Supplementary Materials for

Sex-specific phenotypic effects and evolutionary history of an ancient polymorphic deletion of the human growth hormone receptor

Marie Saitou, Skyler Resendez, Apoorva J. Pradhan, Fuguo Wu, Natasha C. Lie, Nancy J. Hall, Qihui Zhu, Laura Reinholdt, Yoko Satta, Leo Speidel, Shigeki Nakagome, Neil A. Hanchard, Gary Churchill, Charles Lee, G. Ekin Atilla-Gokcumen, Xiuqian Mu*, Omer Gokcumen*

*Corresponding author. Email: gokcumen@gmail.com (O.G.); xmu@buffalo.edu (X.M.)

Published 24 September 2021, Sci. Adv. 7, eabi4476 (2021)
DOI: 10.1126/sciadv.abi4476

The PDF file includes:

Supplementary Methods
Figs. S1 to S13
Tables S6 to S8, and S10
References

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S5, S9 and S11
SUPPLEMENTARY MATERIALS

Supplementary Figures: This document includes the supplementary figures cited in the main text.

Supplementary tables:
Table S1. SNPs in Linkage disequilibrium with GHRd3 and PheWAS results
Table S2. Sample information and weight data
Table S3. Transcriptome data
Table S4. GO analysis results
Table S5. Lipidomics results
Table S6. Results for 2DSFS analysis (Embedded in supplementary information)
Table S7. Parameters for Bayesian simulations (Embedded in supplementary information)
Table S8. Results for Bayesian inference (Embedded in supplementary information)
Table S9. RELATE and CLUES results
Table S10. GHRd3 Genotypes from Malawi malnutrition cohort (Embedded in supplementary information)
Table S11. Differences detected between the reference genome and the Ghrd3 mice

Supplementary information: This document includes detailed methodologies and supporting analysis.
SUPPLEMENTARY MATERIALS

Supplementary figures

Figure S1. Conservation of GHRd3. A. GHRd3 in modern and ancient hominin genomes. These browser snapshots show the genome assembly (Hg19) of a human with the ancestral, homozygous non-deleted genotype and another with a homozygous deleted genotype that shows no reads mapping to the deletion region (top two rows). Similarly, sequences from 3 Neanderthal and 1 Denisovan assemblies were mapped to this region and show a clear signature of the deletion with breakpoints indistinguishable from the deletion observed in modern humans. B. The coding sequence alignment of exon 3 among mammals, which indicates near-complete preservation of the amino-acid sequence in primates and clear conservation across mammals. C. A conservation score (phastCons46.mammals in UCSC Genome Browser) comparison of GHR exon 3 sequences with 20,000 randomly chosen sites from chromosome 5.

Figure S2. Expression of GHR and its isoforms in humans. The heatmap shows data from the GTEx database (v8) on the read counts from individual exons of GHR for multiple human
tissues. The known transcripts and corresponding exons are shown below the heatmap. It is clear from this figure that Exon 3 (designated Exon 11 in GTEx nomenclature - highlighted by the red box) is expressed in all tissues. Bottom right violin plots showed the significant, additive decrease in the splicing involving exon 3 in individuals carrying rs6873545, which tags GHRd3. The lower values here indicate the relative increase in the expression of the transcript without the exon 3 due to genomic deletion of this sequence in these individuals.
Figure S3. Linkage disequilibrium between SNPs near GHRd3 and the deletion. This figure shows $R^2$ values calculated for the GHR deletion and its neighboring SNPs (black dots, left scale), along with recombination rates for the locus (the pink line, right scale). Recombination rates were retrieved from the 1000 genomes selection browser (http://hsb.upf.edu/). Based on the linkage disequilibrium, we selected Hg19: chr5: 42624748-42628325 as the tag region.
Figure S4. *GHRd3* tag SNP frequency map. Global frequencies of the *GHRd3* tag SNP from the Human Genome Diversity Project (https://www.hagsc.org/hgdp/); The C allele tags the derived deletion, despite being the ancestral allele.
Figure S5. *GHR* exon 3 phylogenetic tree. This phylogenetic tree was constructed using randomly selected haplotypes that harbor the deleted and non-deleted *GHR* exon 3 alleles. It is clear from this phylogeny that haplotypes harboring the deleted allele are more diverse and coalesce earlier than those that harbor the non-deleted allele. It is also noteworthy that both the Altai Neanderthal and Denisova genomes cluster with the haplotypes harboring the deleted allele.
Figure S6. Tajima’s D values for all haplotypes. A. Tajima’s D values between the GHRd3 tag region (Hg19: chr5: 42624748-42628325, blue) and 500 randomly selected regions across the genome (white) for three populations. B. Tajima’s D values are calculated for the deleted and non-deletion alleles in three 1000 Genomes meta-populations. The * and ** indicate 0.05 and 0.01 significance, respectively. C. XP-EHH values calculated for the GHRd3 upstream region.; XP-EHH values are computed for the GHRd3 tag SNPs (r^2 > 0.8) and compared to distributions calculated for 1000 similarly sized regions (shown in grey).
Figure S7. The distribution of GHRd3 in ancient genomes. We used Allen Ancient Genome Resource (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data, accessed May 18, 2021) to consolidate genotype information from 1,888 ancient genomes for the SNP rs4590183 which tags the GHRd3 deletion (R²=~0.9). Other tag SNPs were not genotyped in this database. The top figure shows the allele frequency distribution of GHRd3 across time in 5,000 years bins. The size of each dot represents the number of samples in that bin (ranging from 1 to 375). The archaic human genomes (Neanderthal and Denisovan) in each bin were indicated by blue-green text. Other data is from anatomically modern humans. The bottom map shows the geographical distribution of samples that are older than 5,000 years, with a median age = ~8-10 thousand years. Two important shortcomings/biases in this dataset should be noted. First, there are only a few anatomically modern samples that are older than 30,000 year. Thus, this analysis is not informative in understanding the allele frequency of GHRd3 among anatomically modern humans prior to 30 KYA. However, the results are consistent with the dates of frequency change estimated by the haplotype-level analysis (e.g., Fig. 1D) in the sense that the allele frequency levels are below 50% among modern human populations younger than 30 thousand years. Please note that most of the genotypes here are “pseudo-haploid”, thus status of heterozygotes were not determined.
Figure S8. Approximate Bayesian computation testing different models. Decays of extended haplotype homozygosity (EHH) (left) and site frequency spectrum (SFS) (middle for the derived allele and right for the ancestral allele) in YRI (top) and CHB (bottom) populations. The EHH shows the probability that two randomly chosen haplotypes are homozygous at all SNPs within a given distance from a focal SNP site, which in this case is a tag variant for GHRd3 (rs6873545). This measure is depicted as a value between 0 and 1. To fully capture the haplotype structure in this region, we recorded the physical positions of the variable sites on each side of the focal SNP where the EHH value decreased from 0.9 to 0.1, in steps of 0.1, for derived (red dots) and ancestral alleles (light blue dots) separately. The SFS covers the entire range of allele frequencies from singletons to mutations shared across all chromosomes, providing the local reduction in nucleotide diversity and the distortion of the SFS in a population. There were 113 chromosomes carrying the derived alleles and 103 carrying the ancestral alleles in the YRI samples; For the CHB samples, these numbers were 170 and 36, respectively. The bin size was set as "2". Therefore, the total number of bins for those chromosomes was 57 and 52 for YRI and 85 and 18 for CHB. We used all of the data shown in these plots as a set of summary statistics for the analysis of approximate Bayesian computation (ABC) (see Supplementary information).
Figure S9. Estimation of allele frequency trajectory of GHRd3 using RELATE and CLUES. We used RELATE (22) to estimate the allele frequency of GHRd3 (using tag SNP rs6873545) across time in A. 1000 Genomes and B. Simon Genome Diversity Project populations assuming neutrality. In parallel, we used CLUES (23) to infer the trajectory of GHRd3, this time allowing selection in C. 1000 Genomes and Simon Genome Diversity Project populations (Fig. 1D).
Figure S10. The effect of heterozygote GHRd3 on human traits. We have collected data when available from locus-specific studies (7, 9, 11, 65) that provided genotype-level data. Below we show the distribution/variation of measures of different human traits for each genotype. We are showing only those traits for which significant associations with GHRd3 are reported. The confidence intervals used here are also taken directly from the studies as reported. Most provide a 95% confidence interval, except for Padidela et al., who provides standard deviation.
Figure S11. Phewas analysis on the *GHRd3* tag SNP. The UK biobank PheWas (http://geneatlas.roslin.ed.ac.uk/phewas/) results using the tag SNP (rs4073476).
Figure S12. A gel exemplifying our genotyping approach. A two percent agarose gel showing polymerase chain reaction products. Two sets of primers were used. A sample that showed amplification with the non-deleted primers, but not the deleted primers would be deemed wt/wt; A sample that showed amplification with the deleted primers, but not the non-deleted primers would be deemed d3/d3; A sample that showed amplification with both sets of primers would be deemed wt/d3. Water was used as a negative control and a known heterozygote (904) was used as a positive control.

| Conclusions                                      |     |
|--------------------------------------------------|-----|
| Water - No amplification (Negative Control)      | 210 - wt/d3 |
| 904 - wt/d3 (Positive Control)                   | 221 - d3/d3 |
| 205 - wt/d3                                      | 160 - wt/d3 |
| 206 - d3/d3                                      | 163 - d3/d3 |
| 207 - wt/d3                                      | 201 - wt/wt |
| 208 - d3/d3                                      | 202 - wt/wt |
| 209 - wt/d3                                      | 204 - wt/wt |
**Figure S13.** RNAseq exonic read-depth from a wt/wt and d3/d3 mouse for exons 3 and 4. Each horizontal grey line represents a single read. The d3/d3 mouse expresses **GHR** exon 4 in similar levels as compared to the wt/wt mouse, but does not express exon 3.
SUPPLEMENTARY METHODS

1. POPULATION GENETICS ANALYSES

1.1. Haplotype network and read-depth analysis
We used the program vcftools (0.1.16) (66) to calculate the $R^2$ values between GHRd3 (esv3604875) and flanking SNPs to set a target region (Hg19: chr5: 42624748-42628325). A vcf file of the target region obtained from the 1000 Genome Project phase 3 dataset (1), the hg19 reference genome, the chimpanzee reference genome (67), the Altai Neanderthal genome, and the Denisovan genome (68, 69) were all used by the program VCTtoTree (V3.0.0) (70) to draw the haplotype networks. PopART (Version 1.7) (17) by Minimum Spanning Network method (71) was used for the visualization of the network (Fig. 1B). The program rworldmap (72) was used to visualize the global allele frequency data from the 1000 Genome Project phase 3 dataset (1).

1.2. Conservation Analysis
We obtained the GHR coding sequence alignment from Fig. S1B through the UCSC genome browser by utilizing its “Other Species Alignments” function (MAF table: multiz100way) (73). This alignment was then viewed and exon 3 was highlighted using the program Molecular Evolutionary Genetics Analysis (MEGA v 10.0.5) (74). The software program bedtools (v2.27.1) (75) was used to obtain 20,000 random sites from chromosome 5, where the GHR gene is located, for the conservation analysis. The phastCons46way.placental datasets (76) from the UCSC Genome Browser were used to compare the conservation scores of GHR exon 3 and the randomly selected regions (Fig. S1C).

1.3. Empirical comparisons of Tajima’s D and XP-EHH
To obtain random single nucleotide variants, we used bedtools (v2.27.1) (75) to construct random chromosomal coordinates on chromosome 5. We used 500 random regions for the GHRd3 locus comparison. Tajima’s D (62) values for merged haplotypes (Fig S6A and Fig S6B) and XP-EHH (Fig. S6C) (19) values were downloaded from the 1000 Genomes selection browser (77) and used for the bins containing the target region (Hg19: chr5: 42624748-42628325) and control region. To increase sample numbers for our Tajima’s D calculations we extended the analysis beyond the CEU, YRI, and CHB populations. The European category was expanded to include the following populations: Utah residents with Northern and Western European ancestry (CEU), Toscani in Italy (TSI), Finnish in Finland (FIN), British in England and Scotland (GBR), and Iberian populations in Spain (IBS); The African category was expanded to include the following populations: Gambian in Western Division in the Gambia (GWD), Mende in Sierra Leone (MSL), Esan in Nigeria (ESN), Yoruba in Ibadan, Nigeria (YRI), and Luhya in Webuye, Kenya (LWK); The east Asian category was expanded to include the following populations: Han Chinese in Beijing, China (CHB), Japanese in Tokyo, Japan (JPT), Southern Han Chinese, China (CHS), Chinese Dai in Xishuangbanna, China (CDX), and Kinh in Ho Chi Minh City, Vietnam (KHV). The visualization was constructed through ggplot2 (78).

1.4. Calculation of Fc
For the analysis of LD patterns in different site frequency spectra, we followed the approach outlined by Fujito et al. (20). We used the two-dimensional site frequency spectrum (20, 79) to detect the signature of selection on GHRd3. This method can eliminate the effect of recombinations, which affect haplotype structures. We used the 1000 Genome meta-populations (Africa, East Asia, and Europe) for this analysis. By combining the site frequency spectrum method and coalescence simulations, we detected the signature of selection on the GHR non-deletion allele ($p < 0.05$) with the estimated date of selection onset being 27.5ky in Asian populations. The results of 2DSFSanalysis are summarized in Table S6.
Table S6 Measures for the 2DSFS method to detect selective sweep. The basic concept of this method is to measure the IAV (Intra-allelic variability) of a particular haplotype (allelic group) containing the derived allele (der.) or ancestor allele (anc.) of a target site. This group is called D-group or A-group and variability of each group is represented through several measurements, mainly four, FC, LC0, GC0 (GC0*), C. FC is the number of derived SNPs other than the target within the D or A group, relative to the total. LC0 is the number of derived SNPs other than the target within the D or A group, but excluding those at recombinant sites, relative to the total. GC0 is the average number of derived SNPs including the target site per non-recombining SNP site within the D or A group. GC0* is the GC0 excluding singleton in its calculation. C is the total number of derived alleles including the target site at non-recombining SNP sites within the D or A group and the number divided by m, the number of samples with the target allele, is the average height of the D or A group. This number corresponds to TMRCA of the D or A group. Among the four measurements, the statistical significance of the first three are attained to be compared with values under a neutral case which takes the demographic model into consideration. The probability that simulated values are less than each observed value is given in the right column.

1.5. Approximation of the mode and tempo of natural selection
To further clarify the population dynamics of GHRd3 and the non-deleted alleles in YRI and CHB, we applied an approximate Bayesian computation framework developed in a previous study (21) to infer the mode and tempo of natural selection. We first simulated data from three different models using mssel, a modified version of ms (80): i) selection on a new mutation, ii) selection on a standing variant, and iii) neutrality. There are four parameters defining these scenarios: allele frequency at present (f₀), the onset of natural selection (t), selection coefficient (s), and allele frequency at t (fᵣ). The details of our simulation conditions are shown in Table S7 below. We then calculated the goodness-of-fit of each model to the observed summaries (Fig. S8) as an approximate marginal likelihood (aML) using a kernel density estimate (81). Once we identified the best fitting model in terms of approximate Bayes factors (aBFs), we estimated the parameters under the model by kernel ABC (82).

| Population | Demography                                                                 | Prior conditions |
|------------|----------------------------------------------------------------------------|------------------|
| YRI        | Constant (N = 12,000)                                                     |                  |
|            |                                                                             |                  |
| CHB        | Population bottleneck and expansion ~1600 generations ago: N = 12,000      | log₁₀U(-3, -0.5) |
|            | 1100~1600: N = 1,200                                                      | Log-normal(1600, 1600²) |
|            | 500~1100: N = 12,000                                                      |                  |
|            | 500~Present: N = 120,000                                                  | U(0.0, 0.2)      |

Table S7. Details for conditions used for the mssel simulations.

In the YRI population, the standing variation model showed the highest aML. However, this aML was almost comparable to the one under the neutral model (see the table below). The aML calculated for the CHB population was significantly higher in the standing variation model than in the other models, suggesting that the non-deleted allele became advantageous during the expansion into East Eurasia. To test this hypothesis, we further estimated the onset of natural selection under the standing variation model (i.e. the best fitting model chosen in CHB) (see Fig. 1C). The estimates (Table S8) showed that selection on the non-deleted allele occurred around 30,000 years ago (95% confidence interval 5,103 to 46,710) when this allele was present at 11%. The frequency increased over time with a selection coefficient of 1.13%.

| Population | Models                                |
|------------|---------------------------------------|
|            | Neutral | Selection on new mutation | Selection on standing variant |
| YRI        | Log(aML¹) | -214.687 | -249.968 | -212.995 |

Table S8. Estimates for the onset of selection in CHB.
Table S8. The results for ABC simulations - 1Approximated marginal likelihood under a given model. 2Approximated Bayes factor shows how much one model is better fit to the observed data than the other.

|            | Log\(_{10}\)(aBF\(^2\)) between a given model and the model with the highest aML | 0.735 | 16.057 | - |
|------------|-----------------------------------------------------------------------------------|-------|--------|---|
| CHB        | Log(aML)                                                                          | -206.437 | -224.875 | -194.349 |
|            | Log\(_{10}\)(aBF) between a given model and the model with the highest aML      | 5.250 | 13.257 | - |

1.6. RELATE and CLUES

We used existing genome-wide genealogies for the 1000 Genomes Project data set (1000G) (1) and Simon’s Genome Diversity Project (SGDP) (2) (available from https://www.dropbox.com/sh/jmxzp2om4nh3p7u/AABigOmwr3laYEQaabpGNHeta?dl=0 and https://www.dropbox.com/sh/2gjyxe3kzh932o/AAAQcipCHnySqEB873t9EqjNa?dl=0), inferred using Relate (22), with branch lengths and coalescence rates fitted jointly using an MC-EM approach (83). We then inferred the allele frequency trajectories of SNP rs6873545 using two approaches (Table S9):

First, we inferred allele frequencies through time using the CLUES approach (23). For this approach, we first subsetted the local genealogical tree to the group of interest (one of YRI, CHB, CEU in 1000G and one of Africa, West Eurasia, East Asia in SGDP). We then resampled branch lengths 200 times using the genome-wide coalescence rates for that group and a mutation rate of \(1.25 \times 10^{-8}\) per base per generation. These were fed into CLUES. Because the allele of interest was expected to be old and selection was therefore expected to be acting on standing variation <40k years ago, we chose to set the selection coefficient to 0 for times >40k years before present, but non-zero more recently. We set the parameter tCutoff to 5000, i.e. only modeled allele frequencies in the last 5000 generations (140,000 years, assuming 28 years per generation).

The CLUES approach currently fits a separate allele frequency trajectory to each population and therefore does not guarantee consistency across populations. We therefore also opted to infer allele frequencies using a second approach, which is expected to yield more consistent estimates across populations. This approach is only unbiased under neutrality and overestimates allele frequency further back in time under positive selection of the derived allele because the derived allele is overrepresented.

In this approach, we first extracted the marginal genealogical tree describing SNP rs6873545 as before. We then repeatedly resampled branch lengths for this tree (100 times), given a mutation rate of \(1.25 \times 10^{-8}\) per base per generation and the joint genome-wide coalescence rates for all groups in each data set. For each tree with resampled branch lengths, we then subsetted the tree
to a population of interest (one of YRI, CHB, CEU in 1000G and one of Africa, West Eurasia, East Asia in SGDP) and computed, at predefined time points, the proportion of lineages carrying the derived allele, relative to the total number of lineages remaining at that time. We restricted our inference to time points where >95% of sampled trees have at least 5 lineages left, which excluded times >100,000 years before present for most groups.

1.7. Phenotype-wide association studies

One of the single nucleotide variants (rs4073476) upstream of the exon 3 deletion almost perfectly tags GHRd3 in European populations ($R^2 >0.95$). This variant was entered into the GeneATLAS PheWAS database (http://geneatlas.roslin.ed.ac.uk/phewas/) for an association analysis (27). This database identifies associations between SNPs and 742 phenotypic traits using the UK Biobank cohort (https://www.ukbiobank.ac.uk/). In parallel, we examined the GWAS database (available through https://atlas.ctglab.nl/PheWAS) among 152 phenotypes., we found that the derived GHRd3 haplotype is strongly associated with bone mineral density. The next most significant association was height with a nominal p-value < 10$^{-5}$ (Table S1). To further the association analysis, another database was utilized (http://www.nealelab.is/uk-biobank/). This database, created by the Neale lab, provides GWAS summary statistics for 4,203 phenotypes from the UK Biobank cohort. Summary statistics were downloaded for males, females, and both datasets combined for the top ten associations identified through the previous GeneATLAS analysis. These summary statistics allowed for sex-specific association analyses.

1.8. Analysis of GHRd3 in a cohort of severe acute malnutrition

Using available genotyping platforms from previous work (9), we analyzed Malawian children with edematous severe acute childhood malnutrition (edematous SAM or ESAM) and non-edematous SAM (NESAM) (32). The former is a more severe condition. Specifically, we focused on 2 SNPs flanking the GHR exon 3 deletion, rs4590183, and rs6873545, which are in strong ($R^2>0.9$) LD with the GHRd3 in 1000 Genomes data. However, we found that the LD between these two SNPs was broken ($R^2=0.55$) in this east-central African cohort. Thus, instead of using these SNPs as a proxy for the deletion, we directly genotyped this cohort for the deletion using PCR primers and conditions described earlier (9). Briefly, we used the following primers to genotype both the deleted and non-deleted haplotypes.

G1: 5’-TGTGCTGGTCTGTTGGTCTG;
G2: 5’-AGTCGTTCCTGGGACAGAGA;
G3: 5’-CCTGGATTAACACTTTGCAGACTC

G1-G3 would not amplify deleted haplotypes but will give a band for non-deleted haplotypes. G1-G2 would not amplify non-deleted haplotypes (too large to amplify) but will give a band for the deleted haplotypes. Overall, we were able to successfully genotype the homozygous and heterozygous deletion with high confidence in 176 samples, consisting of 84 ESAM children (48 males, 36 females) and 92 NESAM children (52 males, 40 females) using these two PCR
reactions (Table S10). We then tested our expectation that GHRd3 is enriched in the less severe form of malnutrition (NESAM). We found strong evidence of an association between the deletion and the ESAM phenotype generally (all individuals, freq in cases vs freq in controls; p<0.01; test), driven by evidence of association in the subset of males (p<0.05).

| Genotype                | Phenotype | Sex | Number | Frequency |
|-------------------------|-----------|-----|--------|-----------|
| Non-Deleted             | ESAM      | F   | 7      | 0.1944    |
| Heterozygous            | ESAM      | F   | 25     | 0.6944    |
| Homozygous Ghrd3        | ESAM      | F   | 4      | 0.1111    |
| Non-Deleted             | NESAM     | F   | 6      | 0.15      |
| Heterozygous            | NESAM     | F   | 26     | 0.65      |
| Homozygous Ghrd3        | NESAM     | F   | 8      | 0.2       |
| Non-Deleted             | ESAM      | M   | 15     | 0.3125    |
| Heterozygous            | ESAM      | M   | 27     | 0.5625    |
| Homozygous Ghrd3        | ESAM      | M   | 6      | 0.125     |
| Non-Deleted             | NESAM     | M   | 12     | 0.2308    |
| Heterozygous            | NESAM     | M   | 30     | 0.5769    |
| Homozygous Ghrd3        | NESAM     | M   | 10     | 0.1923    |

Table S10. The genotypes for NESAM and ESAM cohorts in Malawi malnutrition cohort. The number of observed genotypes for each phenotype and sex are provided.

We looked for any deviations from Hardy-Weinberg equilibrium in our NESAM cohort. We did notice a slight but not significant (p>0.05) elevation in the observed heterozygotes as compared to expectation in both males and females. This increase in heterozygotes, albeit interesting, is likely due to small sample sizes or small genotyping errors. Regardless, we think that the HWE analysis does not change our observation that the NESAM cohort has a significantly increased allele frequency of GHRd3 as compared to the ESAM cohort given that we have used the same PCR assay to genotype both cohorts. Even if there is any bias in our genotyping (e.g., over-calling heterozygotes), it would affect both NESAM and ESAM cohorts in a similar manner. Clearly, larger cohorts are needed to definitively establish the association between GHRd3 and test for the potential effect of strong selection due to the outcome of severe malnutrition.

2. MOUSE MODEL

2.1. Design of the GHRd3 mouse

To model GHRd3, we designed sgRNAs that flank both sides (5’ and 3’) of the mouse Ghr exon orthologous to the human GHR exon 3 with the help of the Gene Targeting and Transgenic Resource at Roswell Park Comprehensive Cancer Center. Specifically, 18 5’ and 11 3’ primers were designed. These primers were screened for common single nucleotide polymorphisms in dbSNP (https://www.ncbi.nlm.nih.gov/snp/) and tested for off target cuts using a standard T7
endonuclease 1 (T7E1) mismatch detection assay (84). Based on the results of this assay, the following 4 sgRNA sequences were used for the downstream analysis:

mGHR5'.g5 - AATACAATTGGCTAATACCGNGG
mGHR5'.g6 - TTGGCTAATACCGAGGTGAGNGG
mGHR3'.g10 - TAAGATTTTTAGTGATGTAANGG
mGHR3'.g11 - AAACATGACCATTGAATTAANGG

These guides were further validated using Bioanalyzer and Qubit for adequate concentrations (~300ng/µL). C57BL/6N fertilized embryos were harvested at day 0.5 dpc. Fertilized embryos were then injected with guide RNAs and Cas9 into the pronucleus to perform genome editing.

In vitro fertilization (IVF) and calorie restriction: To quickly generate a cohort of mice for calorie restriction, d3/d3 males were sent to The Jackson Laboratory for sperm cryopreservation followed by in vitro fertilization of C57BL/6NJ oocytes and embryo transfer. This resulted in 62 d3 heterozygous mice from which a cohort of 40 mice were generated for the calorie restriction study outlined in Fig. 3A. The cohorts were kept with their mothers which were fed standard (Ad libitum) diets until 1 weaning (~1 month). After weaning, half of the mice were fed 40% calorie restricted chow for another month while the other half were fed Ad Libitum diets. All mice were sacrificed at ~2 months.

2.2. Genotyping

The Vindija and Chagyrskaya Neanderthal BAM files were analyzed using the Integrative Genomics Viewer software (v2.3.75) (85) to determine if these genomes contained GHRd3 (https://www.eva.mpg.de/genetics/genome-projects/chagyrskaya-neandertal/home.html?Fsize=0Svea-Developmental). These results were then added to previous findings regarding the Altai Neanderthal and Denisovan genomes (3) (Fig. S1A).

The mice were screened for GHRd3 using the Kapa Biosystems KAPA2G Fast HS Genotyping kit alongside two sets of primers (Fig. S12), following the producer’s protocol. One set of primers leads to amplification if the deletion is present; the other set of primers leads to amplification if the non-deleted, ancestral haplotype is present. For heterozygous individuals, amplification occurs with both sets of primers. For each experiment, water was used as a negative control and a DNA sample from a known heterozygote was used as a positive control.

Deletion Primers: Forward - SM444CEL F (AGAGTACCCAGTGTATGGCCT). Reverse - SM445CEL R (TGCTGTCTGGCACACATGAT)
Non-Deletion Primers: Forward SM444CEL F (AGAGTACCCAGTGTATGGCCT). Reverse - SM444CEL R (AGTTCTGTGAGCTGGTGTAGC)

In parallel, to validate the successful genome-editing of the mice at the transcriptome level, we conducted a read-depth analysis on BAM files made from the transcriptomes. To create BAM files, FASTQ files were aligned to the GRCm38 (mm10) mouse reference genome using TopHat2 (86) and Bowtie2 (87). We then observed the read depths of transcripts from exon 3 and exon 4 (the latter for comparison) of the mouse Ghr by using the following Samtools command (88):
Based on the read-depth information, the presence or absence of exon 3 expression was readily detectable (Fig. S13).

2.3. Checking for in vitro fertilization effects
To double-check the potential off-target effects from CRISPR-cas9 gene editing, we sequenced 3 male homozygous GHRd3 mice (0063, 0008, and 0077) to at least 50X coverage on an Illumina Novaseq S1 flowcell. We used BWA (89) to align these reads against the GRCm38 reference genome and called single nucleotide and insertion-deletion variants by GATK recommended pipeline (https://gatk.broadinstitute.org/hc/en-us), using the dbSNP dataset as known variable sites. Assuming that the off-target effects that are not proximate to the GHR locus will be eliminated by the multiple backcrossings, we specifically checked 250kb upstream and downstream of the GHR gene (~1Mb region in total including the GHR gene). In this region, we found 12 variants (2 SNPs and 10 indels) that pass the routine QC thresholds (removing sites with QD < 2.0, QUAL < 30.0, SOR > 4.0, FS > 60.0 or ReadPosRankSum < -8.0) and observed in all 6 haplotypes (i.e., homozygous in all three GHRd3 mice that we sequenced) (Table S11). The 2 SNPs are previously reported in dbSNP, thus they are most likely drift effects. Of the 10 indels, 8 of them are expansion or contraction of simple repeat arrays and 2 of them fall in recent retrotransposons. Thus, it is more likely that these indels are recent mutations that drifted to fixation in our colony, alignment artifacts in our calling pipeline, or false-negatives in existing databases rather than off-target effects of the CRISPR-Cas9 intervention. None of these variants hit any known functional sequences. Overall, there is no reason to think that any off-target effects bias the reported results.

2.4. Transcriptomics data and analysis
28 to 34-day-old mice were sacrificed (Table S2) and liver samples were taken and directly put into RNAlater (Thermo Fisher). They were then sent for RNA extraction and sequencing by GENEWIZ. RNA sequencing was performed via Illumina HiSeq, 2x150bp configuration. Quality control of the obtained sequences was performed using FastQC (Andrews S. FASTQC. A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 7/2019), and the results of all the samples were reviewed by MultiQC (90). Adaptor sequences and low quality bases were discarded by Trimomatic (97). Filtered reads were mapped to mouse transcriptome reference (GRCm38) from Ensembl (92) and quantified using Kallisto (63). The transcripts were merged into genes with tximport and Ensembl BioMart (93). Differential expression analyses were performed by DESeq2 (64), which calculates the fold change of each gene using the Wald test and a correction for multiple hypotheses. The gene expressions of the samples are also provided in Table S3. We then defined genes that were upregulated or downregulated using the adjusted (i.e., multiple hypothesis corrected) p-value threshold of 0.05 (Wald test in DeSeq2).

The genes that were upregulated or downregulated in Ghrd3 mice (Table S3) were used against a whole genome background to investigate gene ontology enrichment for all available datasets using ShinyGo (94).

2.5. Lipidomics Analysis
A standard mix was prepared in CHCl₃ containing exogenous lipids d₉ oleic acid (2 μM), d₁₁ arachidonic acid (2 μM), c₁₃ oleic acid (2 μM), d₇₀ distearoylphosphatidylcholine (2 μM), d₃₁ C₁₆ sphingomyelin (2 μM), C₁₇ glucosylceramide (2 μM), C₃₉ triacylglycerol (1 μM), C₆ ceramide (1 μM), C₁₇ ceramide (1μM) and c₁₃ cholesterol (10 μM).
Lipid extraction and analysis were carried out as described in previous studies (95, 96). 28 to 34-day-old mice were sacrificed and serum samples were syphoned. The serum was collected using standard centrifugation methods (97) and later stored at -20 degrees. For the LC-MS analysis, serum samples were thawed on ice, gently mixed by pipetting, then 50 μL samples were transferred to 1 mL cold PBS contained in a 2-dram glass vial on ice. Samples were then resuspended well to mix and a 30 μL aliquot was taken for protein measurement using a Coomassie assay. The relative standard deviation in protein concentration among the samples was 5%. 1 mL of methanol and 2 mL of CHCl₃ were added to the remainder. All solutions were kept cold on ice. Samples were vortexed for 10 seconds, followed by 1 min on ice, three times. Samples were then centrifuged for 30 minutes at 3000 rcf at 4°C. The bottom (CHCl₃) layer was transferred to a 1-dram glass vial and kept on ice. The upper layer was re-extracted with 2 mL of additional CHCl₃ and the two CHCl₃ layers were combined. 3.75 mL of the combined CHCl₃ layers were transferred to a new 1-dram vial and the solvent was removed by rotary evaporation.

The lipid levels of the samples are provided in Table S5.
REFERENCES AND NOTES

1. P. H. Sudmant, T. Rausch, E. J. Gardner, R. E. Handsaker, A. Abyzov, J. Huddleston, Y. Zhang, K. Ye, G. Jun, M. Hsi-Yang Fritz, M. K. Konkel, A. Malhotra, A. M. Stütz, X. Shi, F. P. Casale, J. Chen, F. Hormozdiari, G. Dayama, K. Chen, M. Malig, M. J. P. Chaisson, K. Walter, S. Meiers, S. Kashin, E. Garrison, A. Auton, H. Y. K. Lam, X. J. Mu, C. Alkan, D. Antaki, T. Bae, E. Cerveira, P. Chines, Z. Chong, L. Clarke, E. Dal, L. Ding, S. Emery, X. Fan, M. Gujral, F. Kahveci, J. M. Kidd, Y. Kong, E.-W. Lameijer, S. McCarthy, P. Flicek, R. A. Gibbs, G. Marth, C. E. Mason, A. Menelaou, D. M. Muzny, B. J. Nelson, A. Noor, N. F. Parrish, M. Pendleton, A. Quitadamo, B. Raeder, E. E. Schadt, M. Romanovitch, A. Schlattl, R. Sebra, A. A. Shabalina, A. Untergasser, J. A. Walker, M. Wang, F. Yu, C. Zhang, J. Zhang, X. Zheng-Bradley, W. Zhou, T. Zichner, J. Sebat, M. A. Batzer, S. A. McCarroll; 1000 Genomes Project Consortium, R. E. Mills, M. B. Gerstein, A. Bashir, O. Stegle, S. E. Devine, C. Lee, E. E. Eichler, J. O. Korbel, An integrated map of structural variation in 2,504 human genomes. Nature 526, 75–81 (2015).

2. S. Mallick, H. Li, M. Lipson, I. Mathieson, M. Gymrek, F. Racimo, M. Zhao, N. Chennagiri, S. Nordenfelt, A. Tandon, P. Skoglund, I. Lazaridis, S. Sankararaman, Q. Fu, N. Rohland, G. Renaud, Y. Erlich, T. Willems, C. Gallo, J. P. Spence, Y. S. Song, G. Poletti, F. Balloux, G. van Driem, P. de Knijff, I. G. Romero, A. R. Jha, D. M. Behar, C. M. Bravi, C. Capelli, T. Hervig, A. Moreno-Estrada, O. L. Posukh, E. Balanovska, O. Balanovsky, S. Karachanak-Yankova, H. Sahakyan, D. Toncheva, L. Yepiskoposyan, C. Tyler-Smith, Y. Xue, M. S. Abdullah, A. Ruiz-Linares, C. M. Beall, A. Di Rienzo, C. Jeong, E. B. Starikovskaya, E. Metspalu, J. Parik, R. Vilems, B. M. Henn, U. Hodoglugil, R. Mahley, A. Sajantila, G. Stamatoyannopoulos, J. T. S. Wee, R. Khusainova, E. Khusnutdinova, S. Litvinov, G. Ayodo, D. Comas, M. F. Hammer, T. Kivisild, W. Klitz, C. A. Winkler, D. Labuda, M. Bamshad, L. B. Jorde, S. A. Tishkoff, W. S. Watkins, M. Metspalu, S. Dryomov, R. Sukernik, L. Singh, K. Thangaraj, S. Pääbo, J. Kelso, N. Patterson, D. Reich, The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature 538, 201–206 (2016).

3. Y.-L. Lin, P. Pavlidis, E. Karakoc, J. Ajay, O. Gokcumen, The evolution and functional impact of human deletion variants shared with archaic hominin genomes. Mol. Biol. Evol. 32, 1008–1019 (2015).
4. A. J. Brooks, M. J. Waters, The growth hormone receptor: Mechanism of activation and clinical implications. *Nat. Rev. Endocrinol.* **6**, 515–525 (2010).

5. M. Zoledziewska, C. Sidore, C. W. K. Chiang, S. Sanna, A. Mulas, M. Steri, F. Busonero, J. H. Marcus, M. Marongiu, A. Maschio, D. Ortega Del Vecchyo, M. Floris, A. Meloni, A. Delitala, M. P. Concas, F. Murgia, G. Biino, S. Vaccargiu, R. Nagaraja, K. E. Lohmueller, UK10K consortium, N. J. Timpson, N. Soranzo, I. Tachmazidou, G. Dedoussis, E. Zeggini; Understanding Society Scientific Group, S. Uzzau, C. Jones, R. Lyons, A. Angius, G. R. Abecasis, J. Novembre, D. Schlessinger, F. Cucca, Height-reducing variants and selection for short stature in Sardinia. *Nat. Genet.* **47**, 1352–1356 (2015).

6. S. Amselem, P. Duquesnoy, B. Duriez, F. Dastot, M.-L. Sobrier, S. Valleix, M. Goossens, Spectrum of growth hormone receptor mutations and associated haplotypes in Laron syndrome. *Hum. Mol. Genet.* **2**, 355–359 (1993).

7. R. Padidela, S. M. Bryan, S. Abu-Amero, R. E. Hudson-Davies, J. C. Achermann, G. E. Moore, P. C. Hindmarsh, The growth hormone receptor gene deleted for exon three (GHRd3) polymorphism is associated with birth and placental weight. *Clin. Endocrinol.* **76**, 236–240 (2012).

8. K. Sørensen, L. Aksglaede, J. H. Petersen, H. Leffers, A. Juul, The exon 3 deleted growth hormone receptor gene is associated with small birth size and early pubertal onset in healthy boys. *J. Clin. Endocrinol. Metab.* **95**, 2819–2826 (2010).

9. D. Ben-Avraham, D. R. Govindaraju, T. Budagov, D. Fradin, P. Durda, B. Liu, S. Ott, D. Gutman, L. Sharvit, R. Kaplan, P. Bougnères, A. Reiner, A. R. Shuldiner, P. Cohen, N. Barzilai, G. Atzmon, The GH receptor exon 3 deletion is a marker of male-specific exceptional longevity associated with increased GH sensitivity and taller stature. *Sci. Adv.* **3**, e1602025 (2017).

10. K. Sørensen, L. Aksglaede, T. Munch-Andersen, N. J. Aachmann-Andersen, H. Leffers, J. W. Helge, L. Hilsted, A. Juul, Impact of the growth hormone receptor exon 3 deletion gene polymorphism on glucose metabolism, lipids, and insulin-like growth factor-I levels during puberty. *J. Clin. Endocrinol. Metab.* **94**, 2966–2969 (2009).
11. C. Dos Santos, L. Essioux, C. Teinturier, M. Tauber, V. Goffin, P. Bougnères, A common polymorphism of the growth hormone receptor is associated with increased responsiveness to growth hormone. *Nat. Genet.* **36**, 720–724 (2004).

12. A. Dauber, Y. Meng, L. Audi, S. Vedantam, B. Weaver, A. Carrascosa, K. Albertsson-Wikland, M. B. Ranke, A. A. L. Jorge, J. Cara, M. P. Wajnrajch, A. Lindberg, C. Camacho-Hübner, J. N. Hirschhorn, A genome-wide pharmacogenetic study of growth hormone responsiveness. *J. Clin. Endocrinol. Metab.* **105**, 3203–3214 (2020).

13. K. Prüfer, C. de Filippo, S. Grote, F. Mafessoni, P. Korlević, M. Hajdinjak, B. Vernot, L. Skov, P. Hsieh, S. Peyrégne, D. Reher, C. Hopfe, S. Nagel, T. Maricic, Q. Fu, C. Theunert, R. Rogers, P. Skoglund, M. Chintalapati, M. Dannemann, B. J. Nelson, F. M. Key, P. Rudan, Ž. Kučan, I. Gušić, L. V. Golovanova, V. B. Doronichev, N. Patterson, D. Reich, E. E. Eichler, M. Slatkin, M. H. Schierup, A. M. Andrés, J. Kelso, M. Meyer, S. Pääbo, A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017).

14. F. Mafessoni, S. Grote, C. de Filippo, V. Slon, K. A. Kolobova, B. Viola, S. V. Markin, M. Chintalapati, S. Peyrégne, L. Skov, P. Skoglund, A. I. Krivoshapkin, A. P. Derevianko, M. Meyer, J. Kelso, B. Peter, K. Prüfer, S. Pääbo, A high-coverage Neandertal genome from Chagyrskaya Cave. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15132–15136 (2020).

15. J. Pantel, K. Machinis, M.-L. Sobrier, P. Duquesnoy, M. Goossens, S. Amselem, Species-specific alternative splice mimicry at the growth hormone receptor locus revealed by the lineage of retroelements during primate evolution. *J. Biol. Chem.* **275**, 18664–18669 (2000).

16. S. B. Gabriel, S. F. Schaffner, H. Nguyen, J. M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S. N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E. S. Lander, M. J. Daly, D. Altshuler, The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).

17. J. W. Leigh, D. Bryant, POPART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
18. F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).

19. P. C. Sabeti, P. Varilly, B. Fry, J. Lohmueller, E. Hostetter, C. Cotsapas, X. Xie, E. H. Byrne, S. A. McCarroll, R. Gaudet, S. F. Schaffner, E. S. Lander; International HapMap Consortium, Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007).

20. N. T. Fujito, Y. Satta, T. Hayakawa, N. Takahata, A new inference method for detecting an ongoing selective sweep. *Genes Genet. Syst.* **93**, 149–161 (2018).

21. S. Nakagome, R. R. Hudson, A. Di Rienzo, Inferring the model and onset of natural selection under varying population size from the site frequency spectrum and haplotype structure. *Proc. R. Soc. B* **286**, 20182541 (2019).

22. L. Speidel, M. Forest, S. Shi, S. R. Myers, A method for genome-wide genealogy estimation for thousands of samples. *Nat. Genet.* **51**, 1321–1329 (2019).

23. A. J. Stern, P. R. Wilton, R. Nielsen, An approximate full-likelihood method for inferring selection and allele frequency trajectories from DNA sequence data. *PLOS Genet.* **15**, e1008384 (2019).

24. R. B. Jensen, M. Boas, J. E. Nielsen, L. L. Maroun, A. Jørgensen, T. Larsen, K. M. Main, A. Juul, A common deletion in the growth hormone receptor gene (d3-GHR) in the offspring is related to maternal placental GH levels during pregnancy. *Growth Horm. IGF Res.* **55**, 101360 (2020).

25. F. Schreiner, S. Stutte, P. Bartmann, B. Gohlke, J. Woelfle, Association of the growth hormone receptor d3-variant and catch-up growth of preterm infants with birth weight of less than 1500 grams. *J. Clin. Endocrinol. Metab.* **92**, 4489–4493 (2007).

26. R. J. Strawbridge, L. Kärvestedt, C. Li, S. Efendic, C. G. Ostenson, H. F. Gu, K. Brismar, GHR exon 3 polymorphism: Association with type 2 diabetes mellitus and metabolic disorder. *Growth Horm. IGF Res.* **17**, 392–398 (2007).
27. O. Canela-Xandri, K. Rawlik, A. Tenesa, An atlas of genetic associations in UK Biobank. Nat. Genet. 50, 1593–1599 (2018).

28. N. B. Metcalfe, P. Monaghan, Compensation for a bad start: Grow now, pay later? Trends Ecol. Evol. 16, 254–260 (2001).

29. D. S. Millar, M. D. Lewis, M. Horan, V. Newway, D. A. Rees, T. E. Easter, G. Pepe, O. Rickards, M. Norin, M. F. Scanlon, M. Krawczak, D. N. Cooper, Growth hormone (GH) gene variation and the growth hormone receptor (GHR) exon 3 deletion polymorphism in a West-African population. Mol. Cell. Endocrinol. 296, 18–25 (2008).

30. S. C. Antón, R. Potts, L. C. Aiello, Evolution of early Homo: An integrated biological perspective. Science 345, 1236828 (2014).

31. Y. Czorlich, T. Aykanat, J. Erkinaro, P. Orell, C. R. Primmer, Rapid sex-specific evolution of age at maturity is shaped by genetic architecture in Atlantic salmon. Nat. Ecol. Evol. 2, 1800–1807 (2018).

32. K. V. Schulze, S. Swaminathan, S. Howell, A. Jajoo, N. C. Lie, O. Brown, R. Sadat, N. Hall, L. Zhao, K. Marshall, T. May, M. E. Reid, C. Taylor-Bryan, X. Wang, J. W. Belmont, Y. Guan, M. J. Manary, I. Trehan, C. A. McKenzie, N. A. Hanchard, Edematous severe acute malnutrition is characterized by hypomethylation of DNA. Nat. Commun. 10, 5791 (2019).

33. T. E. Forrester, A. V. Badaloo, M. S. Boyne, C. Osmond, D. Thompson, C. Green, C. Taylor-Bryan, A. Barnett, S. Soares-Wynter, M. A. Hanson, A. S. Beedle, P. D. Gluckman, Prenatal factors contribute to the emergence of kwashiorkor or marasmus in severe undernutrition: Evidence for the predictive adaptation model. PLOS ONE 7, e35907 (2012).

34. A. Bergström, C. Stringer, M. Hajdinjak, E. M. L. Scerri, P. Skoglund, Origins of modern human ancestry. Nature 590, 229–237 (2021).

35. T. M. Smith, C. Austin, D. R. Green, R. Joannes-Boyau, S. Bailey, D. Dumitriu, S. Fallon, R. Grün, H. F. James, M.-H. Moncel, I. S. Williams, R. Wood, M. Arora, Wintertime stress, nursing, and lead exposure in Neanderthal children. Sci. Adv. 4, eaau9483 (2018).
36. P. J. Danneman, M. A. Suckow, C. Brayton, *The Laboratory Mouse* (CRC Press, 2012).

37. Y. Fan, R. K. Menon, P. Cohen, D. Hwang, T. Clemens, D. J. DiGirolamo, J. J. Kopchick, D. Le Roith, M. Trucco, M. A. Sperling, Liver-specific deletion of the growth hormone receptor reveals essential role of growth hormone signaling in hepatic lipid metabolism. *J. Biol. Chem.* **284**, 19937–19944 (2009).

38. X. Yang, E. E. Schadt, S. Wang, H. Wang, A. P. Arnold, L. Ingram-Drake, T. A. Drake, A. J. Lusis, Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* **16**, 995–1004 (2006).

39. J. F. Sassin, D. C. Parker, J. W. Mace, R. W. Gotlin, L. C. Johnson, L. G. Rossman, Human growth hormone release: Relation to slow-wave sleep and sleep-walking cycles. *Science* **165**, 513–515 (1969).

40. J. O. Jansson, S. Edén, O. Isaksson, Sexual dimorphism in the control of growth hormone secretion. *Endocr. Rev.* **6**, 128–150 (1985).

41. H. Vakili, Y. Jin, P. A. Cattini, Evidence for a circadian effect on the reduction of human growth hormone gene expression in response to excess caloric intake. *J. Biol. Chem.* **291**, 13823–13833 (2016).

42. A. Chaudhari, R. Gupta, S. Patel, N. Velingkaar, R. Kondratov, Cryptochromes regulate IGF-1 production and signaling through control of JAK2-dependent STAT5B phosphorylation. *Mol. Biol. Cell* **28**, 834–842 (2017).

43. F. J. Steyn, L. Huang, S. T. Ngo, J. W. Leong, H. Y. Tan, T. Y. Xie, A. F. Parlow, J. D. Veldhuis, M. J. Waters, C. Chen, Development of a method for the determination of pulsatile growth hormone secretion in mice. *Endocrinology* **152**, 3165–3171 (2011).

44. M. G. Holloway, G. D. Miles, A. A. Dombkowski, D. J. Waxman, Liver-specific hepatocyte nuclear factor-4α deficiency: Greater impact on gene expression in male than in female mouse liver. *Mol. Endocrinol.* **22**, 1274–1286 (2008).
45. D. Lau-Corona, A. Suvorov, D. J. Waxman, Feminization of male mouse liver by persistent growth hormone stimulation: Activation of sex-biased transcriptional networks and dynamic changes in chromatin states. *Mol. Cell. Biol.* 37, e00301-17 (2017).

46. S. Duran-Ortiz, V. Noboa, J. J. Kopchick, Tissue-specific disruption of the growth hormone receptor (GHR) in mice: An update. *Growth Horm. IGF Res.* 51, 1–5 (2020).

47. C. Ober, D. A. Loisel, Y. Gilad, Sex-specific genetic architecture of human disease. *Nat. Rev. Genet.* 9, 911–922 (2008).

48. A. M. Lichanska, M. J. Waters, How growth hormone controls growth, obesity and sexual dimorphism. *Trends Genet.* 24, 41–47 (2008).

49. E. A. Lucotte, C. Albiñana, R. Laurent, C. Bhérer, Genome of the Netherland Consortium, T. Bataillon, B. Toupance, Detection of sexually antagonistic transmission distortions in trio datasets. bioRxiv 2020.09.11.293191 [Preprint]. 11 September 2020. https://doi.org/10.1101/2020.09.11.293191.

50. J. E. Mank, Population genetics of sexual conflict in the genomic era. *Nat. Rev. Genet.* 18, 721–730 (2017).

51. D. Charlesworth, Balancing selection and its effects on sequences in nearby genome regions. *PLOS Genet.* 2, e64 (2006).

52. A. M. Andrés, M. Y. Dennis, W. W. Kretzschmar, J. L. Cannons, S.-Q. Lee-Lin, B. Hurle; NISC Comparative Sequencing Program, P. L. Schwartzberg, S. H. Williamson, C. D. Bustamante, R. Nielsen, A. G. Clark, E. D. Green, Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *PLOS Genet.* 6, e1001157 (2010).

53. P. Pajic, Y.-L. Lin, D. Xu, O. Gokcumen, The psoriasis-associated deletion of late cornified envelope genes *LCE3B* and *LCE3C* has been maintained under balancing selection since Human Denisovan divergence. *BMC Evol. Biol.* 16, 265 (2016).
54. E. J. Hollox, J. A. L. Armour, Directional and balancing selection in human beta-defensins. *BMC Evol. Biol.* **8**, 113 (2008).

55. C. Sun, D. Huo, C. Southard, B. Nemesure, A. Hennis, M. Cristina Leske, S.-Y. Wu, D. B. Witonsky, O. I. Olopade, A. Di Rienzo, A signature of balancing selection in the region upstream to the human *UGT2B4* gene and implications for breast cancer risk. *Hum. Genet.* **130**, 767–775 (2011).

56. G. Pasvol, D. J. Weatherall, R. J. Wilson, Cellular mechanism for the protective effect of haemoglobin S against *P. falciparum* malaria. *Nature* **274**, 701–703 (1978).

57. S. Xu, R. Nielsen, E. Huerta-Sanchez, The history and evolution of the Denisovan-EPAS1 haplotype in Tibetans. *Proc. Natl. Acad. Sci. U.S.A.* e2020803118 (2021).

58. R. Potts, R. Dommain, J. W. Moerman, A. K. Behrensmeyer, A. L. Deino, S. Riedl, E. J. Beverly, E. T. Brown, D. Deocampo, R. Kinyanjui, R. Lupien, R. B. Owen, N. Rabideaux, J. M. Russell, M. Stockhecke, P. deMenocal, J. T. Faith, Y. Garcin, A. Noren, J. J. Scott, D. Western, J. Bright, J. B. Clark, A. S. Cohen, C. B. Keller, J. King, N. E. Levin, K. Brady Shannon, V. Muiruri, R. W. Renaut, S. M. Rucina, K. Uno, Increased ecological resource variability during a critical transition in hominin evolution. *Sci. Adv.* **6**, eabc8975 (2020).

59. A. Powell, S. Shennan, M. G. Thomas, Late Pleistocene demography and the appearance of modern human behavior. *Science* **324**, 1298–1301 (2009).

60. P. Roberts, B. A. Stewart, Defining the ‘generalist specialist’ niche for Pleistocene *Homo sapiens*. *Nat. Hum. Behav.* **2**, 542–550 (2018).

61. P. U. Clark, A. S. Dyke, J. D. Shakun, A. E. Carlson, J. Clark, B. Wohlfarth, J. X. Mitrovica, S. W. Hostetler, A. M. McCabe, The Last Glacial Maximum. *Science* **325**, 710–714 (2009).

62. F. Tajima, Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* **135**, 599–607 (1993).

63. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
64. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

65. M. G. Wegmann, A. Thankamony, E. Roche, H. Hoey, J. Kirk, G. Shaikh, S.-A. Ivarsson, O. Söder, D. B. Dunger, A. Juul, R. B. Jensen, The exon3-deleted growth hormone receptor gene polymorphism (d3-GHR) is associated with insulin and spontaneous growth in short SGA children (NESGAS). *Growth Horm. IGF Res.* **35**, 45–51 (2017).

66. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

67. The Chimpanzee Sequencing Consortium, Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005).

68. K. Prüfer, F. Racimo, N. Patterson, F. Jay, S. Sankararaman, S. Sawyer, A. Heinze, G. Renaud, P. H. Sudmant, C. de Filippo, H. Li, S. Mallick, M. Dannemann, Q. Fu, M. Kircher, M. Kuhlwilm, M. Lachmann, M. Meyer, M. Ongyerth, M. Siebauer, C. Theunert, A. Tandon, P. Moorjani, J. Pickrell, J. C. Mullikin, S. H. Vohr, R. E. Green, I. Hellmann, P. L. F. Johnson, H. Blanche, H. Cann, J. O. Kitzman, J. Shendure, E. E. Eichler, E. S. Lein, T. E. Bakken, L. V. Golovanova, V. B. Dorendichev, M. V. Shunkov, A. P. Derevianko, B. Viola, M. Slatkin, D. Reich, J. Kelso, S. Pääbo, The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).

69. D. Reich, R. E. Green, M. Kircher, J. Krause, N. Patterson, E. Y. Durand, B. Viola, A. W. Briggs, U. Stenzel, P. L. F. Johnson, T. Maricic, J. M. Good, T. Marques-Bonet, C. Alkan, Q. Fu, S. Mallick, H. Li, M. Meyer, E. E. Eichler, M. Stoneking, M. Richards, S. Talamo, M. V. Shunkov, A. P. Derevianko, J.-J. Hublin, J. Kelso, M. Slatkin, S. Pääbo, Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060 (2010).

70. D. Xu, Y. Jaber, P. Pavlidis, O. Gokcumen, VCFtoTree: A user-friendly tool to construct locus-specific alignments and phylogenies from thousands of anthropologically relevant genome sequences. *BMC Bioinformatics* **18**, 426 (2017).
71. H. J. Bandelt, P. Forster, A. Röhl, Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).

72. A. South, rworldmap: A new R package for mapping global data. *R J.* **3**, 35–43 (2011).

73. R. M. Kuhn, D. Haussler, W. J. Kent, The UCSC genome browser and associated tools. *Brief. Bioinform.* **14**, 144–161 (2013).

74. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

75. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

76. K. S. Pollard, M. J. Hubisz, K. R. Rosenbloom, A. Siepel, Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010).

77. M. Pybus, G. M. Dall’Olio, P. Luisi, M. Uzkudun, A. Carreño-Torres, P. Pavlidis, H. Laayouni, J. Bertranpetit, J. Engelken, 1000 Genomes Selection Browser 1.0: A genome browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Res.* **42**, D903–D909 (2014).

78. H. Wickham, *Ggplot2: Elegant Graphics for Data Analysis* (Springer Publishing Company Incorporated, ed. 2, 2009).

79. Y. Satta, W. Zheng, K. V. Nishiyama, R. L. Iwasaki, T. Hayakawa, N. T. Fujito, N. Takahata, Two-dimensional site frequency spectrum for detecting, classifying and dating incomplete selective sweeps. *Genes Genet. Syst.* **94**, 283-300 (2020).

80. R. R. Hudson, Generating samples under a Wright–Fisher neutral model of genetic variation. *Bioinformatics* **18**, 337–338 (2002).

81. N. Osada, S. Nakagome, S. Mano, Y. Kameoka, I. Takahashi, K. Terao, Finding the factors of reduced genetic diversity on X chromosomes of *Macaca fascicularis*: Male-driven evolution, demography, and natural selection. *Genetics* **195**, 1027–1035 (2013).
82. S. Nakagome, K. Fukumizu, S. Mano, Kernel approximate Bayesian computation in population genetic inferences. *Stat. Appl. Genet. Mol. Biol.* **12**, 667–678 (2013).

83. L. Speidel, L. Cassidy, R. W. Davies, G. Hellenthal, Inferring population histories for ancient genomes using genome-wide genealogies. bioRxiv 2021.02.17.431573 [Preprint]. 17 February 2021. https://doi.org/10.1101/2021.02.17.431573.

84. R. D. Mashal, J. Koontz, J. Sklar, Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases. *Nat. Genet.* **9**, 177–183 (1995).

85. H. Thorvaldsdóttir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).

86. D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S. L. Salzberg, TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).

87. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**, 357–359 (2012).

88. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

89. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

90. P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

91. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
92. D. R. Zerbino, P. Achuthan, W. Akanni, M. R. Amode, D. Barrell, J. Bhai, K. Billis, C. Cummins, A. Gall, C. G. Girón, L. Gil, L. Gordon, L. Haggerty, E. Haskell, T. Hourlier, O. G. Izuogu, S. H. Janacek, T. Juettemann, J. K. To, M. R. Laird, I. Lavidas, Z. Liu, J. E. Loveland, T. Maurel, W. McLaren, B. Moore, J. Mudge, D. N. Murphy, V. Newman, M. Nuhn, D. Ogeh, C. K. Ong, A. Parker, M. Patricio, H. S. Riat, H. Schuilenburg, D. Sheppard, H. Sparrow, K. Taylor, A. Thormann, A. Vullo, B. Walts, A. Zadissa, A. Frankish, S. E. Hunt, M. Kostadima, N. Langridge, F. J. Martin, M. Muffato, E. Perry, M. Ruffier, D. M. Staines, S. J. Trevanion, B. L. Aken, F. Cunningham, A. Yates, P. Flicek, Ensembl 2018. *Nucleic Acids Res.* **46**, D754–D761, 2018.

93. S. Durinck, Y. Moreau, A. Kasprzyk, S. Davis, B. De Moor, A. Brazma, W. Huber, BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440 (2005).

94. S. Ge, D. Jung, ShinyGO: A graphical enrichment tool for animals and plants. bioRxiv 315150 [Preprint]. 4 May 2018. https://doi.org/10.1101/315150.

95. V. del Solar, D. Y. Lizardo, N. Li, J. J. Hurst, C. J. Brais, G. E. Atilla-Gokcumen, Differential regulation of specific sphingolipids in colon cancer cells during staurosporine-induced apoptosis. *Chem. Biol.* **22**, 1662–1670 (2015).

96. A. Saghatelian, S. A. Trauger, E. J. Want, E. G. Hawkins, G. Siuzdak, B. F. Cravatt, Assignment of endogenous substrates to enzymes by global metabolite profiling. *Biochemistry* **43**, 14332–14339 (2004).

97. M. K. Tuck, D. W. Chan, D. Chia, A. K. Godwin, W. E. Grizzle, K. E. Krueger, W. Rom, M. Sanda, L. Sorbara, S. Stass, W. Wang, D. E. Brenner, Standard operating procedures for serum and plasma collection: Early detection research network consensus statement standard operating procedure integration working group. *J. Proteome Res.* **8**, 113–117 (2009).