5Mb

regions, the iSAFE developers showed that the selected SNP was amongst the top 20 ranked SNPs in 94% of cases[27]. Accordingly, amongst all genes in a sweep window, those with the highest proportions of the top 20 ranking iSAFE alleles are the most likely target selection. Extending this logic, the proportion of top 20 ranked iSAFE SNPs carried by a gene provides a coarse estimate of the probability that the gene actually contains the causal allele somewhere within the annotated gene boundaries. Thus, for all 50 multigene sweeps, for searched for instances were a single gene has a majority (≥50%) of the top iSAFE SNPs and where no other gene in the same sweep has ≥20% (at least 4) of the top 20 iSAFE SNPs.

We applied iSAFE to the 52 multigene sweeps regions in phased genomes from five modern European populations and five modern African populations from phase 3 of the 1,000 genomes dataset, which included samples from populations used our SweepFinder2 assays (i.e. CEU, FIN, TSI, and YRI). African samples were included in the iSAFE estimation in order to provide a sample of non-sweep haplotypes, as this has been shown to increase power to identify the causal allele when the sweep haplotype is at high frequencies[27]. We also extended the boundaries of each sweep by 250kb on each side, as the inclusion of sweep shoulders and flanking regions also improves the power of iSAFE[27]. Application of the iSAFE method to the 52 multigene sweeps resulted in 25 sweeps where a single candidate driver gene could be identified that contained at least 50% of the top 20 iSAFE alleles (Extended Data Table 4). Amongst these 25 sweeps, three sweep regions had at least one other gene that carried a moderate (≥20%) proportion of iSAFE SNPs (i.e. sweeps *CCDC138*, *ARFGEF1*, and *COL4A3BP*; Extended Data Table 4). Because such sweeps provided less convincing evidence for a single focal driver gene overall, these were flagged as edge cases. Three additional sweeps (*PNLIPRP3*, *TP53BP1*, and *FMO2*) were also included as edge cases on the basis that these regions had a single gene with ≥30% of the top 20 iSAFE SNPs, but no other gene in the sweep contained any of these top ranked iSAFE SNPs（Extended Data Table 4). Accordingly, subsequent tests based on putative driver genes were performed using the full repertoire of driver genes (32 in total, 28 iSAFE derived sweeps plus four sweeps containing single genes; *LINC01153* being omitted due to having too few marker SNPs for other analyses, and also having no recorded function in medical literature; Extended Data Table 5).

3.2 Comparative data sets

Population migrations often place organisms in unique conditions that pose new physiological challenges, like those experienced by contemporary human populations that have recently moved into high altitude situations who suffer increased perinatal mortality and morbidity compared to high altitude-adapted populations[129–131]. Similarly, the OoA migration by AMH introduced is thought to have given rise to a variety of adaptations to the cool and dry environments of Pleistocene Arabia and Eurasia[132]. Previous investigations of mammalian adaptation to cold environments have reported a number of putatively adaptive changes to regulate temperature, including neurocranial restructuring (face structure in cold living nonhuman primates[133]), as well as changes to the volume and distribution of fat tissue together with modifications in energy expenditure[134–136], and selection for skin traits[137]. Cold temperatures are also known to affect respiratory health and sensitivity to infection[138,139].

To determine if our driver loci were potentially part of an adaptive response to living in cold climates during the AMH occupation of Arabian Peninsula and Eurasia, we compared the functional characteristics of the driver loci with candidate loci identified from four previous studies of positive selection in Arctic human populations living in Greenland[28], Siberia[140], and North America[141,142]. We obtained the previously curated set of candidate genes from the Greenland and Siberian populations from a recent meta-analysis of mammalian cold adaptation[30], adding three candidate genes from ref. [142] and another 15 candidate genes from ref. [141]. In the latter case, the 15 candidate genes were obtained from a larger set of the ~80 reported outlier genes, by first removing all genes that did not contain any non-synonymous, splicing, or UTR-based variants, and then removing any remaining genes that were within 1Mb of another outlier gene. Compiling the candidate genes from all four studies resulted in a set of 76 candidate adaptive genes from modern Arctic human groups that were considered for functional analyses (Extended Data Table 5).

Additionally, previous studies have demonstrated that a large number of Neandertal and Denisovan loci are segregating in modern human populations as a result of previous admixture events[42,143–146]. While many of these archaic loci are thought to have survived simply due to being in selectively neutral areas of the genome, a small number exhibit signs of positive selection suggesting they have conferred an adaptive advantage in some human populations – *e.g.* the Denisovan variant of the *EPAS1* gene which confers high-altitude tolerance in modern Tibetans[147]. Given the potential advantage of cold-adapted Neandertal

and Denisovan genes to the AMH population spreading across Eurasia, we also compiled a set of candidate AI genes from three studies that performed explicit tests for AI loci[41,148,149]. To ensure that the likely driver gene was identified, we restricted our analyses to AI regions comprising a single gene that did not lie within 1Mb of a neighbouring adaptively-introgressed region from the same study. Further, due to the large number of reported AI loci in ref. [41], we restricted AI regions from this study to those with an FDR < 0.2. This resulted in a set of 56 AI loci, which were further supplemented with a five additional loci that had been specifically singled out for further functional examination in the four AI studies and another recent study investigating the distribution of introgressed hominin DNA in humans[146], producing in a final set of 61 candidate AI loci that were considered in further functional analyses (Extended Data Table 5).

3.3 Sweeps are enriched with genes intolerant of deletion in present day populations

The availability of large genomic datasets across multiple human populations has allowed for rigorous assessment of the 'essentiality' of all annotated human genes[44]. In contrast to genes with high levels of redundancy, such as olfactory genes, that exhibit large numbers of segregating putative loss of function alleles (pLoF) in the human populations[44,150], genes with essential functions are expected to be less tolerant of such inactivating mutations and subsequently should carry fewer pLoF alleles on average[44]. Following this logic, a recent study developed a continuous metric – LOEUF (for loss-of-function observed/expected upper bound fraction) – that quantifies tolerance to pLoF variation across human genes based on segregating pLoF variants estimated from >140,000 exomes from diverse human populations[44]. Notably, genes with low LOEUF scores (i.e. more essential genes) are more likely to be haploinsufficient (i.e. heterozygotes do not produce the wild-type phenotype) and tend to be lethal when heterozygous in mice and human cell lines[44]. Additionally, segregating variants are more likely to be deleterious in genes with low LOEUF scores in general, regardless of their population frequency[44], and LOEUF scores are correlated with two codon-based measures of gene conservation in human and vertebrate lineages (called synRVIS and synGREP, respectively[151]). Thus, genes with low LOEUF scores tend to be intolerant of new variants in general, and preservation of their core functionality appears to be a crucial determinant in maintaining human health.

We used the LOEUF scores for ~20,000 annotated human genes to test if any of the candidate gene sets (i.e. AI, Arctic modern human, and driver genes from the present study) were more

essential than expected by chance, by contrasting the LOEUF scores in each candidate set against all non-candidate genes using a Wilcoxon Sign Rank Test. The LOEUF scores were significantly lower for all three candidate gene sets (driver genes: $p \leq 0.0009$; AI genes: $p \leq 0.0001$; Arctic modern human genes: $p \leq 0.0028$; Extended Data Fig. 7), and this result remained significant after removing the six edge cases from the driver genes ($p \leq 0.0009$), indicating that these candidate gene sets tend to be less permissive to pLoF mutations in human lineages than expected by chance.

Notably, it is possible that the observed tendency toward low LOEUF scores for three putative selected gene sets was a result of positive selection removing pLoF variation at these loci. To test this scenario, we compared another pLoF-based metric that was correlated with LOEUF that was also calculated in ref [44] – namely *p,* the probability that a haplotype carries a pLoF mutation at a particular gene – for genomes from African and non-African populations. Specifically, we computed a modified value of *p* that captures changes that followed OoA movement out of Africa, by performing a LOESS regression of p_{EUR} (i.e. *p* derived from individuals with non-Finnish European ancestry) upon p_{AFR} (i.e. *p* calculated for individuals with African or African American ancestry) using the loess() function provided by the R statistical programming language with default parameter settings[108]. All *p* values were logarithmically transformed (base 10), which required adding a small random number uniformly distributed between 10^{-5} and 10^{-6} to account for *p* values equal to 0. After computing the LOESS model on these log-transformed *p* values, we obtained corrected values, p_{corr}, for each gene by subtracting the observed p_{EUR} from the predicted value in the LOESS model; i.e. the corrected value is the residual score for each gene, whereby positive p_{corr} values indicate that a gene has fewer LoF mutations than expected following the movement of AMH into Eurasia (and vice versa for negative values). We then used a one-sided Wilcoxon Signed Rank Test to determine if the mean p_{corr} was significantly inflated across any of the three categories of candidate genes (i.e. AI, Arctic modern human, and driver genes) relative to the remaining 'null' set of genes. The observed mean p_{corr} values did not differ from expectations for any of the three candidate gene sets (AI genes: $p \leq 0.40$; Arctic modern human genes: $p \leq 0.96$; driver genes: $p \leq 0.16$; Extended Data Fig. 7). Indeed, the deleterious nature of LoF variants imply that most will have evolutionarily recent origins due to new LoF mutants being quickly removed by purifying selection. Our results are consistent with this expectation and indicate that the LOEUF scores for our driver genes are probably only weakly impacted by the putative positive selection events that largely occurred

>10,000 years ago, and therefore should still provide a robust measure of the essentiality of genes putatively affected by positive selection following the OoA movements.

3.4 Functional characterisation of driver genes

The biological function of the 32 putative Eurasian driver genes identified within our hard sweeps was assessed by a systematic review of human disease literature, as well as animal and cell knockout phenotypic data sets, available on OMIM, PubMed and GeneCards databases. For 21 driver genes we could define a biological function based upon its phenotypic information in human carriers of crippling or haploinsufficient alleles or from animal studies investigating the impact of deletions or loss-of-function mutations. The majority of these animal studies were rodent 'knock-out' assays, though zebra fish and *Drosophila* studies were also included in some cases (Extended Data Table 5). For the 11 genes lacking this type of phenotypic information, functions were assigned to four via analysis of human and or animal genome-wide association trait studies (GWAS) and biochemical and cell-based studies were available (Extended Data Table 5). Functions were assigned for the remaining 7 genes by analysis of biochemical and cell-based studies alone. For one driver gene, *LINC01153*, function could not be assigned due to lack of currently available data, and it was excluded from this analysis for this and because there were too few marker SNPs to make accurate measurements[26] (see Extended Data Table 5 and section 3.1).

Using this approach, 32 Eurasian driver genes could be assigned a clear functional role under a surprising small number of higher level categories: Neurological (31% of total genes); Development (31%); Metabolism (28%); Reproductive Fitness (6%); and Angiogenesis (3%); while no genes were associated with immune function (Extended Data Table 5). Further, just over half of all the selected genes (59%) were associated with a major physiological impact in humans and animals when mutated, with phenotypes including premature lethality (e.g. *GTSE1*, *FBN1*), spontaneous neuro-developmental defects (e.g. *ARFGEF1*, *NFASC*), premature ageing, skeletal and organ malformations and rearrangements (e.g. *TBC1D7*, *DNAH7*) and other whole body defects (e.g. absence of a vascular network, *BCAS3;* reduced sperm output and fertility defects *ELMO1*) (Extended Data Table 5). Further evidence for the biological importance of the driver genes is provided by their significantly reduced LOEUF scores (see section 3.2) and the observation that 25% are associated with a lethal phenotype when functionally crippled (Tables 1, Extended Data Table 5). All of these

parameters point to the selected genes having relatively high biological significance and performing non-redundant functions in human physiology.

A second striking observation was that the large majority of the driver genes (~88%) encode intracellular proteins (Extended Data Table 5), and many of these are involved in related evolutionary conserved processes or gene regulatory networks. For example, *CAND1* and *FBXO15* are both components of the Skp1-cullin-F-box (*SCF*) complex[152] – a cellular E3 ligase machine conserved across yeast, plants, nematodes, drosophila and humans – which is critical to regulation of cell cycle progression[153]. Other highly conserved genes include *TP53BP1*[154,155], which was first identified in yeast, which together with *GTSE1*[156] and *FANCD2*[157] regulate components of the evolutionarily-ancient DNA damage response and repair pathways. Further, the driver gene *TAF15* is an evolutionary conserved transcription factor (related to the Drosophila homolog *cabeza/SARFH* associated factor[158]) that cooperates with the RNA polymerase II complex to regulate gene transcription, including regulating the expression of the selected gene *DNAH7*[159] (Fig. 4D). This suggests that recent human evolution has included repeated adaptive modifications of proteins involved in core cellular functions. Notably, cell cycle progression and DNA repair are intertwined processes, implying that our driver genes not only fall within a relatively small set of high-level functional classifications, but also share mechanistic associations at the subcellular functional level.

3.5 Driver genes form coordinated functional pathways with adaptively introgressed hominin genes and positively selected Arctic human genes

Adopting the same criteria and analyses that were used to functionally characterise the driver genes identified in this study, we were able to assign a function to 54 of the 61 AI loci and 49 of the 76 Arctic modern human loci resulting in a set of 103 loci that were subjected to further functional investigation (Extended Data Table 5). Intriguingly, the high-order functional classifications identified for the AI and modern Arctic candidate genes closely align with the distribution of functions observed for the driver genes – with the majority (≥80%) of all AI and Arctic modern human genes also grouping within neurological, development, or metabolic functional domains (Extended Data Table 5). Moreover, many of the candidate genes within either group were associated with severe physiological phenotypes (i.e. lethality and or Mendelian disease) when disrupted (60% and 63% of all AI and Arctic genes, respectively; Extended Data Table 5). Even more strikingly, several of the Eurasian

driver genes also formed parts of larger integrated molecular networks that included both AI and Arctic human candidate genes. Comparing the functions across the three datasets revealed distinct interacting gene networks – fat metabolism, cilia formation and skin physiology – each producing a phenotype with a potentially adaptive advantage in cold conditions, which are discussed in the main text and accompanying figure (Fig. 4, Extended Data Table 5) and outlined in more detail below.

There was an unexpected concentration of genes involved in the developmental formation of cilia and cilia motility in the ancient Eurasian driver genes and the Arctic and AI candidate gene sets (Fig. 4). Interestingly, cilia functional genes were enriched within our driver loci and genes under selection in multiple Arctic mammals (see section 3.7). Cilia are evolutionarily-conserved hair-like cell structures that can function as cellular environmental sensors or provide locomotion, but are also important for lung health in cold and dry environments[160]. The Eurasian driver genes *DNAH6*, *DNAH7,* and *CCDC138* along with *DNAH2*, *DNAH3,* and *DNAH17* from the Greenland Inuit study, all encode molecular structural components of cilia while the *SDCCAG8* gene identified in the Siberian study controls cilia development[26] (Extended Data Table 5). Similarly, while the archaic hominin AI gene *WDR88* has no known function, it's close paralogue *DAW1* is involved in building the cilia structures containing *DNAH6* & *DNAH7*[161]. Further, Eurasion driver genes *DNAH6*, *DNAH7*, and Arctic selected genes *DNAH2*, *DNAH3* and *DNAH17* all comprise specific components of the axonemal dynein motor complex collectively providing the force generation for cilia movement[162].

Another concentration of selected genes across the three groups relates to skin physiology and pigmentation (Fig. 4). Skin pigmentation is dictated by melanosomes where melanin is produced and then transported to intracellular and extracellular sites[163]. In this process the ancient Eurasian driver gene *MLPH* interacts biochemically with the archaic hominin AI gene *RAB27A,* to shuttle melanin containing melanosomes to the cellular periphery[164]. Further, melanin synthesis, which occurs in melanosomes, requires the "P protein" encoded by the hominin AI and Arctic selected gene OCA2 [165]. Finally, melanosome formation (melanogenesis) requires the Arctic cold-adapted gene *SLC24A5*[166]. Mutations in either *MLPH* or *RAB27A* cause the hypopigmentation condition Griscelli syndrome[167,168], while *SLC24A5* and *OCA2* are associated with albinism in humans[165,169]. It is interesting to note these 'colour' genes appear to be functionally conserved throughout vertebrates, as polymorphisms in *MLPH* contribute to coat colour in domestic cats[170] and mutations in

SLC24A5 and *OCA2* result in the 'golden' phenotype in zebrafish and cichlids respectively[171,172]. Selection for pigmentation in human populations during the OoA migration has been noted previously and suggested to be a response to UV/sunlight levels as a selective force. It is noteworthy that a number of the Eurasian sweep genes are involved in the DNA damage response (e.g. *TP53BP1*, *GTSE1, FANCD2)* which may also relate to selection due to high UV/sunlight levels. In this regard, *MLPH, TP53BP1*, and *GTSE1* all occur within the Arabian Standstill period, while *FANCD2* is first detected in the Aurignacian and undated eastern populations.

In addition to selection for skin pigmentation genes, genes regulating physiological properties and homeostasis of skin also showed evidence for selection. The gene *EDAR* is an Arctic modern human candidate adaptive gene that is also part of a large sweep in the present study that results in ectodermal dysplasia (characterised by primary defects in hair, nails, teeth as well as sweat glands and skin) when mutated[173]. Other genes regulating skin and ectodermal properties were also part of the Arctic modern human candidate genes including *SREBF2*, which associates with impaired skin wound repair and cataracts[174], and the *DSP* gene which regulates skin thickness and the 'woolly' phenotype[26].

3.6 Neurological function as a underappreciated adaptive target during human selection

Around one third of the 32 Eurasian driver genes that could be assigned a clear physiological role were associated with neurological function. The majority of those neuronal genes (8/10) underwent strong selection during the Arabian Standstill period of ~80-50ka. The remaining two neuronal sweeps (*WWOX* and *DOCK3)* arise immediately afterwards and are observed in the initial Upper Paleolithic (~45-40ka). Neuronal adaptations therefore appear to have been essential components of the selective environment during the Arabian Standstill and during the early phase of AMH dispersal across Eurasia.

Selection for neurological function stands in apparent contrast to previous evidence of positive selection in humans where the link between the selective force and adaptive response are more readily obvious e.g. changes in diet resulted in selection for genes controlling lactose tolerance; infectious disease resulted in selection for genes controlling malaria resistance; changes in climate resulted in selection for genes controlling pigmentation. While neurological functions initially appeared surprising, we propose that this observation most likely relates to the critical role the nervous system and brain play in coordinating, integrating, and subsequently regulating diverse physiological processes, which are impacted

by cold environments. A key influence of the brain in this context is its ability to synthesize immediate environmental contextual cues and integrate these with the present physiological state. This is exemplified by evidence that the brain plays a central role in adjusting fuel metabolism and appetite to food resource availability[175], regulating blood flow in response to changed external conditions (e.g. cold) for thermal homeostasis [176], directly aids immune defense against infection and injury[177], and ensures the fetus is protected, safely delivered, and cared for to facilitate successful pregnancy[37]. Indeed, both cognitive performance[34,178–180] and multiple aspects of neural physiology[181–184] are sensitive to thermal variation, suggesting that sustained occupation of cold environments by AMH may have required recalibration of many essential neurological processes towards new optimal settings, potentially leading to some of the adaptive changes observed in this study.

Thus, the Eurasian sweeps point to the brain as a central mechanism facilitating human adaptation to changing environmental conditions where cold temperatures are a stressor, a hypothesis that is further supported by our finding that neurological genes also represent ~33% of candidate genes observed in modern Arctic human groups. It is intriguing to speculate that neurological adaptation may have helped to facilitate the spread of AMH through the colder Eurasian environments ~50kya.

3.7 Testing functional enrichment in annotated gene sets

For the 32 sweeps for which a putative driver gene could be determined, we used GOrilla[185,186] to test if any gene ontology (GO) terms are enriched amongst our driver genes compared to the remaining background genes. Eight GO terms were found with $p < 0.001$, which included two cilia-based GO processes – *cilium or flagellum-dependent cell motility* and *cilium-dependent cell motility*; (Extended Data Table 8). Though neither of these processes were significant after multiple testing, the *cilium or flagellum-dependent cell motility* was found to be significantly enriched amongst selected genes observed in Greenland Inuit ($q \leq 0.0104$; see Extended Data Table 6 in ref. [28]). Notably the genes contributing to the enrichment observed for Greenland Inuit (*DNAH2*, *DNAH3*, *DNAH17*) differed from those contributing to the enrichment in our study (*DNAH6*, *DNAH7*), which is likely partly due to the different candidate detection criteria used between the two studies (see section 3.5).

Enriched cilia GO terms were also reported for a collated set of 416 putatively selected genes from Arctic mammals (*regulation of cilium beat frequency*, *regulation of cilium movement*, *non-motile cilium assembly*, and *ciliary basal body docking*; all $q \leq 0.4$; see Extended Data Table 8 in ref. [30]), which combined outlier loci from Yakutian horses[187], polar bears[188], minke whales[189], Arctic foxes[190], extinct mammoths[191], and two Arctic human populations[28,140]. This collated gene set was also enriched for genes that are highly expressed in the lungs, and a further GO analysis on the subset of lung-expressed genes also resulted in enrichment of cilia based GO processes (see tables S9 and S10 in ref. [30]).

3.8 Comparison with selected genes reported in Arctic human populations

No candidate genes were shared between ancient Eurasian driver genes and the combined candidate gene set from the four Arctic human studies, suggesting that while selection has targeted common biological functions in both ancient Eurasian and modern Arctic human groups, different sets of loci were the focus. However, this lack of overlap is perhaps not unexpected given the differences between the candidate locus detection criteria employed by the Arctic human studies and the current study. All Arctic human studies used criteria that favoured the detection of regions under 'local' selection in the Arctic populations – i.e. genomic regions that showed large frequency changes in these Arctic populations relative to other Eurasian groups. Accordingly, while we expect that a subset of the 57 ancient Eurasian sweeps from the current study would exist in Arctic populations (see the next part of this section), the fact that these loci are also expected to be at high frequencies in other Eurasian groups means that they are unlikely to be amongst the extreme outliers selected, (i.e. candidate) loci in the Arctic populations. Specifically, two of the Arctic studies (refs. [28,141]) used the population branch statistic (PBS[192]) to test for genomic regions showing excessive differentiation relative to modern Han Chinese (CHB) populations, with Northern Europeans (CEU) as an outgroup. The other two studies used PBS tests and another statistic that scans for regions with depleted haplotype diversity in the focal populations (iHS[193]), with ref. [140] also using a further divergence statistic XP-EHH[194] that contrasts haplotype diversity in the focal population with a Eurasian group (in this case an ethnic Vietnamese population). Notably, while iHS statistic should in principle be capable of detecting sweeps that arose before the separation of each Arctic populations from other Mainland Eurasian groups, both studies use post-test filtering strategies that greatly reduce the chances of observing such sweeps – ref. [140] ignores the top 5% of iHS windows from the ethnic Vietnamese population,

and ref. [142] only reports outliers in the top 1% of PBS and iHS statistics. Thus, because all of these tests use other Eurasian populations as contrasts, they will favour the detection of locally selected regions and potentially bias against the detection of the ancient sweep loci reported in the present study.

Nonetheless, it is possible that there is still some statistical correspondence between the selection scans from these studies and our own sweep regions when looking at more general signals beyond the candidate loci. Further, the extensive Holocene admixture suggested in the present study has likely reduced the sweep haplotype frequencies in many modern European populations, making it possible that some of these sweeps might have moderate PBS signals if they remained at high frequencies in the Arctic populations. To examine these scenarios in more detail, we used the permutation framework outlined in section 1.8 to test if the outlier windows reported for the Arctic populations in ref. [140] were enriched within the sweep regions detected in the present study, with separate tests performed for each of the three statistics employed in this study (i.e. PBS, XP-EHH, and iHS) as well as the combined set of results from all three statistics.

Intriguingly, while the outlier iHS windows were found to be marginally enriched across our 57 sweeps for both unscaled (observed overlap = 7 sweeps; expected overlap = 3.71 sweeps; $p \leq 0.0701$) and recombination-scaled (expected overlap = 3.96 sweeps; $p \leq 0.0905$) genomes, we also observed an significant enrichment for XP-EHH outliers in both scenarios (Unscaled: observed overlap = 11 sweeps; expected overlap = 4.51 sweeps; $p \leq 0.0028$; Scaled: expected overlap = 4.77 sweeps; $p \leq 0.0051$), but not for PBS outliers (Unscaled: observed overlap = 16 sweeps; expected overlap = 20.06 sweeps; $p \leq 0.9197$; Scaled: expected overlap = 19.36 sweeps; $p \leq 0.8779$). The PBS result indicates that there were fewer of these outlier regions overlapping with our sweeps than expected – which is consistent with our hypothesis that inclusion of a European population in this statistic has biased it against the sweeps detected in our study. However, the inclusion of the ethnic Vietnamese population as a contrast in the XP-EHH statistic did not produce a similar bias. Both Vietnamese and Siberian populations are descendents of an East Asian branch that split from Main Eurasians > 40kya[195,196], and therefore are expected to have inherited the same subset of the 57 sweeps relative to Siberians. Accordingly, the enrichment of outlier XP-EHH signals in Siberian populations suggests that the underlying selection signals have dissipated in modern Vietnamese populations, in comparison to modern European populations, possibly through

one or more admixture events that occurred after their split from Siberian lineages, or perhaps because later occupation of a subtropical climate has resulted in selection against any cold-sensitive haplotypes in the Vietnamese.

3.9 Comparison with adaptively introgressed archaic hominin loci

The growing evidence for adaptively introgressed loci in AMH populations[41,148,149,197] suggests that some of our sweeps could be driven by introgressed hominin-derived variants. To explore this possibility in more detail we compared the location of our 57 sweeps with previously identified AI hominin loci from four recent publications[41,148,149,197], and tested if there was an significant excess of AI loci overlapping our sweeps and also if any of the nine previously published hominin genomes were a close match to the marker alleles in the overlapping region.

Combining all reported outlier loci in the four AI studies resulted in 2267 distinct loci, which were refined further by combining all loci with identical chromosomal positions from the same study into a single locus, resulting in a final set of 862 distinct AI loci. We compared our 57 sweep regions with the combined set of 862 putative AI archaic hominin loci, identifying 25 distinct AI loci overlapping 14 different hard sweeps (~25% of all sweeps) (Extended Data Fig. 5). Note that because only central positions were reported in candidate loci in ref. [149], we created intervals for each locus by adding 100kb flanking regions either side of this reported position (100kb being the approximate mean size of AI loci in ref. [41]). We tested if the observed number of sweeps that overlap one or more AI hominin locus was greater than expected by chance, applying the permutation test procedure outlined in section 1.8 to the complete set of 862 AI hominin loci, with additional testing also performed for the AI loci specific to each study. These tests confirmed that our sweeps were significantly more likely to overlap the combined set of 862 AI regions than expected for both the standard genome ($p \leq 0.0092$) and the recombination rescaled genome ($p \leq 0.0113$) analyses. Moreover, excluding AI loci with Denisovan or ambiguous origins from these analyses showed that our hard sweeps were also more likely to overlap Neandertal-specific AI loci than expected ($p \leq 0.0364$ and 0.0475 for standard and recombining genome tests, respectively). The number of AI loci overlapping sweeps consistently exceeded expectations across all four studies though these were only significant ($p < 0.05$) or marginally significant ($p < 0.10$) in the case of refs. [148,149]. Notably, these two studies also reported much fewer

candidate AI loci overall, suggesting that these significant patterns may partly relate to the higher ratio of true positives to false positives in these two studies.

Next, we examined if any of the nine previously published hominin genomes available on AADR v42.4 closely matched the set of marker SNPs in each overlapping AI region, extending the marker alleles in each sweep to include all SNPs that exceeded the 80th percentile of mean neutral F_{st} values (see section 2.1) in order to provide a reasonable amount of SNPs for this calculation. Amongst the set of 25 AI regions that overlapped a sweep, 18 also overlapped at least one marker SNP across 12 separate sweeps, though many only covered a handful of marker SNPs (Extended Data Fig. 5). When considering AI regions overlapping at least 4 marker SNPs (after accounting for missingness), the maximum proportion of matching marker alleles at a AI locus for any hominin genome was ~82%, with the majority of cases having less than half the total number of marker alleles (median proportion of shared marker alleles ~40%, dependent on the hominin allele being present; Extended Data Fig. 5).

Our results suggest that AI variants are more common amongst our hard sweeps than expected, but that available hominin genomes do not provide particularly close matches for the putative causal haplotypes when a reasonable number of marker SNPs are available. Further investigation of the overlapping AI loci showed that they tend to occur on the periphery of the sweep regions that they overlap, often at a considerable distance from the peak selection signals from SweepFinder2 CLR scores (Extended Data Fig. 6). The latter finding suggests that introgressed hominin variants are unlikely to be the causal alleles for most sweeps – rather the enrichment of AI signals amongst our hard sweeps may largely be an artefact caused by introgressed hominin alleles hitchhiking on a beneficial human-derived variant.

Because our sweeps might have been caused by an as yet undetected AI locus, we also applied the recently published *admixfrog*[48] method to the 27 Anatolian EF samples to investigate if any introgressed hominin haplotypes were segregating at high frequencies near to the CLR peaks in our sweeps. We chose the Anatolian samples for this test as this population had the most significant hard sweep signals amongst the three ancestral populations that make up modern European ancestry (WHG and Steppe being the other two), and causal allele frequencies in this population are unaffected by subsequent Holocene admixture events. To identify Neandertal introgressed segments in the Anatolian EF genomes

we used *admixfrog* version 0.6.2.dev6 (https://github.com/BenjaminPeter/admixfrog/)[48] applying the parameters used in ref. [21]. In short, we used three high-coverage Neandertal genomes[198–200] and 46 African genomes from the Simons Genome Diversity Project[99] as reference genomes, and the panTro4 chimpanzee genome to assign the ancestral hominin allele state for each of the ~1.1 million probe SNPs used in this study.

We detected thousands of small to moderate length introgressed hominin loci (mean introgressed locus length = 69.6kb, each sample having around 3,890 individual introgressed loci > 10kb on average) amongst the 28 Anatolian EF samples, all coming from Neandertals. We calculated the frequency of the introgressed loci overlapping our sweep regions by counting the number of introgressed hominin alleles at each site in a sweep (noting that when multiple regions overlapped in a single individual, we kept only the highest scoring introgressed region from the overlapping set of loci). Several introgressed loci were found to be segregating at intermediate frequencies – 34 of 56 (~60%) sweeps had a segregating introgressed locus above 20% frequency – with the maximum frequency observed at any introgressed site that overlapped a sweep being ~78% (*MLPH* and *ZMYM6*; Figs. 3, Extended Data 6). However, all introgressed loci segregating at intermediate frequencies (>20%) tended to be hundreds of thousands of base pairs distant from the maximum CLR score of any sweep (Figs. 3, Extended Data 6), echoing the patterns observed for the overlapping AI loci obtained from previous studies. Moreover, the frequency of introgressed loci was significantly reduced near to the peak CLR score in each sweep relative to regions >100kb from this peak (the mean frequency of introgressed fragments minus two times the local standard error being greater mean frequency of introgressed fragments at the peak CLR score), which is consistent with positive selection on human-derived variants resulting in the removal of introgressed hominin loci near to the selected site (purifying selection on introgressed variants not expected to show the same relationship with the peak CLR score). Accordingly, we find no strong support that any of our sweeps are caused by introgressed archaic hominin variants, and posit that the majority of these overlapping AI regions are likely to be artefacts created by introgressed neutral alleles that are hitchhiking on beneficial human backgrounds.

Extended Data Figures

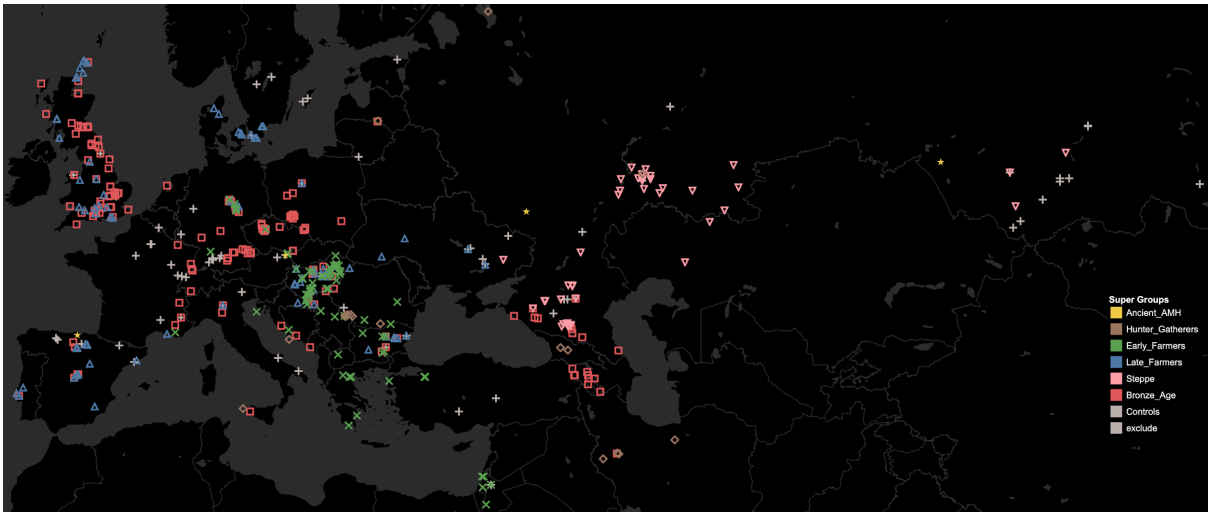

Figure 1. Full geographic distribution of the 1,162 ancient Eurasian samples used in this study with different populations (indicated by coloured symbols) classified into broader groupings according to archaeological records of material culture and lifestyle.

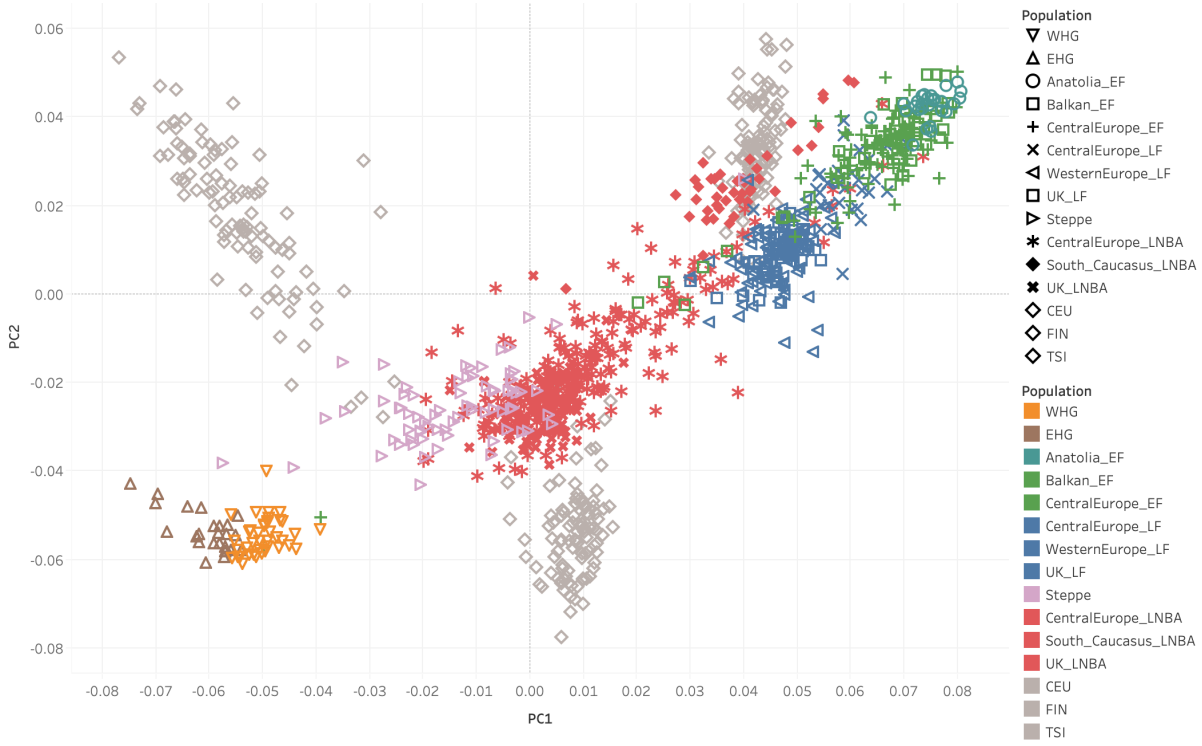

Figure 2. PCA analysis showing the genetic relationships between the different populations used in this study with all ancient populations distinguished (see key).

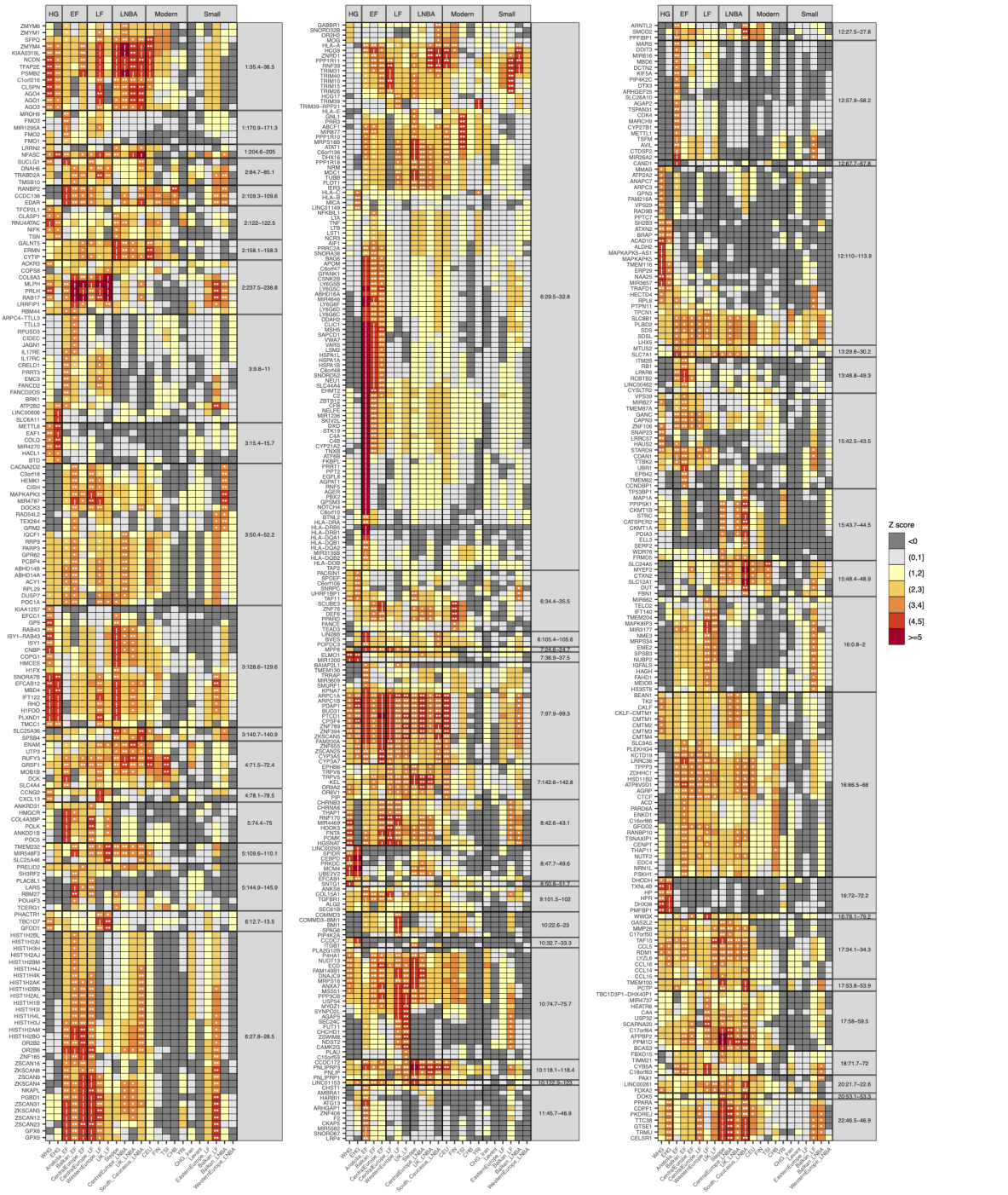

Figure 3. The *Z*-score for each outlier gene observed amongst the 57 candidate sweeps (y-axis), for all modern and ancient populations (x-axis). Populations are organized chronologically according to broad archaeological categorizations, with populations that had insufficient data for robust inferences labelled as 'Small' (i.e. $n_{\text{eff}} < 10$). Sweeps and genes are organized by chromosomal position. Outlier genes are indicated with a specific symbol according to their q-value ($0.05 \leq q < 0.10$ = *; $0.01 \leq q < 0.05$ = **; $q < 0.01$ = !)

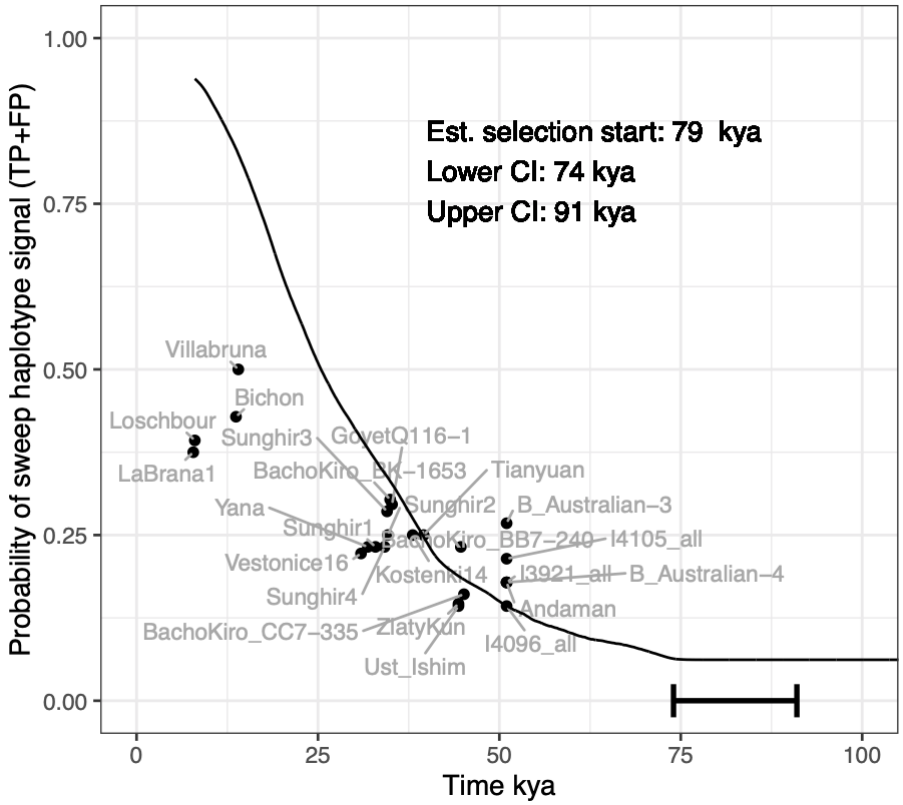

Figure 4a. Model-based estimate of sweep accumulation across time. Sweeps were simulated under a demographically-informed coalescent model and the resulting trajectories were aggregated across time and compared to individual sweep haplotype proportions (out of the 57 inferred sweep loci) observed in a mixture of samples from dated samples from Eurasia samples and Oceania (the latter samples dated at the plausible point of separation from the Main Eurasian branch; i.e. ~51kya). Sweeps starting ≤ 79kya show the best fit with real data (black line). Notably, samples from the Mesolithic (which were not included in the model) have >30% fewer sweeps being observed than expected if sweeps continued to aggregate at the same rate. This may result from our model conditioning on the presence of sweep signatures in a hypothetical single panmictic Mainland Eurasian population during the Neolithic, whereas the sampled individuals come from structured Mesolithic West Eurasian populations. Each population is expected to accumulate local sweeps, increasing the number of sweeps overall but depressing the proportion of observed sweep haplotypes in any single subpopulation below what is expected under a panmictic model. Also note that the expected probability of sweep presence does not decrease below 6%, as this is the FPR for sweep detection which is directly incorporated into the model.

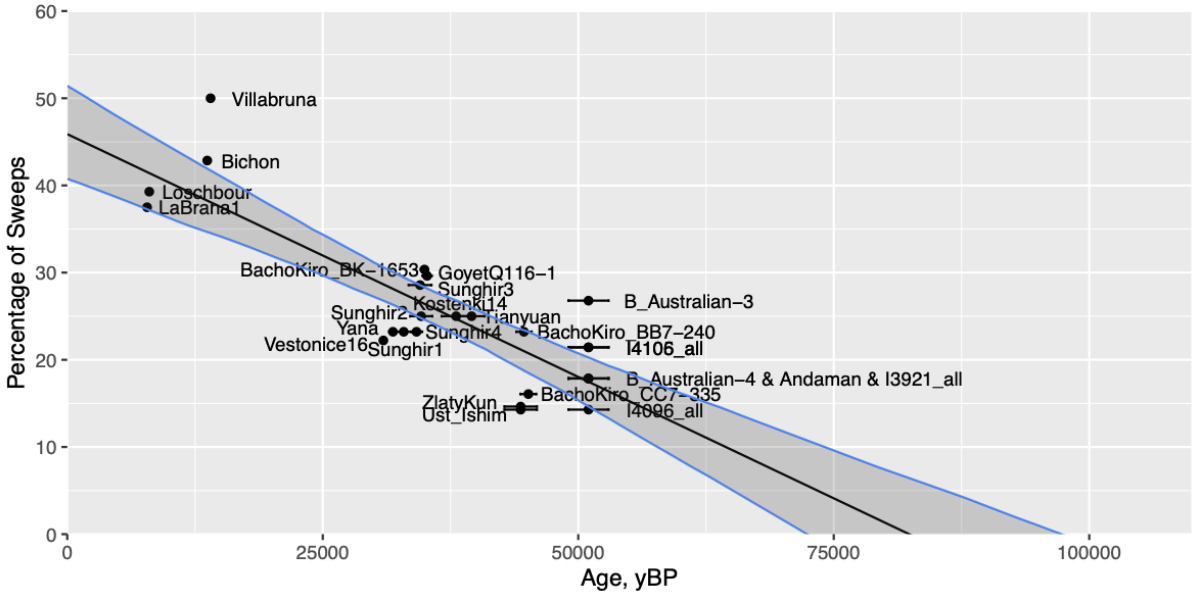

Figure 4b. Linear Regression estimate of sweep accumulation across time. The proportion of sweeps detected in dated samples from Eurasia samples and Oceania (the latter samples dated at the plausible point of separation from the Main Eurasian branch; i.e. ~51kya) were explored under the assumption that sweeps have aggregated in an approximately linear fashion since humans first left Africa. We explored the age when this linear model suggests there would have been 0 sweeps (i.e. the predicted x-intercept in the regression model), as this provides a rough estimate for the initial isolation of the OOA population. The uncertainties in the ages of each sample are represented on the figure as horizontal lines, and were accommodated in the modelling through use of bootstrap techniques. The mean x-intercept was determined to be 83,141 yBP and the associated empirical 95% confidence interval to be between 72,369 and 97,284 yBP.

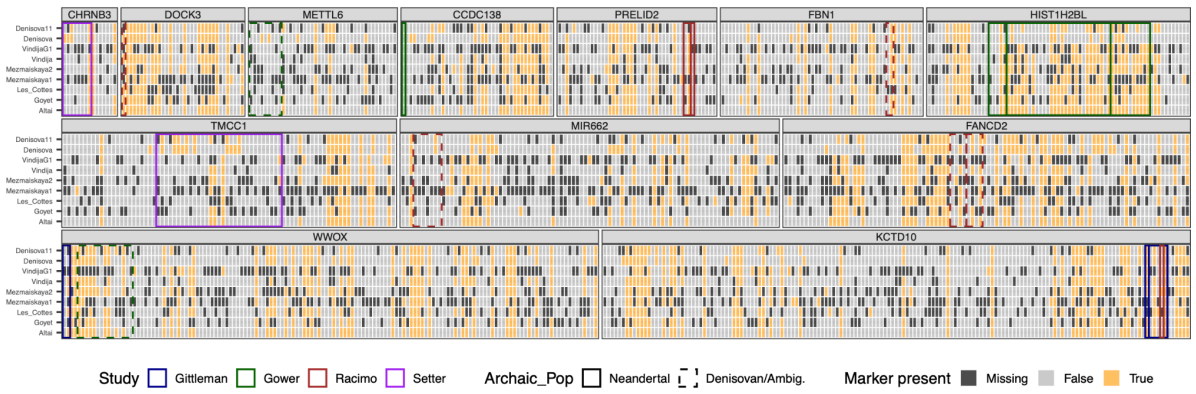

Figure 5. Putatively adaptively introgressed regions overlapping sweeps. The allele carried at marker SNPs is shown for all 12 sweep regions where an AI locus overlaps at least one marker SNP (orange = same allele as inferred sweep allele, grey = alternate allele from inferred sweep allele, black = no allele called in sample). AI regions are indicated by colored boxes, with different colours for each study (see key); boxes with solid lines depict AI loci observed in Neandertals, dashed lines indicate AI loci observed in Denisovans or where the hominin donor was ambiguous.

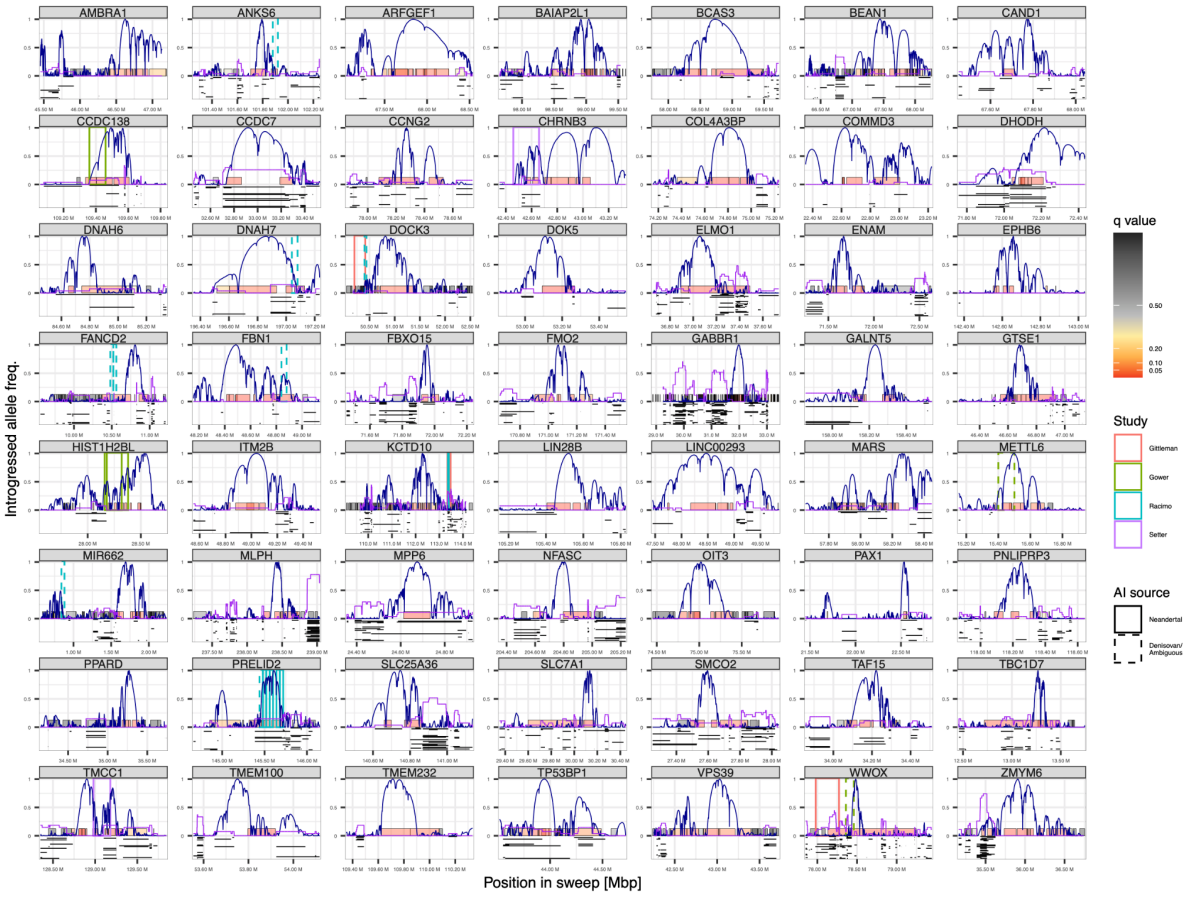

Figure 6. Introgressed archaic hominin loci in the vicinity of ancient hard sweeps. Similar to Fig. 3, but showing results for all 56 sweeps that were used in dating analyses. See Fig 3. for a full explanation of the plot items. As with Figure S5, the coloured boxes indicate the position of putative AI loci, with different colours for each study (see key).

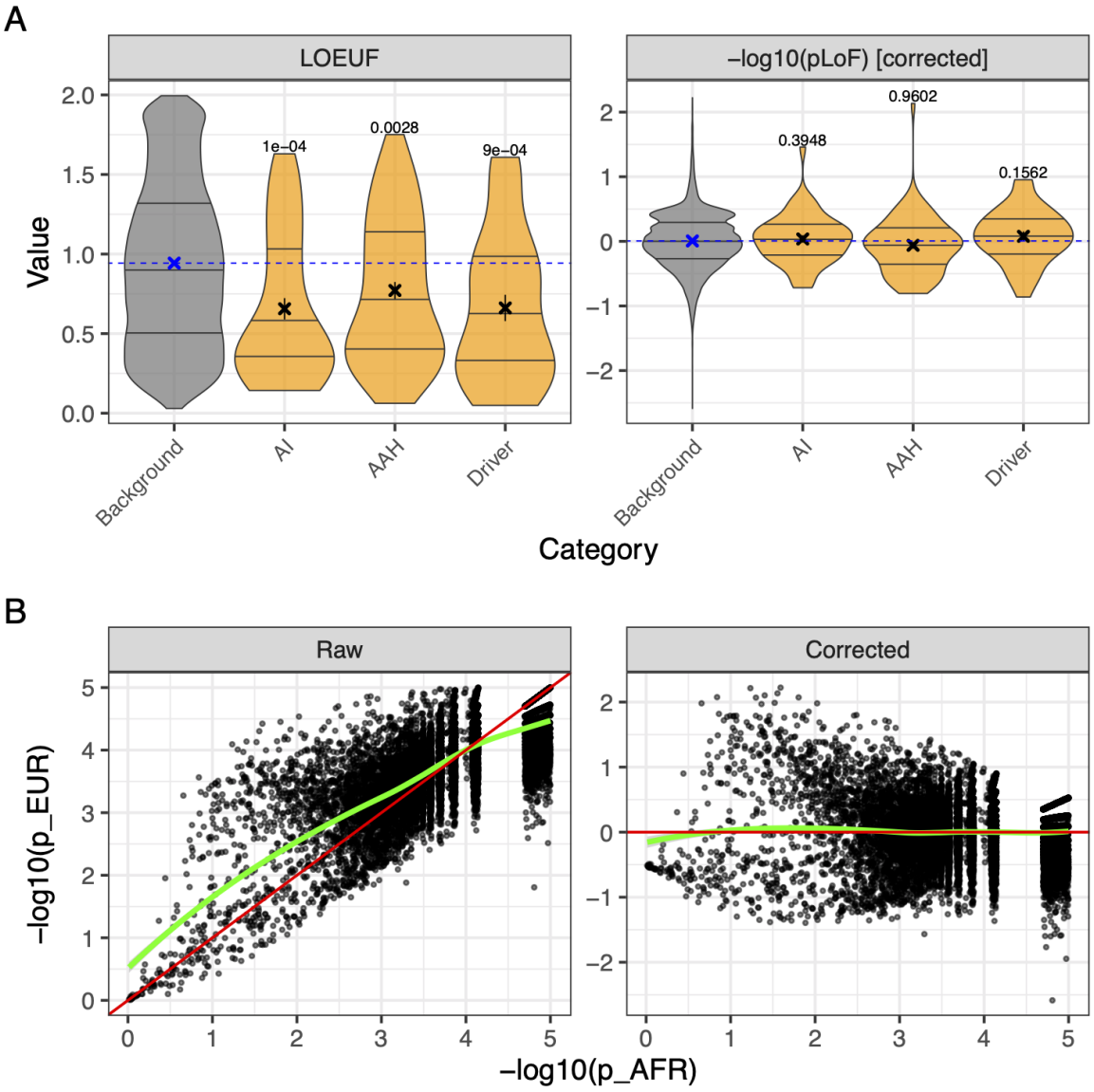

Figure 7. Distribution of LOEUF and rescaled *p* statistics. A. The distribution of LOEUF scores and rescaled *p* statistics across putative driver genes detected for the sweeps in this study, and also adaptively introgressed homini loci and selected genes previously reported from Artic human populations (orange violins), relative to the scores on all remaining genes (grey violins). P values from permutation tests are shown at the top of each violin. LOEUF scores are significantly lower in each of three categories, but this is not observed for the related p statistic, suggesting that the significant LOEUF scores are not due to the loss of deleterious variation due to positive selection at each sweep locus. **B.** Raw and corrected *p* statistic values for European populations after rescaling to account for African values.

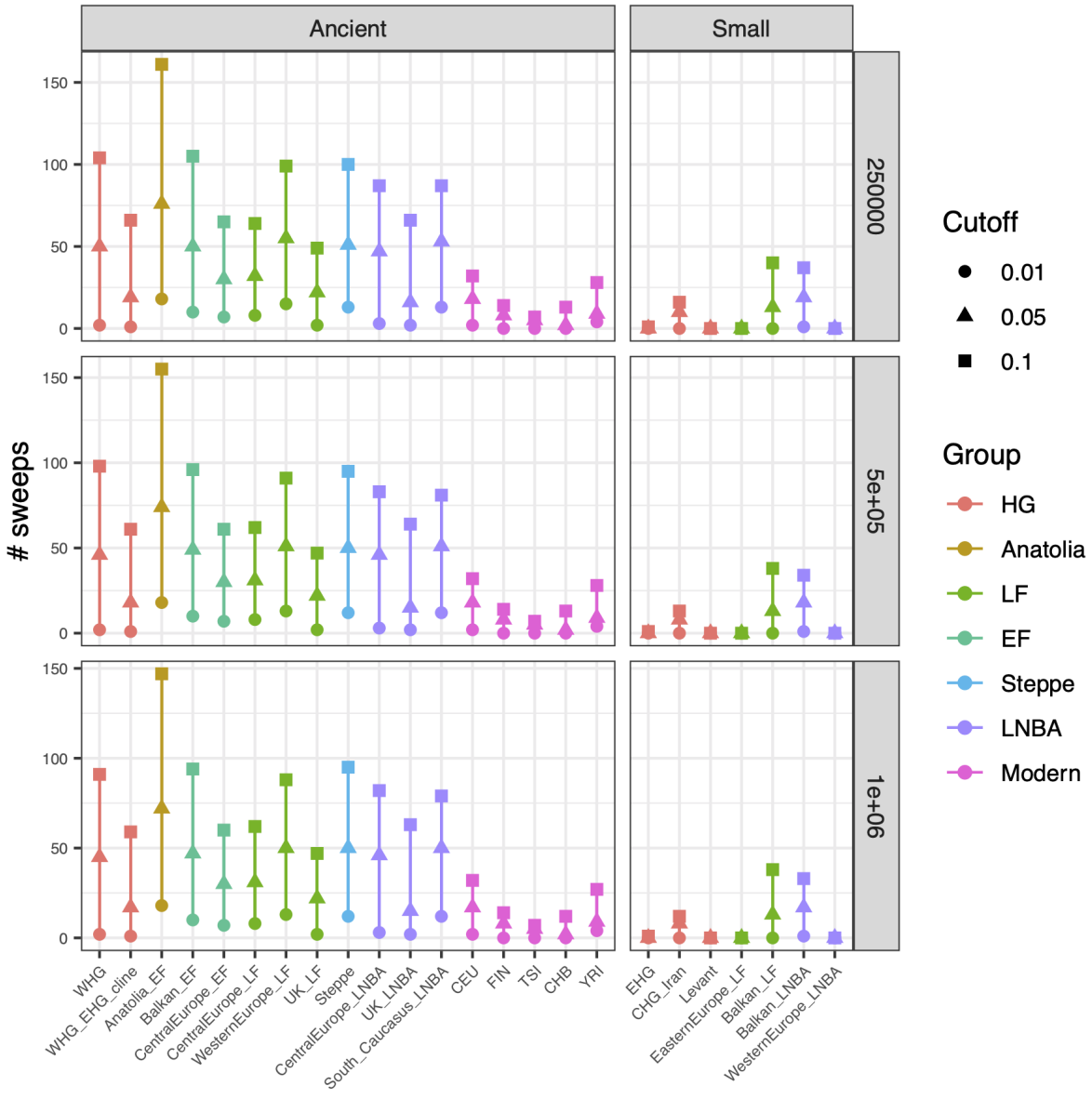

Figure 8. The frequency of sweeps across populations according to different *q*-value thresholds and joint widths. Sweeps were defined by grouping together genes within 250kb, 500kb, or 1Mb (*i.e.* join width; row panels) of each other that fell under a specific *q*-value threshold (see key). Populations are split into two groups based on whether they had sufficient data to make robust inferences ($n_{eff} \geq 10$ = 'Large') or not ($n_{eff} < 10$ = 'Small').

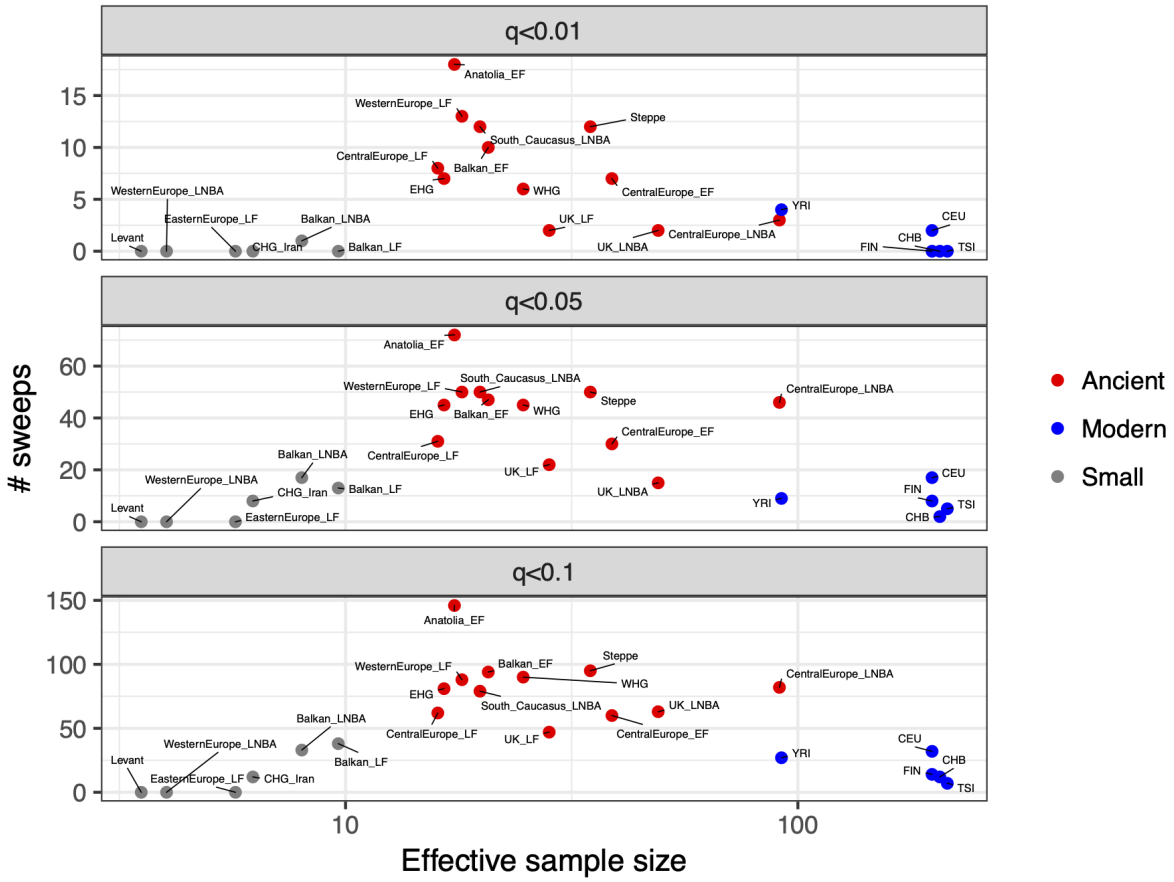

Figure 9. The frequency of sweeps across populations relative to the effective sample size for each population. Horizontal panels show different *q*-values used to determine the outlier sweeps. Notable decreases in the number of observed sweeps occur for small populations (i.e. $n_{\text{eff}} < 10$; gray dots) and also modern European populations (blue dots), though only the former is likely to have been a consequence of low statistical power.

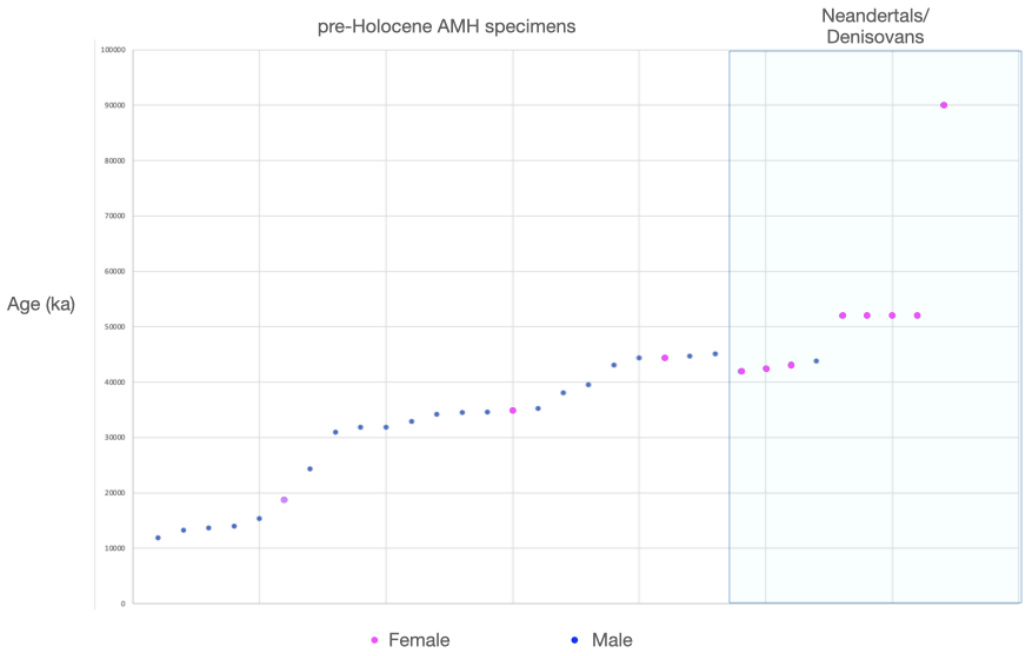

Figure 10. Age and genetically determined sex of pre-Holocene AMH and archaic hominin specimens. All pre-Holocene AMH samples from the Allen Ancient DNA Resource version 42.4 with sufficient genomic coverage to pass the quality filtering criteria. Most are western Eurasian in origin, and include specimens associated with IUP, Aurignacian, Gravettian, and Magdalenian cultures. There is a marked dominance of male specimens (87%) in the pre-Holocene AMH specimens, whereas in comparison available archaic hominins with sufficient genomic coverage appear to show the reverse pattern. A number of the older hominid ages are mid-point estimates from a sizable range.

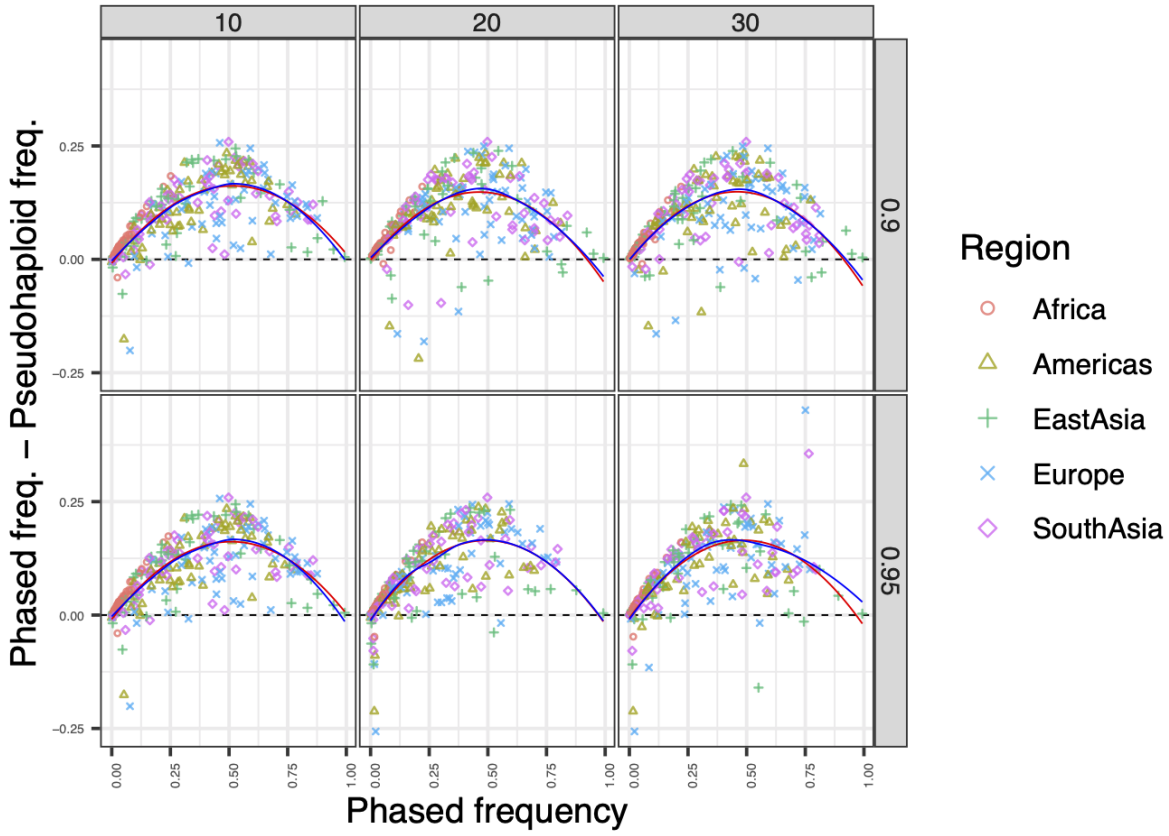

Figure 11. Sweep haplotype frequency differences between phased and pseudohaploid samples. Phased samples are shown on the x-axis, and the difference between pseudohaploid and phased frequencies on the x-axis for modern human samples from the Thousand Genomes Project (averaged across continental regions; see key). The imposition of pseudohaploidy results in an underestimation of the sweep haplotype that peaks at intermediate frequencies, as a direct result of mixing of sweep and non-sweep alleles in increasingly prevalent heterozygous individuals. This bias was consistent across all combinations of SNP marker number (column panels) and minimum marker allele presence (row panels). The blue and red lines show the fit for LOESS and quadratic regressions, respectively.

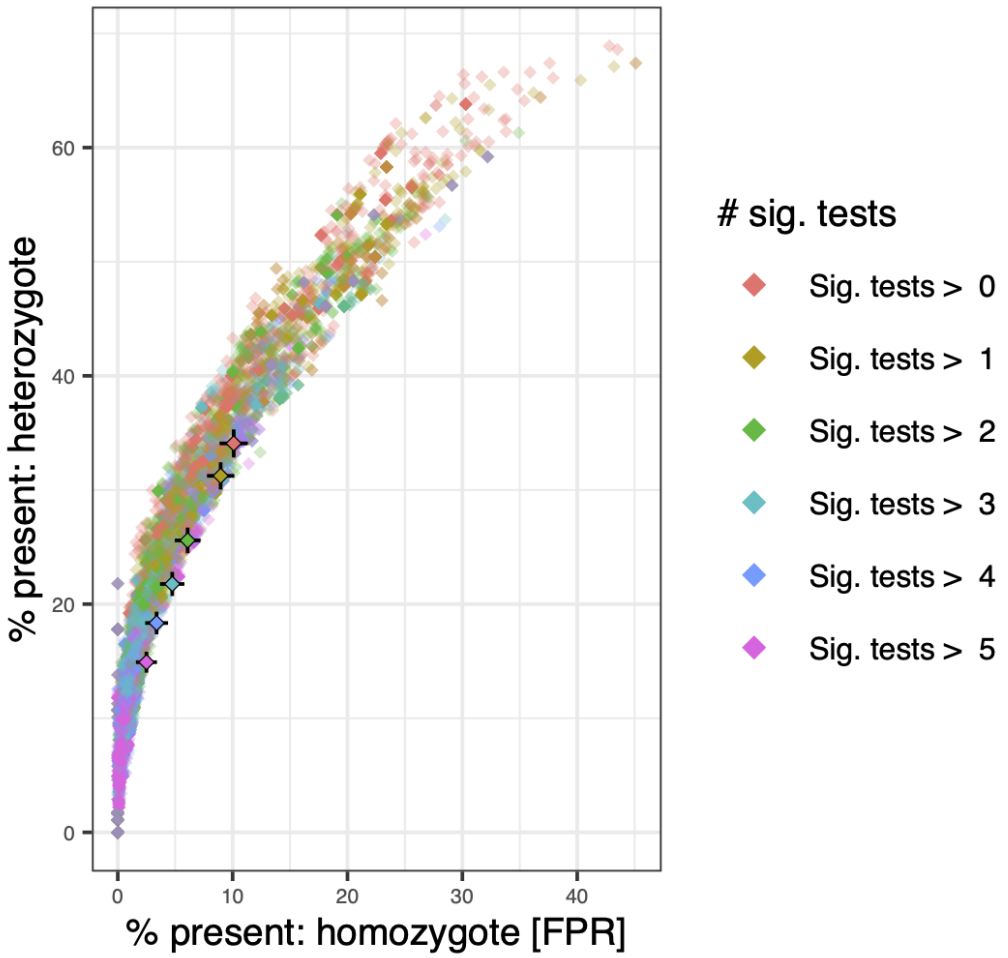

Figure 12. Estimated sweep presence in non-sweep homozygotes and heterozygotes for simulated ancient samples. Each individual sweep and sample combination are indicated and are colored according to the different sweep detection criteria. Sweep detection becomes more stringent in accordance with the minimum number of significant tests it must satisfy, which equates to the number of criteria combinations (out of six) in which the sweep was detected. For the criteria used in the current study, we expect sweeps will be called in around 6% of non-sweep homozygotes and 26% of sweep heterozygotes. Sweeps are always expected to be called in sweep homozygotes, see methods.

Extended Data Table Captions

Extended Data Table 1. Full set of metadata for all ancient western Eurasian samples used for hard sweep detection in this study. Data includes sample ages (largely calibrated with IntCal13), geographic locations, as well as associated cultural and historical affinities for each individual. This information is combined with genetic relationships between samples in order to categorise the populations that were used for selection analyses.

Extended Data Table 2. List of 57 outlier sweeps at *q*-value < 0.01 and associated metadata. Information includes chromosomal position; sweep labels (positional = SweepID, gene-based = Sweep_label); functional status (Functional_criterion); estimated sweep onset (oldest observed sample = Earliest_ancient_Eurasian; classification based on combined data = Sweep_time_consensus); previous evidence of selection (i.e. evidence of positive selection in European or ancient Eurasian lineages taken from the three previous studies: P=doi.org/10.1101/gr.087577.108[201], R=doi.org/10.1101/gr.246777.118[202], S=doi.org/10.1371/journal.pgen.1005928[203]); populations where at least one sweep gene was observed at $q < 0.1$ (Population_Q<0.10); and list of all genes observed in the sweep boundaries (AllGenes).

Extended Data Table 3. Estimates of sweep haplotype presence across ancient and modern samples used in this study. Related cultural, geographical, and chronological metadata from the Allen Ancient DNA Resource version 42.4 are included for each sample. For individual ancient samples, sweep haplotype presence is indicated using a binary indicator (i.e. present = 1 and absent = 0), whereas sweep frequencies are provided for modern populations (after pseudohaploidization). Separate tables are included for three different sweep detection criteria; the current study is based on sweep detection in ≥ 2 criteria (see methods). The additional summary table highlights the sharing of sweeps within and between Oceanian samples and early Eurasian samples, and details of IUP, Aurignacian, Gravettian and Magdalenian samples, and sex data from pre-Holocene specimens used in the criteria analyses.

Extended Data Table 4. Summary of *iSAFE* analyses. For each sweep, genes with at least one of the top 20 ranked SNPs (i.e. those with the highest 20 iSAFE scores) are indicated (up to five genes in each sweep). Genes are ranked according to the number of the top 20 SNPs they carry. The proportion of top ranked SNPs carried by each gene are shown on the left

hand columns. Gene names are listed in the right hand columns, along with the ratio of the top 20 SNPs to the total number of SNPs found in the gene (in parentheses).

Extended Data Table 5. Summary of the functional analyses. Results are provided for (a) driver genes identified within Eurasion hard sweep genes, (b) adaptively introgressed hominin loci, and (c) adaptive genes in Arctic human populations on separate sheets. The biological significance [function] of genes in each set was assessed by a systematic review of human disease literature, as well as animal and cell knockout phenotypic data sets, available on GnomAD, OMIM, PubMed and GeneCards databases.

Extended Data Table 6. List of ancient and modern samples used to detect and quantify sweep haplotypes dynamics. Abbreviated labels used in the study are provided along with the full names and associated metadata from the Allen Ancient DNA Resource version 42.4.

Extended Data Table 7. Radiocarbon ages of key ancient Eurasian samples recalibrated with IntCal20. Key samples with large temporal impacts on the quantification of sweep dynamics were recalibrated with IntCal20 [82].

Extended Data Table 8. Gorilla enrichment test results for the 32 driver genes from the current study. All results were obtained from http://cbl-gorilla.cs.technion.ac.il/, and are described in section 3.7.

Supplementary Data Figure Captions

Supplementary Data Figures 1-56. **Selection information for all sweeps.** Each figure contains selection information for a single sweep, with the top two panels showing the iSAFE scores used to determine driver genes and F_{st} scores used to define marker SNPs in sweep haplotype analyses, respectively, in relation to their genomic coordinates (x-axis). The remaining panels show the genomic distribution of the local SweepFinder2 composite likelihood ratio (CLR) scores (blue line). Genes within the sweep region are represented by colored boxes, with the color intensity indicating their Z-scores. The population with the single highest gene score is indicated with the [Focal] label, with all genes with $q < 0.1$ are labelled for this population. The figures are available on figshare (https://figshare.com/s/7d3711649677746bc231).

References

1. Pritchard, J. K., Pickrell, J. K. & Coop, G. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20**, R208–15 (2010).
2. Huber, C. D., Nordborg, M., Hermisson, J. & Hellmann, I. Keeping it local: evidence for positive selection in Swedish Arabidopsis thaliana. *Mol. Biol. Evol.* **31**, 3026–3039 (2014).
3. Harris, R. B., Sackman, A. & Jensen, J. D. On the unfounded enthusiasm for soft selective sweeps II: Examining recent evidence from humans, flies, and viruses. *PLoS Genet.* **14**, e1007859 (2018).
4. Haak, W. *et al.* Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
5. Schrider, D. R. & Kern, A. D. Soft Sweeps Are the Dominant Mode of Adaptation in the Human Genome. *Mol. Biol. Evol.* **34**, 1863–1877 (2017).
6. Sohail, M. *et al.* Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *Elife* **8**, (2019).
7. Souilmi, Y. *et al.* Ancient human genomes reveal a hidden history of strong selection in Eurasia. *Cold Spring Harbor Laboratory* 2020.04.01.021006 (2020) doi:10.1101/2020.04.01.021006.
8. Marciniak, S. & Perry, G. H. Harnessing ancient genomes to study the history of human adaptation. *Nat. Rev. Genet.* **18**, 659–674 (2017).
9. Huber, C. D., DeGiorgio, M., Hellmann, I. & Nielsen, R. Detecting recent selective sweeps while controlling for mutation rate and background selection. *Mol. Ecol.* (2015) doi:10.1111/mec.13351.
10. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
11. Szpak, M., Xue, Y., Ayub, Q. & Tyler-Smith, C. How well do we understand the basis of classic selective sweeps in humans? *FEBS Lett.* **593**, 1431–1448 (2019).
12. Bergström, A., Stringer, C., Hajdinjak, M., Scerri, E. M. L. & Skoglund, P. Origins of modern human ancestry. *Nature* **590**, 229–237 (2021).
13. Lipson, M. & Reich, D. A Working Model of the Deep Relationships of Diverse Modern Human Genetic Lineages Outside of Africa. *Mol. Biol. Evol.* **34**, 889–902 (2017).
14. Bergström, A. *et al.* Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, (2020).
15. O'Connell, J. F. *et al.* When did Homo sapiens first reach Southeast Asia and Sahul? *Proc. Natl. Acad. Sci. U. S. A.* **115**, 8482–8490 (2018).
16. Sun, X.-F. *et al.* Ancient DNA and multimethod dating confirm the late arrival of anatomically modern humans in southern China. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
17. Hallast, P., Agdzhoyan, A., Balanovsky, O., Xue, Y. & Tyler-Smith, C. A Southeast Asian origin for present-day non-African human Y chromosomes. *Hum. Genet.* **140**, 299–307 (2021).
18. Posth, C. *et al.* Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr. Biol.* **26**, 827–833 (2016).
19. Groucutt, H. S. *et al.* Homo sapiens in Arabia by 85,000 years ago. *Nat Ecol Evol* **2**, 800–809 (2018).
20. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
21. Hajdinjak, M. *et al.* Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry. *Nature* **592**, 253–257 (2021).
22. Higham, T. *et al.* Precision dating of the Palaeolithic: a new radiocarbon chronology for the Abri Pataud (France), a key Aurignacian sequence. *J. Hum. Evol.* **61**, 549–563 (2011).
23. Banks, W. E. *et al.* An application of hierarchical Bayesian modeling to better constrain the

chronologies of Upper Paleolithic archaeological cultures in France between ca. 32,000–21,000 calibrated years before present. *Quat. Sci. Rev.* **220**, 188–214 (2019).
24. Cooper, A. *et al.* A global environmental crisis 42,000 years ago. *Science* **371**, 811–818 (2021).
25. Dannemann, M., Andrés, A. M. & Kelso, J. Introgression of Neandertal- and Denisovan-like Haplotypes Contributes to Adaptive Variation in Human Toll-like Receptors. *Am. J. Hum. Genet.* **98**, 22–33 (2016).
26. Materials and methods are available as supplementary materials.
27. Akbari, A. *et al.* Identifying the favored mutation in a positive selective sweep. *Nature Methods* vol. 15 279–282 (2018).
28. Fumagalli, M. *et al.* Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* **349**, 1343–1347 (2015).
29. Saltykova, M. M. The Main Physiological Mechanisms of Cold Adaptation in Humans. *Neuroscience and Behavioral Physiology* vol. 48 543–550 (2018).
30. Yudin, N. S., Larkin, D. M. & Ignatieva, E. V. A compendium and functional characterization of mammalian genes involved in adaptation to Arctic or Antarctic environments. *BMC Genet.* **18**, 111 (2017).
31. Fang, L. *et al.* PPARgene: A Database of Experimentally Verified and Computationally Predicted PPAR Target Genes. *PPAR Res.* **2016**, 6042162 (2016).
32. Will, M., Krapp, M., Stock, J. T. & Manica, A. Different environmental variables predict body and brain size evolution in Homo. *Nat. Commun.* **12**, 4116 (2021).
33. Corbetta, M., Patel, G. & Shulman, G. L. The reorienting system of the human brain: from environment to theory of mind. *Neuron* **58**, 306–324 (2008).
34. Mäkinen, T. M. Human cold exposure, adaptation, and performance in high latitude environments. *Am. J. Hum. Biol.* **19**, 155–164 (2007).
35. Tierney, J. E., deMenocal, P. B. & Zander, P. D. A climatic context for the out-of-Africa migration. *Geology* **45**, 1023–1026 (2017).
36. Williams, M. A. J. *et al.* Reply to the comment on 'Environmental impact of the 73ka Toba super-eruption in South Asia' by M. A. J. Williams, S. H. Ambrose, S. van der Kaars, C. Ruehlemann, U. Chattopadhyaya, J. Pal, P. R. Chauhan [Palaeogeography, Palaeoclimatology, Palaeoecology 284 (2009) 295–314]. *Palaeogeography, Palaeoclimatology, Palaeoecology* vol. 296 204–211 (2010).
37. Brunton, P. J. & Russell, J. A. The expectant brain: adapting for motherhood. *Nat. Rev. Neurosci.* **9**, 11–25 (2008).
38. Krashes, M. J., Lowell, B. B. & Garfield, A. S. Melanocortin-4 receptor--regulated energy homeostasis. *Nat. Neurosci.* **19**, 206–219 (2016).
39. Nakamura, K. & Morrison, S. F. A thermosensory pathway that controls body temperature. *Nat. Neurosci.* **11**, 62–71 (2008).
40. Zammit, N. W. *et al.* Denisovan, modern human and mouse TNFAIP3 alleles tune A20 phosphorylation and immunity. *Nat. Immunol.* **20**, 1299–1310 (2019).
41. Gittelman, R. M. *et al.* Archaic Hominin Admixture Facilitated Adaptation to Out-of-Africa Environments. *Curr. Biol.* **26**, 3375–3382 (2016).
42. Teixeira, J. C. & Cooper, A. Using hominin introgression to trace modern human dispersals. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15327–15332 (2019).
43. Hublin, J.-J. *et al.* Initial Upper Palaeolithic Homo sapiens from Bacho Kiro Cave, Bulgaria. *Nature* **581**, 299–302 (2020).
44. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
45. Reiter, J. F. & Leroux, M. R. Genes and molecular pathways underpinning ciliopathies. *Nat. Rev.*

Mol. Cell Biol. **18**, 533–547 (2017).
46. Saklayen, M. G. The Global Epidemic of the Metabolic Syndrome. *Curr. Hypertens. Rep.* **20**, 12 (2018).
47. Hebert, L. E., Weuve, J., Scherr, P. A. & Evans, D. A. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. *Neurology* **80**, 1778–1783 (2013).
48. Peter, B. M. 100,000 years of gene flow between Neandertals and Denisovans in the Altai mountains. *bioRxiv* 2020.03.13.990523 (2020) doi:10.1101/2020.03.13.990523.
49. Rasmussen, S. O. *et al.* A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy. *Quaternary Science Reviews* vol. 106 14–28 (2014).
50. Mathieson, I. *et al.* Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
51. Mathieson, I. *et al.* The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
52. Eisenmann, S. *et al.* Reconciling material cultures in archaeology with genetic data: The nomenclature of clusters emerging from archaeogenomic analysis. *Sci. Rep.* **8**, 13003 (2018).
53. Smith, J. M. & Haigh, J. The hitch-hiking effect of a favourable gene. *Genetical Research* vol. 23 23–35 (1974).
54. Hermisson, J. & Pennings, P. S. Soft sweeps and beyond: Understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods in Ecology and Evolution* **8**, 700–716 (2017).
55. Höllinger, I., Pennings, P. S. & Hermisson, J. Polygenic adaptation: From sweeps to subtle frequency shifts. *PLoS Genet.* **15**, e1008035 (2019).
56. Stephan, W. Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. *Mol. Ecol.* **25**, 79–88 (2016).
57. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
58. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* (2013).
59. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
60. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
61. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
62. Kircher, M. Analysis of High-Throughput Ancient DNA Sequencing Data. in *Ancient DNA: Methods and Protocols* (eds. Shapiro, B. & Hofreiter, M.) 197–228 (Humana Press, 2012).
63. Jun, G., Wing, M. K., Abecasis, G. R. & Kang, H. M. An efficient and scalable analysis framework for variant extraction and refinement from population-scale DNA sequence data. *Genome Res.* **25**, 918–925 (2015).
64. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
65. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
66. Patterson, N., Price, A. L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
67. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer

datasets. *Gigascience* **4**, 7 (2015).
68. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
69. DeGiorgio, M., Huber, C. D., Hubisz, M. J., Hellmann, I. & Nielsen, R. S weep F inder 2: increased sensitivity, robustness and flexibility. *Bioinformatics* **32**, 1895–1897 (2016).
70. Nielsen, R. *et al.* Genomic scans for selective sweeps using SNP data. *Genome Res.* **15**, 1566–1575 (2005).
71. Kinsella, R. J. *et al.* Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database* vol. 2011 bar030–bar030 (2011).
72. Storey, J. D. The positive false discovery rate: a Bayesian interpretation and the q -value. *The Annals of Statistics* vol. 31 2013–2035 (2003).
73. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols* vol. 4 1184–1191 (2009).
74. Iglewicz, B. & Hoaglin, D. C. *How to Detect and Handle Outliers*. (Asq Press, 1993).
75. Daub, J. T. *et al.* Evidence for Polygenic Adaptation to Pathogens in the Human Genome. *Molecular Biology and Evolution* vol. 30 1544–1558 (2013).
76. Ferrer-Admetlla, A., Liang, M., Korneliussen, T. & Nielsen, R. On detecting incomplete soft or hard selective sweeps using haplotype structure. *Mol. Biol. Evol.* **31**, 1275–1291 (2014).
77. Choudhury, A. *et al.* High-depth African genomes inform human migration and health. *Nature* **586**, 741–748 (2020).
78. Gel, B. *et al.* regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics* **32**, 289–291 (2016).
79. Bhérer, C., Campbell, C. L. & Auton, A. Refined genetic maps reveal sexual dimorphism in human meiotic recombination at multiple scales. *Nat. Commun.* **8**, 14994 (2017).
80. Bhatia, G., Patterson, N., Sankararaman, S. & Price, A. L. Estimating and interpreting FST: The impact of rare variants. *Genome Res.* **23**, 1514 (2013).
81. Prüfer, K. *et al.* A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia. *Nat Ecol Evol* **5**, 820–825 (2021).
82. Reimer, P. J. *et al.* The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).
83. Gower, G. *et al.* Widespread male sex bias in mammal fossil and museum collections. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 19019–19024 (2019).
84. Petr, M. *et al.* The evolutionary history of Neanderthal and Denisovan Y chromosomes. *Science* **369**, 1653–1656 (2020).
85. Hershkovitz, I. *et al.* A Middle Pleistocene Homo from Nesher Ramla, Israel. *Science* **372**, 1424–1428 (2021).
86. Moorjani, P. *et al.* A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5652–5657 (2016).
87. Iasi, L. N. M., Ringbauer, H. & Peter, B. M. An extended admixture pulse model reveals the limitations to Human-Neandertal introgression dating. *bioRxiv* 2021.04.04.438357 (2021) doi:10.1101/2021.04.04.438357.
88. Tobler, R. *et al.* Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia. *Nature* **544**, 180–184 (2017).
89. David, B. *et al.* 45,610–52,160 years of site and landscape occupation at Nawarla Gabarnmang, Arnhem Land plateau (northern Australia). *Quat. Sci. Rev.* **215**, 64–85 (2019).
90. Maloney, T., O'Connor, S., Wood, R., Aplin, K. & Balme, J. Carpenters Gap 1: A 47,000 year

old record of indigenous adaption and innovation. *Quat. Sci. Rev.* **191**, 204–228 (2018).
91. Surovell, T. A. Early Paleoindian women, children, mobility, and fertility. *Am. Antiq.* **65**, 493–508 (2000).
92. Allen, J. & O'connell, J. F. A different paradigm for the initial colonisation of Sahul. *Archaeol. Ocean.* **55**, 1–14 (2020).
93. Rose, J. I. New Light on Human Prehistory in the Arabo-Persian Gulf Oasis. *Curr. Anthropol.* **51**, 849–883 (2010).
94. Fewlass, H. *et al.* A 14C chronology for the Middle to Upper Palaeolithic transition at Bacho Kiro Cave, Bulgaria. *Nat Ecol Evol* **4**, 794–801 (2020).
95. Harris, K. & Nielsen, R. The Genetic Cost of Neanderthal Introgression. *Genetics* **203**, 881–891 (2016).
96. Petraglia, M. D. Archaeology: Trailblazers across Arabia. *Nature* **470**, 50–51 (2011).
97. Armitage, S. J. *et al.* The southern route 'out of Africa': evidence for an early expansion of modern humans into Arabia. *Science* **331**, 453–456 (2011).
98. Moreno-Mayar, J. V. *et al.* Early human dispersals within the Americas. *Science* **362**, (2018).
99. Mallick, S. *et al.* The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. *Nature* **538**, 201–206 (2016).
100. Lipson, M. *et al.* Population Turnover in Remote Oceania Shortly after Initial Settlement. *Curr. Biol.* **28**, 1157–1165.e7 (2018).
101. Hoffecker, J. F. Out of Africa: modern human origins special feature: the spread of modern humans in Europe. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16040–16045 (2009).
102. Andrew Bennett, E. *et al.* The origin of the Gravettians: genomic evidence from a 36,000-year-old Eastern European. *Cold Spring Harbor Laboratory* 685404 (2019) doi:10.1101/685404.
103. Lazaridis, I. The evolutionary history of human populations in Europe. *Curr. Opin. Genet. Dev.* **53**, 21–27 (2018).
104. Channell, J. E. T., Singer, B. S. & Jicha, B. R. Timing of Quaternary geomagnetic reversals and excursions in volcanic and sedimentary archives. *Quat. Sci. Rev.* **228**, 106114 (2020).
105. Feldman, M. *et al.* Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).
106. Hoffmann, D. L. *et al.* U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science* **359**, 912–915 (2018).
107. Langley, M. C. *et al.* Bows and arrows and complex symbolic displays 48,000 years ago in the South Asian tropics. *Sci Adv* **6**, eaba3831 (2020).
108. R Core Team. R: A Language and Environment for Statistical Computing.
109. Canty, A., Ripley, B. & Others. boot: Bootstrap R (S-Plus) functions. *R package version* **1**, 3–20 (2017).
110. Kamm, J., Terhorst, J., Durbin, R. & Song, Y. S. Efficiently inferring the demographic history of many populations with allele count data. *J. Am. Stat. Assoc.* **115**, 1472–1487 (2020).
111. Ewing, G. & Hermisson, J. MSMS: a coalescent simulation program including recombination, demographic structure and selection at a single locus. *Bioinformatics* **26**, 2064–2065 (2010).
112. Paeth, H. *et al.* Comparison of climate change from Cenozoic surface uplift and glacial-interglacial episodes in the Himalaya-Tibet region: Insights from a regional climate model and proxy data. *Global and Planetary Change* vol. 177 10–26 (2019).
113. Kutzbach, J. E. *et al.* African climate response to orbital and glacial forcing in 140,000-y simulation with implications for early modern human environments. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 2255–2264 (2020).
114. Rostek, F., Bard, E., Beaufort, L., Sonzogni, C. & Ganssen, G. Sea surface temperature and

productivity records for the past 240 kyr in the Arabian Sea. *Deep Sea Research Part II: Topical Studies in Oceanography* vol. 44 1461–1480 (1997).
115. Scussolini, P. *et al.* Agreement between reconstructed and modeled boreal precipitation of the Last Interglacial. *Sci Adv* **5**, eaax7047 (2019).
116. Jennings, R. P. *et al.* The greening of Arabia: Multiple opportunities for human occupation of the Arabian Peninsula during the Late Pleistocene inferred from an ensemble of climate model simulations. *Quat. Int.* **382**, 181–199 (2015).
117. Turney, C. S. M. *et al.* A global mean sea surface temperature dataset for the Last Interglacial (129–116 ka) and contribution of thermal expansion to sea level change. *Earth System Science Data* vol. 12 3341–3356 (2020).
118. Al Harthy, L. & Grenyer, R. Classification and ordination of the main plant communities of the Eastern Hajar Mountains, Oman. *J. Arid Environ.* **169**, 1–18 (2019).
119. Amonkar, A., Iyer, S. D., Babu, E. V. S. S. K. & Manju, S. Extending the limit of widespread dispersed Toba volcanic glass shards and identification of new in-situ volcanic events in the Central Indian Ocean Basin. *J. Earth Syst. Sci.* **129**, 175 (2020).
120. Westgate, J. A. *et al.* Tephrochronology of the Toba tuffs: four primary glass populations define the 75-ka Youngest Toba Tuff, northern Sumatra, Indonesia. *J. Quat. Sci.* **28**, 772–776 (2013).
121. Rampino, M. R. & Self, S. Volcanic winter and accelerated glaciation following the Toba super-eruption. *Nature* **359**, 50–52 (1992).
122. Jones, G. S., Gregory, J. M., Stott, P. A., Tett, S. F. B. & Thorpe, R. B. An AOGCM simulation of the climate response to a volcanic super-eruption. *Clim. Dyn.* **25**, 725–738 (2005).
123. Williams, M. A. J. *et al.* Environmental impact of the 73ka Toba super-eruption in South Asia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **284**, 295–314 (2009).
124. Camillo, E., Quadros, J. P., Santarosa, A. C. A., Costa, K. B. & Toledo, F. A. L. An abrupt cooling event recorded around 73 kyr in western South Atlantic. *Quat. Int.* **542**, 80–87 (2020).
125. Du, W. *et al.* Timing and structure of the weak Asian Monsoon event about 73,000 years ago. *Quat. Geochronol.* **53**, 101003 (2019).
126. Zielinski, G. A. *et al.* Potential atmospheric impact of the Toba Mega-Eruption ~71,000 years ago. *Geophys. Res. Lett.* **23**, 837–840 (1996).
127. Williams, M. Did the 73 ka Toba super-eruption have an enduring effect? Insights from genetics, prehistoric archaeology, pollen analysis, stable isotope geochemistry, geomorphology, ice cores, and climate models. *Quat. Int.* **269**, 87–93 (2012).
128. Robock, A. *et al.* Did the Toba volcanic eruption of ~74 ka B.P. produce widespread glaciation? *J. Geophys. Res.* **114**, (2009).
129. Zhang, X. *et al.* The history and evolution of the Denisovan- haplotype in Tibetans. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
130. Moore, L. G. *et al.* Maternal adaptation to high-altitude pregnancy: an experiment of nature--a review. *Placenta* **25 Suppl A**, S60–71 (2004).
131. Moore, L. G. Human genetic adaptation to high altitude. *High Alt. Med. Biol.* **2**, 257–279 (2001).
132. Rees, J. S., Castellano, S. & Andrés, A. M. The Genomics of Human Local Adaptation. *Trends Genet.* **36**, 415–428 (2020).
133. Buck, L. T. *et al.* Evidence of different climatic adaptation strategies in humans and non-human primates. *Sci. Rep.* **9**, 11025 (2019).
134. Grav, H. J. & Blix, A. S. Brown adipose tissue-a factor in the survival of harp seal pups. *Can. J. Physiol. Pharmacol.* **54**, 409–412 (1976).
135. van der Lans, A. A. J. J. *et al.* Cold acclimation recruits human brown fat and increases nonshivering thermogenesis. *J. Clin. Invest.* **123**, 3395–3403 (2013).
136. Sanchez-Gurmaches, J. *et al.* Brown Fat AKT2 Is a Cold-Induced Kinase that Stimulates

ChREBP-Mediated De Novo Lipogenesis to Optimize Fuel Storage and Thermogenesis. *Cell Metab.* **27**, 195–209.e6 (2018).
137. Yang, Z. *et al.* Darwinian Positive Selection on the Pleiotropic Effects of KITLG Explain Skin Pigmentation and Winter Temperature Adaptation in Eurasians. *Mol. Biol. Evol.* **35**, 2272–2283 (2018).
138. Mäkinen, T. M. *et al.* Cold temperature and low humidity are associated with increased occurrence of respiratory tract infections. *Respir. Med.* **103**, 456–462 (2009).
139. Koskela, H. O. Cold air-provoked respiratory symptoms: the mechanisms and management. *Int. J. Circumpolar Health* **66**, 91–100 (2007).
140. Cardona, A. *et al.* Genome-wide analysis of cold adaptation in indigenous Siberian populations. *PLoS One* **9**, e98076 (2014).
141. Zhou, S. *et al.* Genetic architecture and adaptations of Nunavik Inuit. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 16012–16017 (2019).
142. Reynolds, A. W. *et al.* Comparing signals of natural selection between three Indigenous North American populations. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 9312–9317 (2019).
143. Mondal, M., Bertranpetit, J. & Lao, O. Approximate Bayesian computation with deep learning supports a third archaic introgression in Asia and Oceania. *Nat. Commun.* **10**, 246 (2019).
144. Sankararaman, S., Mallick, S., Patterson, N. & Reich, D. The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Curr. Biol.* **26**, 1241–1247 (2016).
145. Vernot, B. *et al.* Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* **352**, 235–239 (2016).
146. Jacobs, G. S. *et al.* Multiple Deeply Divergent Denisovan Ancestries in Papuans. *Cell* **177**, 1010–1021.e32 (2019).
147. Huerta-Sánchez, E. *et al.* Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194–197 (2014).
148. Gower, G., Picazo, P. I., Fumagalli, M. & Racimo, F. Detecting adaptive introgression in human evolution using convolutional neural networks. doi:10.1101/2020.09.18.301069.
149. Setter, D. *et al.* VolcanoFinder: Genomic scans for adaptive introgression. *PLoS Genet.* **16**, e1008867 (2020).
150. Trimmer, C. *et al.* Genetic variation across the human olfactory receptor repertoire alters odor perception. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 9475–9480 (2019).
151. Dhindsa, R. S., Copeland, B. R., Mustoe, A. M. & Goldstein, D. B. Natural Selection Shapes Codon Usage in the Human Genome. *Am. J. Hum. Genet.* **107**, 83–95 (2020).
152. Zhou, W., Wei, W. & Sun, Y. Genetically engineered mouse models for functional studies of SKP1-CUL1-F-box-protein (SCF) E3 ubiquitin ligases. *Cell Res.* **23**, 599–619 (2013).
153. Cardozo, T. & Pagano, M. The SCF ubiquitin ligase: insights into a molecular machine. *Nat. Rev. Mol. Cell Biol.* **5**, 739–751 (2004).
154. Morales, J. C. *et al.* Role for the BRCA1 C-terminal repeats (BRCT) protein 53BP1 in maintaining genomic stability. *J. Biol. Chem.* **278**, 14971–14977 (2003).
155. Ward, I. M., Minn, K., van Deursen, J. & Chen, J. p53 Binding protein 53BP1 is required for DNA damage responses and tumor suppression in mice. *Mol. Cell. Biol.* **23**, 2556–2563 (2003).
156. Monte, M. *et al.* Cloning, chromosome mapping and functional characterization of a human homologue of murine gtse-1 (B99) gene. *Gene* **254**, 229–236 (2000).
157. Yang, Q. *et al.* Severe Fanconi Anemia phenotypes in Fancd2 depletion mice. *Biochem. Biophys. Res. Commun.* **514**, 713–719 (2019).
158. Catinozzi, M. *et al.* The Drosophila FUS ortholog cabeza promotes adult founder myoblast selection by Xrp1-dependent regulation of FGF signaling. *PLoS Genet.* **16**, e1008731 (2020).
159. Ballarino, M. *et al.* TAF15 is important for cellular proliferation and regulates the expression of a

subset of cell cycle genes through miRNAs. *Oncogene* **32**, 4646–4655 (2013).
160. Bustamante-Marin, X. M. & Ostrowski, L. E. Cilia and Mucociliary Clearance. *Cold Spring Harb. Perspect. Biol.* **9**, (2017).
161. Lesko, S. L. & Rouhana, L. Dynein assembly factor with WD repeat domains 1 (DAW1) is required for the function of motile cilia in the planarian Schmidtea mediterranea. *Dev. Growth Differ.* **62**, 423–437 (2020).
162. Canty, J. T., Tan, R., Kusakci, E., Fernandes, J. & Yildiz, A. Structure and Mechanics of Dynein Motors. *Annu. Rev. Biophys.* **50**, 549–574 (2021).
163. Ohbayashi, N. & Fukuda, M. Recent advances in understanding the molecular basis of melanogenesis in melanocytes. *F1000Res.* **9**, (2020).
164. Hume, A. N., Ushakov, D. S., Tarafder, A. K., Ferenczi, M. A. & Seabra, M. C. Rab27a and MyoVa are the primary Mlph interactors regulating melanosome transport in melanocytes. *J. Cell Sci.* **120**, 3111–3122 (2007).
165. Crawford, N. G. *et al.* Loci associated with skin pigmentation identified in African populations. *Science* **358**, (2017).
166. Yousaf, S. *et al.* Molecular characterization of SLC24A5 variants and evaluation of Nitisinone treatment efficacy in a zebrafish model of OCA6. *Pigment Cell Melanoma Res.* **33**, 556–565 (2020).
167. Ménasché, G. *et al.* Griscelli syndrome restricted to hypopigmentation results from a melanophilin defect (GS3) or a MYO5A F-exon deletion (GS1). *J. Clin. Invest.* **112**, 450–456 (2003).
168. Aksu, G., Kütükçüler, N., Genel, F., Vergin, C. & Omowaire, B. Griscelli syndrome without hemophagocytosis in an eleven-year-old girl: expanding the phenotypic spectrum of Rab27A mutations in humans. *Am. J. Med. Genet. A* **116A**, 329–333 (2003).
169. Lee, S. T. *et al.* Mutations of the P gene in oculocutaneous albinism, ocular albinism, and Prader-Willi syndrome plus albinism. *N. Engl. J. Med.* **330**, 529–534 (1994).
170. Ishida, Y. *et al.* A homozygous single-base deletion in MLPH causes the dilute coat color phenotype in the domestic cat. *Genomics* **88**, 698–705 (2006).
171. Lamason, R. L. *et al.* SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* **310**, 1782–1786 (2005).
172. Kratochwil, C. F., Urban, S. & Meyer, A. Genome of the Malawi golden cichlid fish (Melanochromis auratus) reveals exon loss of oca2 in an amelanistic morph. *Pigment Cell Melanoma Res.* **32**, 719–723 (2019).
173. Döffinger, R. *et al.* X-linked anhidrotic ectodermal dysplasia with immunodeficiency is caused by impaired NF-kappaB signaling. *Nat. Genet.* **27**, 277–285 (2001).
174. Merath, K. M., Chang, B., Dubielzig, R., Jeannotte, R. & Sidjanin, D. J. A spontaneous mutation in Srebf2 leads to cataracts and persistent skin wounds in the lens opacity 13 (lop13) mouse. *Mamm. Genome* **22**, 661–673 (2011).
175. Krashes, M. J., Lowell, B. B. & Garfield, A. S. Melanocortin-4 receptor-regulated energy homeostasis. *Nat. Neurosci.* **19**, 206–219 (2016).
176. Houghton, B. L., Meendering, J. R., Wong, B. J. & Minson, C. T. Nitric oxide and noradrenaline contribute to the temperature threshold of the axon reflex response to gradual local heating in human skin. *J. Physiol.* **572**, 811–820 (2006).
177. Pavlov, V. A. & Tracey, K. J. Neural regulation of immunity: molecular mechanisms and clinical translation. *Nat. Neurosci.* **20**, 156–166 (2017).
178. Palinkas, L. A. Mental and cognitive performance in the cold. *Int. J. Circumpolar Health* **60**, 430–439 (2001).
179. Mäkinen, T. M. *et al.* Effect of repeated exposures to cold on cognitive performance in humans.

Physiol. Behav. **87**, 166–176 (2006).
180. Taylor, L., Watkins, S. L., Marshall, H., Dascombe, B. J. & Foster, J. The Impact of Different Environmental Conditions on Cognitive Function: A Focused Review. *Front. Physiol.* **6**, 372 (2015).
181. Montgomery, J. C. & Macdonald, J. A. Effects of temperature on nervous system: implications for behavioral performance. *Am. J. Physiol.* **259**, R191–6 (1990).
182. Städele, C., Heigele, S. & Stein, W. Neuromodulation to the Rescue: Compensation of Temperature-Induced Breakdown of Rhythmic Motor Patterns via Extrinsic Neuromodulatory Input. *PLoS Biol.* **13**, e1002265 (2015).
183. de la Peña, E. *et al.* The Influence of Cold Temperature on Cellular Excitability of Hippocampal Networks. *PLoS One* **7**, e52475 (2012).
184. Van Hook, M. J. Temperature effects on synaptic transmission and neuronal function in the visual thalamus. *PLoS One* **15**, e0232451 (2020).
185. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
186. Eden, E., Lipson, D., Yogev, S. & Yakhini, Z. Discovering motifs in ranked lists of DNA sequences. *PLoS Comput. Biol.* **3**, e39 (2007).
187. Librado, P. *et al.* Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E6889–97 (2015).
188. Liu, S. *et al.* Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. *Cell* **157**, 785–794 (2014).
189. Yim, H.-S. *et al.* Minke whale genome and aquatic adaptation in cetaceans. *Nat. Genet.* **46**, 88–92 (2014).
190. Kumar, V., Kutschera, V. E., Nilsson, M. A. & Janke, A. Genetic signatures of adaptation revealed from transcriptome sequencing of Arctic and red foxes. *BMC Genomics* **16**, 585 (2015).
191. Lynch, V. J. *et al.* Elephantid Genomes Reveal the Molecular Bases of Woolly Mammoth Adaptations to the Arctic. *Cell Rep.* **12**, 217–228 (2015).
192. Yi, X. *et al.* Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010).
193. Voight, B. F., Kudaravalli, S., Wen, X. & Pritchard, J. K. A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72 (2006).
194. Sabeti, P. C. *et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).
195. McColl, H. *et al.* The prehistoric peopling of Southeast Asia. *Science* **361**, 88–92 (2018).
196. Yang, M. A. *et al.* 40,000-Year-Old Individual from Asia Provides Insight into Early Population Structure in Eurasia. *Curr. Biol.* **27**, 3202–3208.e9 (2017).
197. Racimo, F., Marnetto, D. & Huerta-Sánchez, E. Signatures of Archaic Adaptive Introgression in Present-Day Human Populations. *Mol. Biol. Evol.* **34**, 296–317 (2017).
198. Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
199. Prüfer, K. *et al.* A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017).
200. Mafessoni, F. *et al.* A high-coverage Neandertal genome from Chagyrskaya Cave. *Cold Spring Harbor Laboratory* 2020.03.12.988956 (2020) doi:10.1101/2020.03.12.988956.
201. Pickrell, J. K. *et al.* Signals of recent positive selection in a worldwide sample of human populations. *Genome Res.* **19**, 826–837 (2009).
202. Refoyo-Martínez, A. *et al.* Identifying loci under positive selection in complex population histories. *Genome Res.* **29**, 1506–1520 (2019).

203. Schrider, D. R. & Kern, A. D. S/HIC: Robust Identification of Soft and Hard Sweeps Using Machine Learning. *PLoS Genet.* **12**, e1005928 (2016).