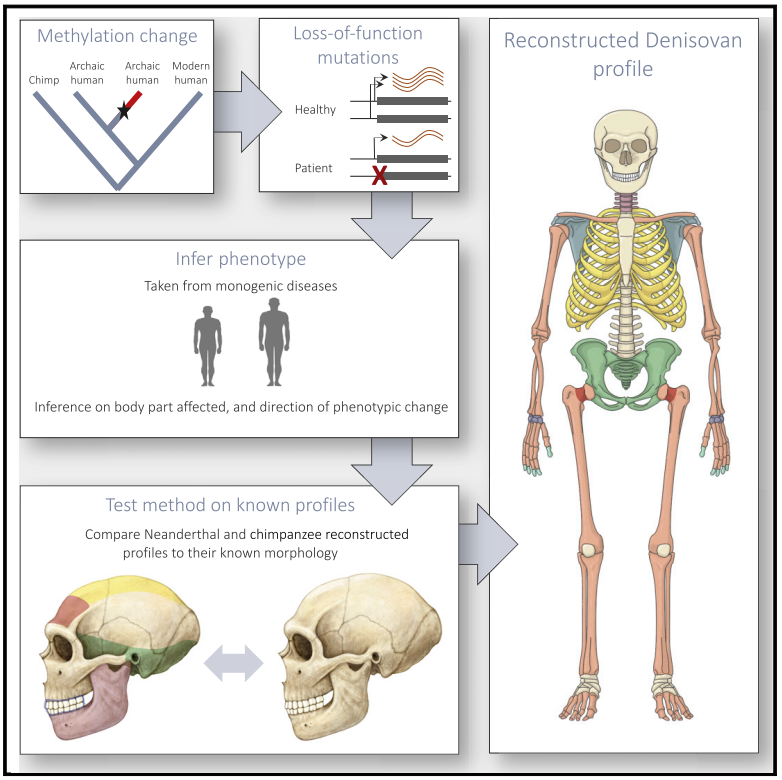
Article

Cell

Reconstructing Denisovan Anatomy Using DNA Methylation Maps

Graphical Abstract



Highlights

- DNA hypermethylation was used to infer gene downregulation
- Phenotype of downregulation was inferred from known loss-of-function phenotypes
- Neanderthal and chimp profiles reconstructed and tested against their known anatomy
- We propose a reconstructed Denisovan anatomical profile

Authors

David Gokhman, Nadav Mishol, Marc de Manuel, ..., Tomas Marques-Bonet, Yoel Rak, Liran Carmel

Correspondence

david.gokhman@mail.huji.ac.il (D.G.), liran.carmel@huji.ac.il (L.C.)

In Brief

DNA methylation maps can be used to predict anatomical features in hominins and chimpanzees, allowing for reconstruction of a putative anatomical profile of the Denisovan, currently absent from the fossil record.

Gokhman et al., 2019, Cell *179*, 180–192
September 19, 2019
https://doi.org/10.1016/j.cell.2019.08.035

CellPress

Cell

Article

Reconstructing Denisovan Anatomy Using DNA Methylation Maps

David Gokhman,[1,*] Nadav Mishol,[1] Marc de Manuel,[2] David de Juan,[2] Jonathan Shuqrun,[1,3] Eran Meshorer,[1,4] Tomas Marques-Bonet,[2,5,6,7] Yoel Rak,[8] and Liran Carmel[1,9,*]

[1]Department of Genetics, The Alexander Silberman Institute of Life Sciences, Faculty of Science, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel

[2]Institute of Evolutionary Biology (UPF-CSIC), Barcelona Biomedical Research Park (PRBB), Dr. Aiguader 88, 08003 Barcelona, Spain

[3]Alpha program, Future Scientists Center, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel

[4]The Edmond and Lily Safra Center for Brain Sciences (ELSC), The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem, 91904, Israel

[5]Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys 23, 08010 Barcelona, Spain

[6]National Center of Genomic Analyses (CNAG)-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, 08028 Barcelona, Spain

[7]Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, Cerdanyola del Vallès, 08193 Barcelona, Spain

[8]Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, 6997801, Israel

[9]Lead Contact

*Correspondence: david.gokhman@mail.huji.ac.il (D.G.), liran.carmel@huji.ac.il (L.C.)

https://doi.org/10.1016/j.cell.2019.08.035

SUMMARY

Denisovans are an extinct group of humans whose morphology remains unknown. Here, we present a method for reconstructing skeletal morphology using DNA methylation patterns. Our method is based on linking unidirectional methylation changes to loss-of-function phenotypes. We tested performance by reconstructing Neanderthal and chimpanzee skeletal morphologies and obtained >85% precision in identifying divergent traits. We then applied this method to the Denisovan and offer a putative morphological profile. We suggest that Denisovans likely shared with Neanderthals traits such as an elongated face and a wide pelvis. We also identify Denisovan-derived changes, such as an increased dental arch and lateral cranial expansion. Our predictions match the only morphologically informative Denisovan bone to date, as well as the Xuchang skull, which was suggested by some to be a Denisovan. We conclude that DNA methylation can be used to reconstruct anatomical features, including some that do not survive in the fossil record.

INTRODUCTION

Very little is known about the anatomy of Denisovans. The first specimen, Denisova 3, comprises a manual phalanx found in the Denisova cave in Siberia, dated between 74 and 82 thousand years ago (kya) (Krause et al., 2010). DNA extracted from this bone indicated that this individual belonged to a sister group of Neanderthals (Meyer et al., 2012), thereafter called Denisovans. These two groups separated 390–440 kya (Prüfer et al., 2017), and their ancestors split from our lineage between 520 and 630 kya (Prüfer et al., 2017), though these datings are still under debate (Mafessoni and Prüfer, 2017; Rogers et al., 2017). Based on this genome, Denisovan ancestry of up to 6% was detected in present-day Melanesians and Aboriginal Australians and to a lesser level in East Asians, Native Americans, and Polynesians (Meyer et al., 2012; Prüfer et al., 2014; Racimo et al., 2015; Reich et al., 2010; Skoglund and Jakobsson, 2011). Some introgressed Denisovan haplotypes might have conferred modern humans (MHs) an adaptive advantage in high-altitude (Beall et al., 2010) and cold climates (Racimo et al., 2017).

Despite our growing understanding of their genetics, findings that provide information on Denisovan anatomy remain scarce. The only confirmed Denisovan samples hitherto are the aforementioned Denisova 3 phalanx, from which a 30x genome was sequenced (Meyer et al., 2012), a lower jawbone (Chen et al., 2019), and several teeth (Chen et al., 2019; Sawyer et al., 2015; Slon et al., 2017). Anatomical studies of the teeth revealed that the Denisovan molars differ in their cusp and root morphology, and their size is outside the range of MHs and mostly outside the range of Neanderthals too (Chen et al., 2019; Sawyer et al., 2015; Slon et al., 2017). The jawbone was shown to be robust, protruding, with a long dental arcade and no chin (Chen et al., 2019). Studying differences in anatomy between human groups is critical in understanding human-specific adaptations, selective pressures, and developmental trajectories as well as the phenotypic effects of introgression events.

While the Denisovan DNA sequence potentially bears ample information on its anatomical features, our current ability to decode these data is very restricted. A direct approach is to examine the biological consequences of substitutions that

alter protein sequence. However, less than 100 fixed nonsynonymous substitutions distinguish MHs from the Denisovan and Neanderthal, whereas the remaining ~30,000 fixed changes are noncoding or synonymous (Prüfer et al., 2014). Although many of the noncoding changes are likely neutral (or nearly so), many others probably alter gene activity and may be highly informative to anatomy. However, pinpointing such variants is notoriously difficult.

A possible approach to circumvent this is to predict the combined effect of SNPs that are known to be associated with various traits. Prediction accuracy for traits such as skin, hair, and eye pigmentation exceeds 80% in Europeans (Walsh et al., 2013), but for the vast majority of traits, genome-wide association study (GWAS)-based predictions reach substantially lower accuracy levels (Price et al., 2015), including in facial morphology (Brinkley et al., 2016; Cole et al., 2016; Erlich, 2017; Liu et al., 2012; Shaffer et al., 2016). Moreover, the ability to extrapolate European-based GWASs to non-European populations was shown to be very limited (Martin et al., 2017). Perhaps most importantly, GWASs are based on within-population variability, which usually reflects variants that emerged more recently. However, older variants that separate more deeply diverged lineages and variants with considerable phenotypic effects are more likely to reach fixation and are therefore unlikely to be pinned down in GWASs, even if their effect is substantial (Martin et al., 2017; Price et al., 2015). Together, these factors limit the applicability of GWAS-detected variants in morphological analyses of deeply diverged groups, such as the Denisovan.

Ideally, to further the understanding of Denisovan anatomy, one would strive to directly measure gene expression, which is more readily interpretable than noncoding sequence changes. However, RNA molecules rapidly degrade in ancient samples and are unavailable for sequencing. Therefore, we used DNA methylation, a key regulatory layer of the genome, as a proxy for gene activity. Here, we developed a method that compares the Denisovan DNA methylation patterns to those of MHs, Neanderthals, and chimpanzees and infers which genes may have become up- or downregulated along each lineage. Next, we linked these changes to potential phenotypic alterations. We did so by analyzing phenotypes that are known to be incurred by loss-of-function mutations in these genes and could therefore be roughly paralleled with reduced activity. Importantly, unlike previous efforts to make quantitative morphological estimations (Claes et al., 2014; Erlich, 2017; Lippert et al., 2017), our aim was considerably more modest; we strove to reconstruct a qualitative skeletal profile by predicting traits that are divergent between the human groups and—when possible—to determine their direction of change. Our rationale was that providing accurate magnitudes of anatomical changes is infeasible mainly because precise activity levels of archaic human genes cannot be determined, and even in present-day samples, the quantitative contribution of each gene to the trait is currently impossible to predict. We quantified the accuracy of our method by applying it to the Neanderthal and the chimpanzee and then comparing our predictions with the known morphology of these groups. We show that we reach prediction precision of 82.8% in reconstructing traits that separate Neanderthals and MHs and 87.9% in predicting their direction of change. In the chimpanzee, we reach a similar performance, with 90.5% precision at predicting which traits are divergent and 90.9% in predicting their direction of change. By applying our method to the Denisovan, we propose a methylation-based profile of Denisovan morphology.

RESULTS

Using Changes in DNA Methylation for Phenotypic Reconstruction

We have previously developed a method to reconstruct full DNA methylation maps of ancient genomes (Gokhman et al., 2014). This method is based on the analysis of damage patterns in ancient DNA, whereby pre-mortem methylated and unmethylated cytosines leave distinct signatures that can be used to differentiate between them (Briggs et al., 2010; Gokhman et al., 2014; Pedersen et al., 2014). Using this method, we reconstructed the methylomes of Denisova 3, two Neanderthals (Altai and Vindija), and five anatomically modern humans from 45 to 7.5 kya (Gokhman et al., 2014, 2017a). Together with bone methylomes from 55 present-day humans and five chimpanzees, we were able to chart a comprehensive map of differentially methylated regions (DMRs) that emerged along the different branches of the hominin tree (Gokhman et al., 2017a).

To be able to reconstruct the morphology of an archaic individual using these DMRs, three nested conditions should be met: (1) the DMRs should reflect lineage-specific changes, rather than within-lineage variation that is driven by factors such as age, sex, or bone type; (2) lineage-specific methylation changes should reflect changes in gene activity levels; and (3) changes in gene activity should be associated with known phenotypic effects.

The first condition was accounted for during the DMR calling phase, where we harnessed the diversity across the 68 modern human, Neanderthal, Denisovan, and chimpanzee samples to filter out bone-type-, age-, disease- and sex-specific DMRs. This was achieved by examining the methylation levels of each DMR and requiring that all samples in a human lineage cluster outside all samples from the other lineages, regardless of age, sex, disease condition, or bone type. In other words, loci whose methylation levels differ across skeletal parts, age, health state, or sex were removed; see full details in Gokhman et al. [2017a]. Importantly, the fact that these DMRs have similar methylation levels across various skeletal regions of the same human group suggests that they are not constrained to a specific bone type but rather exist throughout the skeleton, including in the cranium, limbs, and teeth (Gokhman et al., 2016, 2017a).

To focus on DMRs reflecting the most extensive methylation changes, we also required that DMRs represent at least a 50% change in methylation (for example, from a regional average of 30% methylation to a regional average of 80% methylation) and span a minimum of 50 CpG positions. This filter also removes most environmentally induced DMRs, as their effect size rarely exceeds 10% methylation change (Gokhman et al., 2017b). Therefore, our list of DMRs is likely to capture marked pan-skeletal evolutionary changes between the human lineages and to be devoid of regions that show within-lineage variability

(Figure S1A). These strict filters and the use of over 60 samples are expected to considerably reduce false discoveries at the expense of leaving more subtle lineage-specific methylation changes undetected. Moreover, we conjecture that such substantial changes between lineages are more likely to result in observable phenotypes. The final list of lineage-specific DMRs includes 873 MH-derived DMRs, 939 archaic-derived DMRs (i.e., DMRs that emerged in the ancestors of Neanderthals and Denisovans), 570 Neanderthal-derived DMRs, 443 Denisovan-derived DMRs, and 2,031 DMRs that separate chimpanzees from all human groups (Figure S1B).

In order to fulfill the second condition, we needed to identify DMRs that are associated with changes in expression levels. To this end, we used the fact that the relation between DNA methylation and gene expression is strongest in gene promoters, where DNA methylation tends to be associated with gene silencing (Jones, 2012). Though often insufficient to drive gene silencing on its own (Ford et al., 2018), methylation can nevertheless be used as a marker of downregulation. Estimates of the correlation between promoter methylation and gene-silencing range between 0.3 and 0.7, depending on the surrounding methylation landscape and distance to transcription start site (TSS) (Kapourani and Sanguinetti, 2016; Schultz et al., 2015). These correlation levels are too low to allow precise predictions of gene expression levels based on DNA methylation alone. Here, however, we do not seek to obtain precise quantitative predictions of the changes in expression level associated with methylation changes. Rather, given that our DMRs represent sizeable methylation changes (>50% difference), our goal is to determine whether such substantial increase in promoter methylation is associated with reduced expression levels. Indeed, regardless of the strength of the methylation-expression correlations, an increase in promoter methylation is rarely associated with gene activation (Jones, 2012; Teschendorff and Relton, 2018; Yin et al., 2017). For example, in expression-associated CpGs that are located up to 1 kb from a TSS, methylation is associated with gene silencing in 82.1% of the cases (Wagner et al., 2014). Furthermore, when looking at a wider region of −5–1 kb around the TSS, only 21.6% of genes show more CpGs that are positively, rather than negatively, associated with expression (Wagner et al., 2014). Thus, a change in promoter methylation from hypomethylation to hypermethylation, especially a pronounced change of >50% methylation difference, will generally (but not always) reflect decreased gene activity (Jones, 2012; Kapourani and Sanguinetti, 2016; Schultz et al., 2015; Teschendorff and Relton, 2018; Wagner et al., 2014).

For the purpose of this work, we therefore only took DMRs located −5–1 kb around TSS (hereinafter, promoter DMRs). This yielded 154 MH-derived, 171 archaic-derived, 113 Neanderthal-derived, 55 Denisovan-derived, and 415 chimpanzee-human promoter DMRs (Table S1). To verify that hypermethylation of these promoter DMRs is indeed associated with decreased expression levels of the corresponding genes, we analyzed data from the Roadmap Epigenomics and GTEx Projects, where methylation and expression levels were measured in the same individual. Overall, these data cover 21 tissues. We computed the correlation of DNA methylation and expression for each promoter DMR and found that, for 90.7% of the DMRs where the correlation is significant (false discovery rate [FDR] < 0.05), hypermethylation is associated with gene silencing. Therefore, we used promoter hypermethylation as a marker of reduced gene activity. However, we accounted for exceptions by interpreting hypermethylation in the five promoter DMRs where a significant positive correlation was identified between methylation and expression levels (FDR < 0.05), as a marker of up- rather than downregulation.

Although some methylation patterns are plastic during early stages of development, alterations slow down drastically in late development, and the methylome becomes mostly stable through adulthood, especially in promoter regions. Thus, methylation levels in the majority of CpGs in the adult state reflect earlier developmental stages, and postnatal methylation differences between two individuals often reflect divergence at earlier stages of development (Alisch et al., 2012; Hon et al., 2013; Numata et al., 2012; Ziller et al., 2013).

The third and final condition that should be met in order to carry out anatomical reconstruction is that changes in gene activity should be associated with a known phenotypic change. We define differentially methylated genes (DMGs) as genes whose promoter harbors a DMR. To link DMGs to the trait they might underlie, we used the Human Phenotype Ontology (HPO) database of gene-phenotype associations (Köhler et al., 2014). HPO is based on ~4,000 monogenic human disorders, taken from highly curated databases (OMIM, Orphanet, and DECIPHER), which are translated to over 100,000 gene-phenotype associations. HPO thus serves as a comprehensive source for genetically driven human phenotypes. The use of disease phenotypes as a platform to infer the morphological effects of genes is supported by the observation that genes that underlie genetic disorders tend to underlie morphological variation within humans, as well as between humans and chimpanzees (Claes et al., 2018).

Importantly, most gene-disease associations at the basis of HPO come from works that linked protein sequence, rather than regulatory mutations, to diseases. Many of these mutations result in loss of function, where one or both copies of a gene are dysfunctional. Examples of such mutations include frameshifts, gain of stop codons, or complete gene deletions (Hamosh et al., 2005; Li et al., 2015; MacArthur et al., 2012). Loss-of-function mutations are sometimes counterbalanced by upregulation of the other gene copy (El-Brolosy and Stainier, 2017). However, this is unlikely to be the case in HPO phenotypes, as their underlying mutations were shown to result in an observable phenotype, suggesting that their loss of function is not fully compensated. Therefore, the loss of function that underlies most phenotypes in HPO could be paralleled to a partial or complete decrease in gene activity. Although the exact level of decrease associated with hypermethylation is impossible to evaluate, these phenotypes are key to understanding the direction of phenotypic change when these genes show reduced activity in humans. We use this logic to associate the phenotypes reported in HPO with the hominin group exhibiting the gene-silencing patterns (Figure 1). For all DMGs eventually used for the Denisovan skeletal reconstruction, we have validated the assertion that the observed phenotype is a result of loss of function. We have also asserted that most phenotypes are

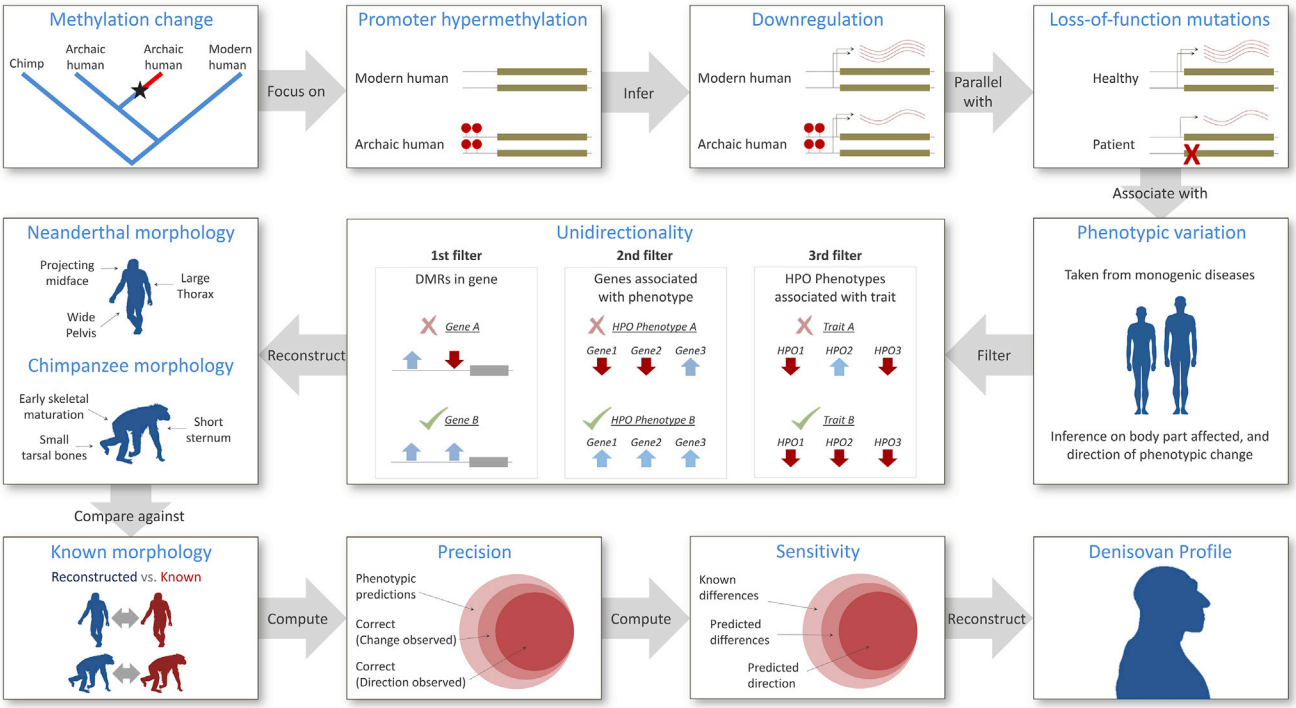


Figure 1. The Pipeline of Reconstructing Denisovan Anatomy Using DNA Methylation Changes

The first five boxes summarize the process of linking methylation changes to phenotypes, the sixth box shows the unidirectionality filters, the seventh to tenth boxes show how the reconstruction accuracy was computed, and the eleventh box reflects the final step of applying our method to the Denisovan. Promoter methylation changes along each of the hominin branches were translated into corresponding gene expression changes using the assertion that hypermethylation of the promoter is associated with downregulation of the gene. Downregulation was paralleled with loss-of-function mutations, for which the phenotypic effect is known. Next, unidirectionality filtering was applied to identify traits for which direction of morphological change could be predicted. Accuracy of the method was evaluated by reconstructing skeletal profiles of Neanderthals and chimpanzees and then matching them against these organisms' known morphology.

known to be driven by heterozygous mutations, suggesting that even partial loss of function results in a phenotype (see STAR Methods). HPO also contains information on the penetrance of each phenotype, dividing them into *frequent* and *non-frequent* based on whether they appear in >50% of patients or not (Köhler et al., 2014). Here, we used only *frequent* HPO phenotypes.

The DNA methylation maps that we use came from bones and teeth. Hence, linking DMRs to expression changes would be most reliable in these tissues. We have therefore used Gene ORGANizer—a database linking genes to the organs they phenotypically affect (Gokhman et al., 2017c)—to further filter the list of phenotypes and discard genes that are unknown to affect bones or teeth. This left a total of 597 skeletal DMGs on the hominin or chimpanzee lineages, linked to 1,528 skeleton-related HPO phenotypes (Table S2).

Filtering for Unidirectionality

We divided all phenotypes into two groups: directional and non-directional. Directional phenotypes were defined as those that could be described along a one-dimensional axis of change, such as higher-lower, accelerated-delayed, etc. Examples of such phenotypes include *Delayed skeletal maturation* (HP:0002750) and *Biparietal narrowing* (HP:0004422). Non-directional phenotypes are phenotypes that could not be described on a one-dimensional scale, such as *Abnormality of the face* (HP:0000271), and *Dental malocclusion* (HP:0000689). Here, we strove to predict the direction of morphological changes. Therefore, non-directional phenotypes were discarded from our analysis, leaving 815 skeleton-related directional HPO phenotypes, each linked to the gene or genes that have been shown to underlie the phenotype (Table S2). We then intersected this list with the skeletal DMGs and determined an expected direction of change in expression (hyperactive or hypoactive) in each hominin relative to MHs (Figure 1; Table S3).

We make a distinction between two types of predictions: predictions of trait divergence (e.g., finger length differs between MHs and Neanderthals) and predictions of trait divergence where a direction of change could be assigned as well (e.g., the fingers are longer in MHs compared to Neanderthals). One can think of each promoter DMR that is linked to an HPO phenotype as a predictor of the direction of divergence of that phenotype. If all predictors that relate to the same trait pointed to the same direction (hereinafter, unidirectionality), we make a directional prediction to this trait. However, if different predictors provided contradicting conclusions regarding the direction of change, we cannot ascertain a direction of change and only predict that the trait is divergent (Figures 2A and S2A). Theoretically, such opposing directions of change could point to no net phenotypic change or to neutral evolution.

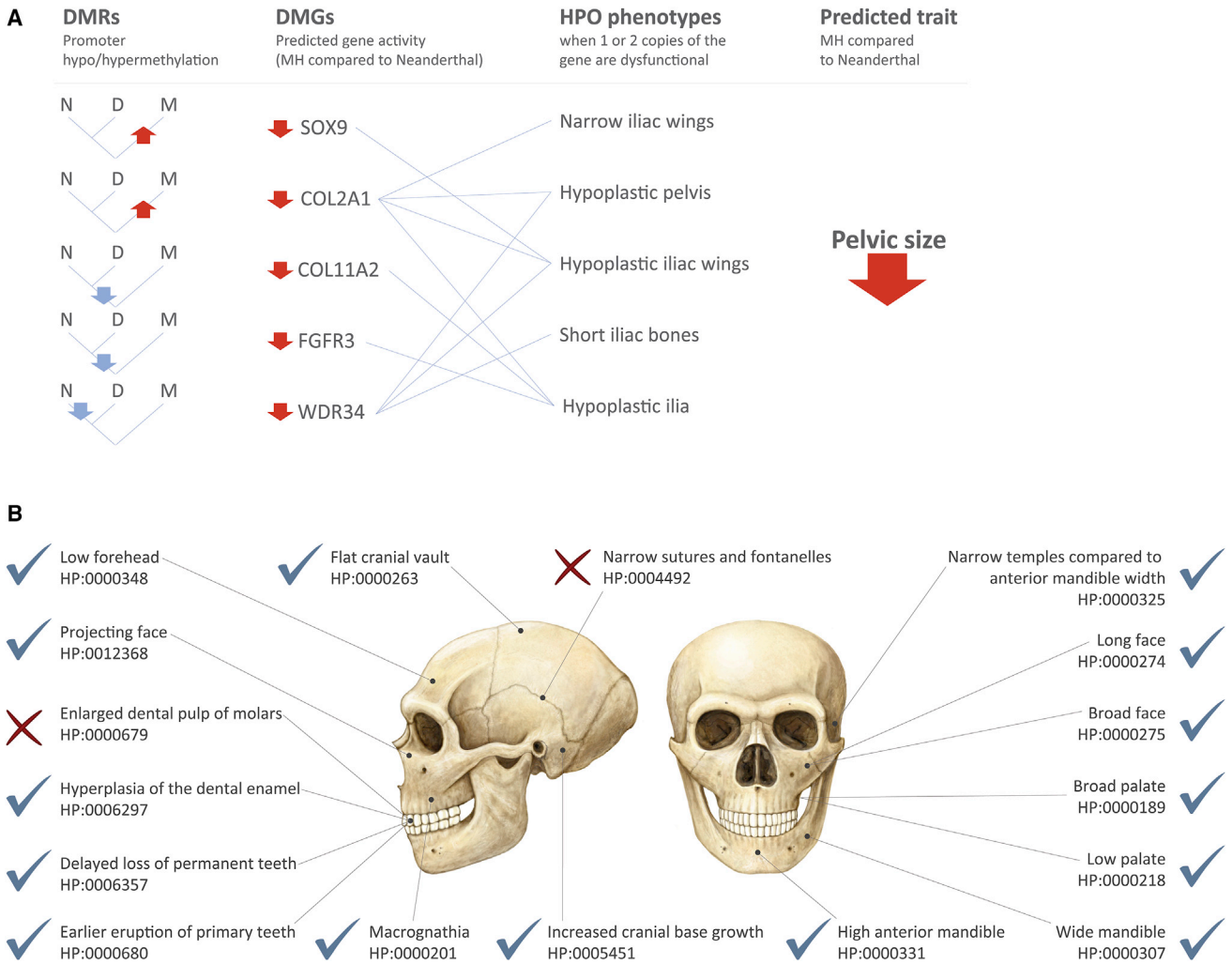


Figure 2. Unidirectional Promoter Methylation Changes Are Predictive of Neanderthal Anatomy

(A) An example of a unidirectional trait. See Figure S2A for an example of a non-unidirectional trait. The DMRs column shows promoter DMRs along the different lineages (N, Neanderthal; D, Denisovan; M, modern human). Up and down arrows mark hyper- and hypomethylation, respectively. The DMGs column shows predicted gene activity change for each of the DMGs. Increased or decreased activity is marked with up or down arrows, respectively. All five genes show patterns of decreased activity in MHs compared to Neanderthals. The HPO phenotypes column shows phenotypes associated with each of the genes. All five phenotypes associated with the genes point to reduced pelvic size in MHs compared to Neanderthals.

(B) The 17 divergent skull phenotypes for which a direction of change could be assigned. See Figure S2B for non-craniofacial phenotypes. Whenever overlapping phenotypes were merged, the displayed HPO ID belonged to the most general phenotype. Each prediction was tested against known Neanderthal morphology. Checkmarks represent correct predictions (trait is divergent and was predicted in the direction that matches known morphology); X marks predictions where the known Neanderthal phenotype is opposite to the prediction.

However, this is less likely in DMRs that are in promoters and show a large methylation change. Overlapping predictions (e.g., *Short iliac bones* and *Hypoplastic ilia*) were grouped together and counted as one trait prediction (see STAR Methods).

Unidirectionality filtering was applied in three nested levels: (1) on DMRs within the same gene promoter, (2) on DMGs linked to the same HPO phenotype, and (3) on HPO phenotypes linked to the same trait (Figures 1 and S3). For example, there are four DMGs associated with the HPO phenotype *High forehead* (HP:0000348): *LETM1* and *SOX9*, whose promoter became hypermethylated along the MH lineage, and *FGFR3* and *NELFA*, whose promoter became hypomethylated in archaic humans. Thus, it passed the unidirectionality filter as all four DMRs point to lower gene activity in MHs and, thus, probably a higher forehead (as indeed observed in MHs compared to Neanderthals [Weaver, 2009]). In cases where the DMRs point to different directions, we predict that the trait is divergent but do not predict a direction of change. As demonstrated later, a focus on unidirectional predictions reduces the number of predictions but at the same time increases prediction strength. Moreover, this filtering is likely to enrich for traits that are under selection, as selection is expected to drive changes in gene activity in the same direction (Fraser et al., 2010).

Performance of Methylation-Based Anatomical Reconstruction

In order to estimate the performance of our approach, we applied it to two groups for which we have extensive skeletal information: the Neanderthal and the chimpanzee. For each group, we evaluated two measures of performance: *precision* (*PRE*), defined as the fraction of true predictions from the total number of predictions, and true positive rate, or *sensitivity* (*SEN*), defined as the fraction of all known divergent traits that we correctly predicted. As mentioned above, we made a distinction between two types of predictions: predictions that a trait has diverged between the hominin groups (*divergence*) and predictions of the direction of change (*direction*).

Focusing first on the Neanderthal, we predicted 64 skeletal traits where Neanderthals are expected to differ from MHs (Table S4). We defined a prediction to be correct if the trait was previously shown to be divergent between Neanderthal and MH skeletons. To this end, we assembled a list of 107 known Neanderthal-MH differences (Table S5; see STAR Methods), including traits whose phenotypic distributions partly overlap between the groups. Of these divergent traits, 75 had at least one corresponding HPO phenotype and could therefore be compared against our predictions. Out of our 64 predictions, 53 are indeed known to be divergent between Neanderthals and MHs, giving $PRE_{divergence}$ = 82.8%. Of the 53 correctly predicted traits, 33 passed all three unidirectionality filters and were thus assigned a direction of change as well. Out of these 33 traits, we correctly predicted the direction of 29, giving $PRE_{direction}$ of 87.9% (Figures 2B, S2B, and 3A; see Quantification and Statistical Analysis).

To assess the sensitivity of our approach, which measures how many of the known Neanderthal traits we are able to detect using DNA methylation, we analyzed the aforementioned list of traits (Table S5) and examined how many of them were predicted by our method. We identified 62 of the 75 divergent traits, giving $SEN_{divergence}$ of 82.7%. Of the 62 correctly predicted divergent traits, we could assign a direction of change to 46. We correctly predicted the direction of 36 out of these 46 traits, giving $SEN_{direction}$ of 78.3% (Figure 3B).

Correct predictions do not necessarily reflect high accuracy but could rather stem from elevated probability to detect divergent traits by chance. Thus, in order to further evaluate the performance of our approach, we compared its precision and sensitivity to random expectation (see STAR Methods). For precision, the random expectation of $PRE_{direction}$ is 0.5, as each trait has an equal probability to change in one direction or the opposite. To assess the random expectation of $PRE_{divergence}$, we annotated all 813 directional skeletal terms on HPO. For 644 of them (79.0%), we were able to determine whether they differ between Neanderthals and MHs, based on previous comparative analyses of Neanderthal morphology (Tables S2 and S5). The rest were discarded due to insufficient information in the fossil record (e.g., *Delayed calcaneal ossification*, HP:0008142), ambiguity (e.g., *Skeletal dysplasia*, HP:0002652), or contradicting evidence (e.g., *Long phalanx of finger*, HP:0006155, where Neanderthals have shorter proximal, but longer intermediate and distal, phalanges [Klenerman and Wood, 2006]). We found that 361 of the 644 HPO phenotypes differ between Neanderthals and MHs, giving a random expected value of $PRE_{divergence}$ of 56.1%. Interestingly, the number of divergent phenotypes varies considerably across anatomical regions. For example, 81.0% of facial phenotypes differ between the groups, suggesting that the facial region is particularly divergent ($p = 2.8 \times 10^{-5}$, hypergeometric test). This was previously proposed based on comparative anatomical studies (Weaver, 2009) as well as DNA methylation patterns (Gokhman et al., 2017a). Our approach performs significantly better than random expectation, both in predicting divergent traits ($PRE_{divergence}$ = 82.8% compared to 56.1%, 1.48x, $p = 1.5 \times 10^{-6}$, hypergeometric test) and in predicting the direction of change ($PRE_{direction}$ = 87.9% compared to 50%, 1.76x, $p = 5.5 \times 10^{-6}$, Figure 3A). We repeated the above analysis while dividing the phenotypes into several skeletal regions. In all cases, our prediction power is significantly higher than random expectation (Figure 3C).

We also evaluated the sensitivity of our method compared to random expectation. To this end, we randomly replaced DMGs with non-DMGs that are associated with skeletal phenotypes and examined the number of divergent traits linked to these random genes. We repeated this process 10,000 times and found that the number of divergent traits detected using our method is significantly higher than expected by chance ($SEN_{divergence}$ = 82.7% compared to 61.0%, 1.36x, $p = 1.0 \times 10^{-4}$).

To assess the contribution of the different filters to the overall performance of our method, we tested four subsets of DMRs with increasing filtering stringency. Each subset of DMRs added a filtering step to the previous subset: fixation, overlap with promoters, and unidirectionality. For each subset of DMRs, we examined its ability to predict Neanderthal skeletal features. We found that the probability of any DMR to be linked to a divergent trait is only 2.3% higher than expected by chance (p = 0.3, hypergeometric test). However, this probability increases with every filtering step, reaching +68.3% for non-variable promoter DMRs that passed the unidirectionality filtering ($p = 9.5 \times 10^{-3}$, Figure 3D). Thus, we conclude that prediction power significantly improves by (1) focusing on fixed or nearly fixed DMRs, (2) focusing on promoter DMRs, and (3) applying unidirectionality filters.

We have also lowered the stringency of our DMR-detection method in order to study how this affects prediction accuracy. We found that while relaxing the criteria for DMR-detection slightly reduces performance, it is still significantly higher than expected by chance (see STAR Methods; Table S1).

To further examine the robustness of our approach, we applied it on chimpanzees, which are ~10x more deeply diverged from MHs compared to Neanderthals. We reached similar precision levels to the ones obtained in the Neanderthal reconstruction: 42 traits were predicted to differ between the groups, of which 38 are indeed known to separate humans and chimpanzees ($PRE_{divergence}$ = 90.5% compared to random expectation of 75.0%, 1.21x, $p = 9.0 \times 10^{-3}$, hypergeometric test). 22 of the traits passed all three unidirectionality filters, and thus, we were also able to predict their direction. In 20 of them, the prediction was correct ($PRE_{direction}$ = 90.9% compared to random expectation of 50%, 1.82x, $p = 6.1 \times 10^{-5}$).

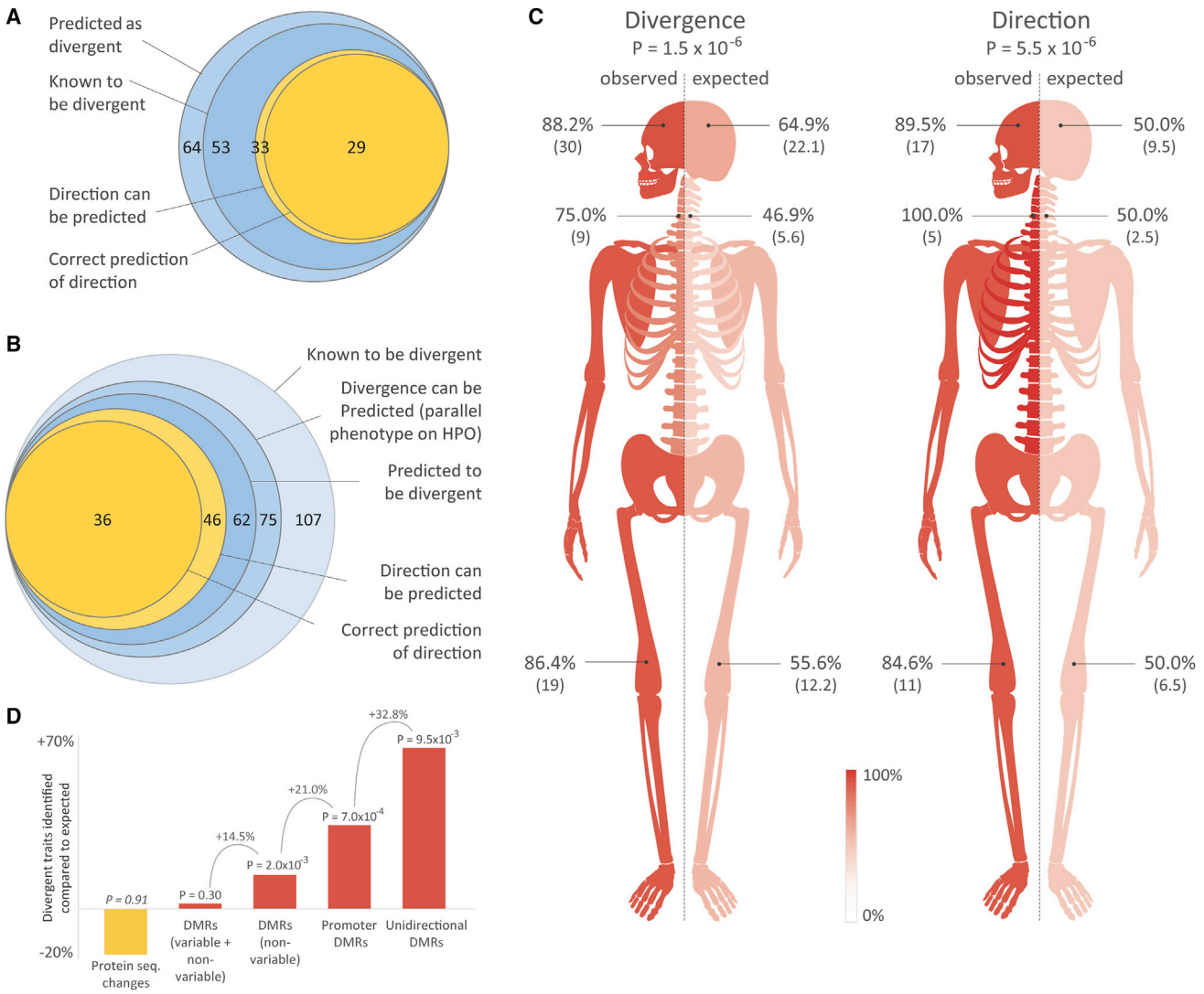


Figure 3. Precision and Sensitivity of Neanderthal Skeletal Reconstruction

(A) Method precision. From 64 predictions of divergent traits, 53 (82.8%) are indeed observed to be derived in the fossil record. Out of the 33 divergent traits for which a direction of change could be assigned, the direction is predicted correctly in 29 (87.9%).

(B) Method sensitivity. Out of 107 traits that are known to separate Neanderthals and MHs, 75 have an equivalent HPO phenotype, 62 of which (82.7%) are identified as divergent by our method. For 46, a direction could be assigned to the morphological change, and in 36 of them (78.3%), the direction matched the observations from the Neanderthal fossil record.

(C) Observed versus expected probability to detect a divergent phenotype ($PRE_{divergence}$, left skeleton) and probability to assign the right direction to the phenotypic change ($PRE_{direction}$, right skeleton) shown for each of the three skeletal regions (skull, axial, and appendicular). Absolute numbers are shown in parentheses. p values are computed using hypergeometric and binomial tests. Prediction accuracy is significantly higher than random predictions, both for $PRE_{divergence}$ and $PRE_{direction}$.

(D) The ability to detect divergent traits was tested for five groups of genes: (1) genes containing protein sequence changes; (2) genes overlapping any DMR in their gene body or promoter; (3) genes overlapping a DMR in their gene body or promoter, for DMRs that show little to no variability within human population; (4) genes whose promoter overlaps a DMR with little variability; and (5) genes that passed the three unidirectionality filters. Genes in each of the five groups were randomly replaced with skeleton-related genes, and the number of divergent traits linked to each group was examined. This permutation test was repeated 10,000 times for each of the five groups. In each group, we computed the ratio $(obs - exp) \div exp$ for the number of divergent traits that are linked to the genes. Fixed unidirectional promoter DMRs are 68.3% more likely to be associated with divergent traits than non-filtered DMRs, and 87.3% more likely than protein sequence changes.

The fraction of known chimpanzee traits that we are able to predict (i.e., sensitivity) is lower compared to the Neanderthal: $SEN_{divergence}$ = 61.4%, 1.15x, p = 0.08; $SEN_{direction}$ = 71.9%, 1.59x, p = 0.04; Table S4; Figures S1C–S1E), as expected due to the deeper divergence time of chimpanzees.

Finally, to examine the performance of our method when applied to other types of data, we tested it on enhancer marks as well as expression changes separating humans and chimps (Prescott et al., 2015). Here too, we reach prediction levels significantly higher than expected by chance, particularly for

the expression data ($PRE_{divergence}$ = 87.5%, p = 0.039; $PRE_{direction}$ = 76.9%, 1.54x, p = 0.047; Table S6; see STAR Methods). To compare between the prediction power of methylation and expression, we down-sampled our DMR list to resemble the number used to produce the aforementioned expression data. With a similar number of samples, expression-based analysis provides better predictions of divergent traits than methylation-based analysis (87.5% compared to 75.4%, see STAR Methods). This is unsurprising given that in our method methylation serves as a proxy for expression.

In summary, using three steps—(1) linking promoter methylation changes to downregulation, (2) associating these changes with phenotypes of loss-of-function mutations, and (3) applying unidirectionality filtering—we are able to reconstruct phenotypic profiles with ~87% precision and ~73% sensitivity. For about half of the traits, we are able to predict the direction of change, with ~89% precision and ~76% sensitivity (Figures 2, 3, S1, and S2; Table S4).

Reconstructing Denisovan Anatomy

Given the precision and sensitivity demonstrated by our approach, we turned to apply our method to the Denisovan. Denisovan traits can be divided into two classes: traits where the Neanderthal and Denisovan both differ from MHs and traits where the Denisovan differs from the Neanderthal. The former class represents traits that have likely evolved along the MH lineage or in the ancestors of Neanderthals and Denisovans, while the latter represents traits that have likely evolved along one of the archaic human branches. Given the number of DMRs identified along each branch and the phylogenetic architecture (Figure S1B), most traits are expected to fall into the first category. These are also expected to be the most reliable predictions, as they are based on DMRs that are fixed across all 60 MH samples and where methylation patterns differ between MHs and archaic humans throughout the skeleton, including the skull (Gokhman et al., 2017a). To enhance our prediction accuracy for this class, we took only traits where the prediction for the Neanderthal was correct; i.e., we considered only DMGs whose methylation patterns (which the Denisovan and Neanderthal share) were shown to be predictive of the phenotypic change. We also conducted a comparison of the Denisovan to the Neanderthal using Denisovan-derived and Neanderthal-derived DMGs to identify traits in which the two hominins differ from one another.

Overall, we identified 56 traits in which the Denisovan is expected to be different from MHs or Neanderthals, and in 32, we were able to predict a direction of change (Table S7). As the method is based on associating HPO phenotypes with the hominin group exhibiting downregulation patterns, we verified that all HPO phenotypes used in the analysis are indeed a result of loss-of-function mutations in human patients (Hamosh et al., 2005) (Table S3). The 32 unidirectional traits allowed us to suggest a reconstruction of Denisovan skeletal features (Figures 4 and 5). As expected, most Denisovan traits (21 out of 32) are predicted to be shared with the Neanderthal. The 21 shared traits include characteristic Neanderthal features such as robust jaws, low cranium, increased cranial base growth, low forehead, thick enamel, wide pelvis, large femoral articulations, wide fingertips, and large ribcage (Aiello and Dean, 2002; Weaver, 2009). The other 11 unidirectional traits are based on DMGs with distinct Denisovan or Neanderthal evolution and therefore represent morphological aspects where the two archaic humans are expected to differ from one another. These 11 traits can be divided into three groups according to the branches in which their associated DMGs emerged. The first group includes changes that emerged along the Denisovan lineage, i.e., traits where Denisovan features are expected to differ from both MHs and Neanderthals, and includes three morphologies: (1) elongated dental arch, (2) enlarged mandibular condyle (the posterior protuberance which is part of the mandibular joint), and (3) biparietal expansion, i.e., increased distance between the parietal bones of the cranium. The second group includes Neanderthal-specific changes and therefore represents traits where the Denisovan is expected to resemble MHs rather than Neanderthals. This group includes two morphologies: (1) broad temporal bones compared to the width of the anterior mandible and (2) premature loss of permanent teeth. The last group includes six changes where several of their underlying DMGs emerged along the MH or archaic branches and others emerged in the Neanderthal. In these traits, the three hominin groups are expected to differ from one another. These traits are bone mineral density, face width, metaphyseal and diaphyseal width, facial protrusion (prognathism), scapular size, and skeletal maturation timing (Figures 4 and 5).

Unlike the Neanderthal skeletal profile, which can be matched against fossil evidence, the Denisovan profile can only be validated at this point with regard to the mandible and teeth. When this manuscript was in the process of peer review, the first confirmed Denisovan jawbone was reported (Chen et al., 2019). This jawbone presents a unique opportunity to test the accuracy of our predictions. We provided four predictions for the mandibular region: (1) high anterior mandible, (2) wide anterior mandible, (3) mandibular protrusion, and (4) long dental arch (Figure 4). Each of these predictions can be tested on the Denisovan mandible twice: once against MHs and once against Neanderthals. We found that seven out of eight predictions match the morphological description of the Denisovan jawbone—the authors report high anterior versus posterior mandible, a very wide anterior mandible, mandibular protrusion, and a long dental arch (Chen et al., 2019). Our profile did not match the reported mandible in the anterior width when compared to Neanderthals—we predicted it to be similar to Neanderthals, whereas the authors report an anterior mandible that is considerably longer than that of Neanderthals (see STAR Methods).

DISCUSSION

Interestingly, many of the Denisovan traits we reconstruct were identified in Middle and Late Pleistocene fossils from China. These fossils display various Neanderthal-like characteristics, but their phylogenetic classification remains undetermined (Bae, 2010; Li et al., 2017). Probably the most Neanderthal-like are the 100,000- to 130,000-year-old crania from Xuchang, eastern China. The similarity of these crania to those of Neanderthals, together with their eastern geographical location, raise the possibility that they might belong to Denisovans. However,

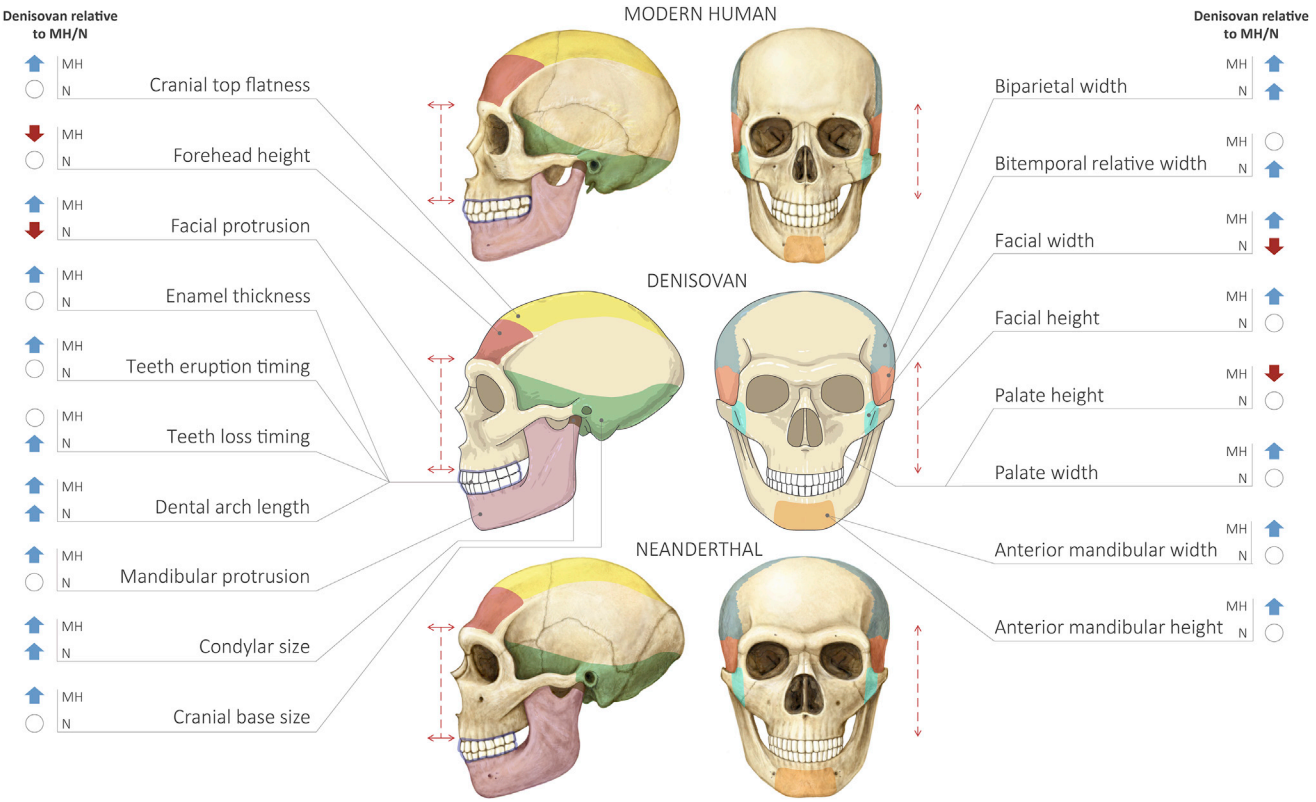


Figure 4. Reconstructed Profile of the Denisovan Skull

Colors on the Denisovan skull mark reconstructed Denisovan traits. The equivalent regions in the modern human and Neanderthal skulls are marked as well. Traits for which direction of change could be determined are labeled. Blue and red arrows represent the direction of change in the Denisovan compared to modern humans (MH) and Neanderthals (N), and empty circles represent no detectable difference. For example, the Denisovan is expected to have a lower forehead compared to MHs and similar to Neanderthals. Upward-facing arrows in the teeth eruption and loss traits represent an earlier timing. Regions for which there is no reconstruction were illustrated in a more general way. Face height (i.e., the vertical length of the face) and face protrusion (how much the face projects forward) are marked with dashed lines. The figure depicts an adult, as reconstruction was based on DMRs that are age independent.

without DNA, this could not be confirmed. The bones include the skull cap and base, but not the face or the jaws, and exhibit the following ten directional morphologies: (1) lateral expansion of the temporal bones; (2) low cranial vault; (3) lateral expansion of the parietal bones, outside the range of Neanderthals and MHs (Suzuki and Takai, 1970); (4) wide cranial base; (5) cranial gracility; (6) prominent supraorbital tori; (7) reduced thickening (restricted nuchal torus) of the occipital bone; (8) sagittal flatness; (9) short inward-sloping mastoid process; and (10) small anterior semicircular canal radii and more superior lateral versus posterior canals. Traits 1–8 have equivalent phenotypes on HPO and could thus be examined against our reconstructed profile. Strikingly, seven of them were identified as divergent traits in our reconstructed Denisovan profile (Figure 4; Table S7). Traits 1–4 are all linked to unidirectional methylation changes, and therefore, their directionality could be predicted too. The directionality we report for all of these traits matches the directionality observed in the Xuchang fossils. Even more outstanding is the fact that we predict no other divergent traits related to the cranial vault (i.e., the region that was preserved in the Xuchang crania) except for the seven that are observed in the Xuchang fossils (Li et al., 2017) ($p = 0.035$, Figure 4). The almost complete overlap between the Xuchang crania and our reconstructed profile provides the first genetic support to the notion that the Xuchang skulls are related to Denisovans.

Denisovan molars are observed to be significantly broader than that of MHs and also mostly outside the range of Neanderthal molars (Chen et al., 2019; Sawyer et al., 2015; Slon et al., 2017). We did not predict larger molars, although this might be related to the fact that, while the HPO database includes some phenotypes that describe changes in tooth size, it does not include phenotypes that are specific to molar size or breadth. However, we do predict longer dental arch, which could potentially be linked to large molars.

The ability to extract phenotypic information from a single gene is limited. Here, we demonstrated that by (1) looking at the set of genes that underlie a trait, (2) considering only marked promoter methylation changes, and (3) adding unidirectionality filters high phenotypic reconstruction accuracy can be attained. Moreover, as our directional predictions pass all three unidirectionality filters, they are likely to represent either higher-level regulatory changes that cascaded to affect several loci (e.g.,

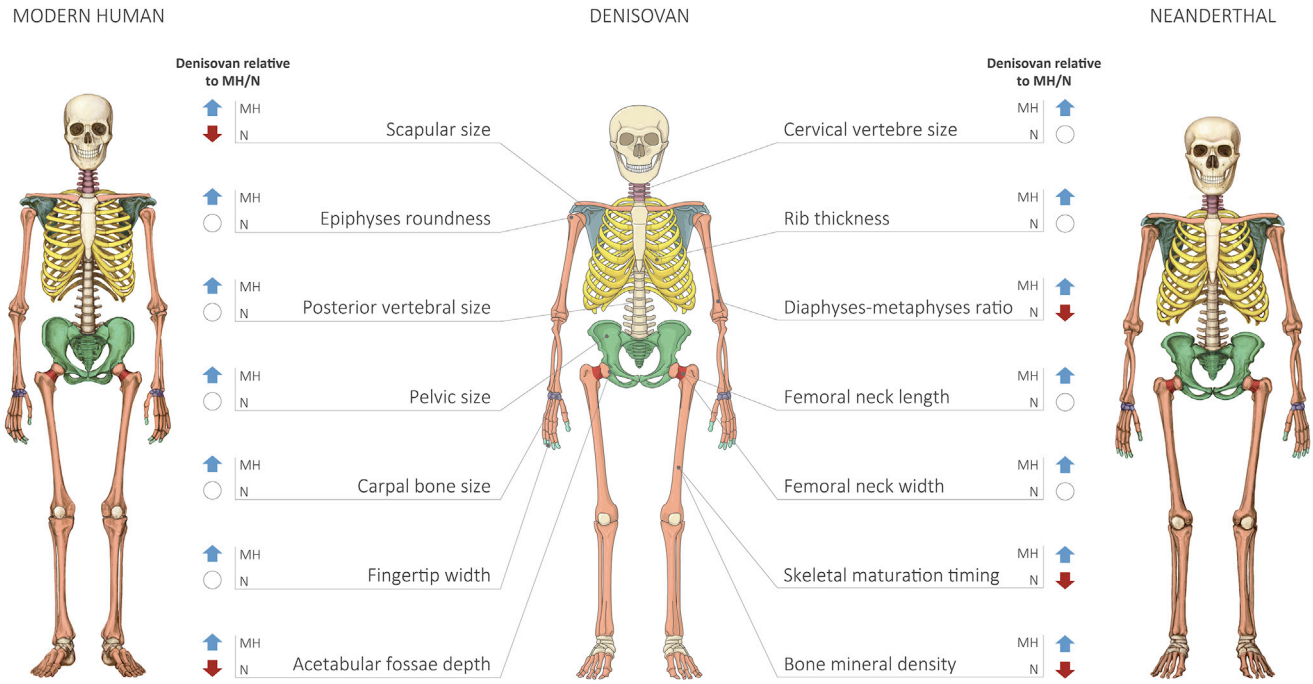


Figure 5. Reconstructed Profile of the Denisovan Skeleton

Colors on the Denisovan skeleton mark reconstructed Denisovan traits. The equivalent regions in the modern human and Neanderthal skeletons are marked as well. Traits for which direction of change could be determined are labeled. Blue and red arrows represent the direction of change in the Denisovan compared to modern humans (MH) and Neanderthals (N), and empty circles represent no detectable difference. Regions for which there is no reconstruction were illustrated in a more general way. The figure depicts an adult, as reconstruction was based on DMRs that are age independent.

silencing of a transcription factor that affects the promoter methylation of multiple target genes) or, alternatively, the process of polygenic adaptation, where one direction of phenotypic change is advantageous, thus resulting in coordinated changes in several loci (Fraser et al., 2010). Both options are likely to drive pronounced phenotypic alterations and, therefore, potentially explain why the predictive power increases when looking at the collective effect of unidirectional changes.

We summarize the main strengths and weaknesses of the proposed method. (1) Its predictions offer a comparative direction of change, rather than precise quantitative evaluation of the extent of phenotypic change. (2) In instances where there is no unidirectionality, the direction of change cannot be determined. (3) Morphologies that have no equivalent term on HPO cannot be reconstructed. (4) Reconstructed traits that are solely based on Denisovan-specific DMRs (3 predicted traits out of 56 in total) could represent only the Denisova 3 individual, rather than the entire Denisovan population. (5) Due to the strict unidirectionality filtering, and because of the varying number of individuals used for DMR detection along each branch, the method is more accurate at identifying traits that emerged earlier and became fixed or nearly fixed, while ignoring traits with high intra-population variability. At the same time, such unidirectional fixed changes are more likely to be driven by selection (Fraser et al., 2010) and could therefore represent a more interesting subset of morphological alterations. (6) The precision and sensitivity levels we report are partly dependent on the way traits are discretized, as we have clustered overlapping traits together, thus creating morphological units (see STAR Methods). Yet many morphologies could nevertheless be developmentally intertwined, as reflected by the overlapping set of genes that underlie them. (7) Finally, because Denisovan-specific DMRs are based on a single sample, it is yet to be determined which of the Denisovan morphologies we report are confined to Denisovan 3 and which reflect the Denisovan population. In this regard, several observations support the notion that the majority of the reconstructed traits are shared throughout the Denisovan population. (1) In the fossil record, it has been shown that traits that separate a single Neanderthal fossil from MHs tend to be shared by all Neanderthals (Aiello and Dean, 2002). (2) Roughly half of the reconstructed traits are based on DMRs that emerged along the MH lineage. Because such traits are derived in MHs, Denisovans are expected to share the ancestral form of the trait. (3) The entire analysis is based on DMRs that are unaffected by age, sex, or bone type, and therefore, a Denisovan individual from a different age and sex and where the sample was obtained from another bone type is expected to exhibit similar methylation patterns (Gokhman et al., 2017a).

It is illuminating that although the reconstruction accuracy levels for the Neanderthal and chimpanzee are very similar, the fraction of known derived traits that have an equivalent HPO phenotype are markedly different. For the Neanderthal, 75 out of 107 derived traits (70%) have a parallel HPO phenotype, whereas for the chimpanzee, where divergence is ~10x deeper, the fraction is significantly lower (83 out of 201, 41%,

$p = 1.5 \times 10^{-6}$, χ^2 test). This raises the possibility that along short timescales, the genes that underlie diseases may also be those that underlie evolutionary phenotypic divergence, possibly accounting for the overlap between the two lists. This is also supported by a study that suggested that the loci involved in craniofacial disorders are likely to underlie normal variation too, both within humans and between humans and chimpanzees (Claes et al., 2018). Given that divergence time between Denisovans and MHs equals that between Neanderthals and MHs, it is likely that the fraction of Denisovan traits that we can predict using HPO phenotypes is high and similar to that in Neanderthals.

The approach we presented relies on two basic hypotheses: (1) pronounced methylation changes (i.e., statistically significant unidirectional promoter methylation changes of >50% that extend across at least 50 CpG positions) are more likely to be associated with phenotypic effects than subtle changes, and (2) the direction of a phenotype driven by downregulation of a gene is expected to be similar to the direction of a phenotype driven by loss-of-function mutations. We demonstrated here that using these assumptions, we are able to reconstruct dozens of the traits that differ between human groups with over 80% accuracy. It might be surprising at first glance that the analysis of DNA methylation alone is sufficient to reach such accuracy. However, it was shown that there is high correlation and interplay between DNA methylation and other regulatory layers, such as transcription-factor binding and histone modifications (Banovich et al., 2014). Therefore, although this analysis was based on a single regulatory layer, the information that is extracted from it might reflect a sizeable portion of the regulatory landscape. Furthermore, as we did not reconstruct the extent of each phenotypic change, but rather only its direction of change, it is likely that the other regulatory layers, as well as protein changes and non-promoter methylation changes, further contribute quantitatively to the observed phenotypes.

We conclude that unidirectional promoter methylation changes can be used to identify phenotypic divergence between closely related organisms. Even though we validated this approach on Neanderthals and chimpanzees, as well as on the Denisovan jawbone, its ultimate test would be to match the reconstructed profile to a more complete collection of Denisovan samples, once they are discovered.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- METHOD DETAILS
  - Anatomically-related protein-coding variants
  - Lineage-specific methylation changes
  - Associating genes with phenotypes
  - Unidirectionality filtering
  - Validation and accuracy estimation
  - Quantifying precision
  - Quantifying sensitivity
  - Testing the method on other types of data
  - Comparing the profile to the jawbone
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - DNA methylation reconstruction
  - DMR detection
  - Lineage assignment
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cell.2019.08.035.

ACKNOWLEDGMENTS

This work has been supported by the National Geographic Society (grant HJ-111R-17 to L.C.). D.G. was supported by the Clore Israel Foundation. T.M.-B. is supported by BFU2017-86471-P (MINECO/FEDER, UE), U01 MH106874 grant, Howard Hughes International Early Career, Obra Social "La Caixa," and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya (GRC 2017 SGR 880). We would like to thank Shiran Bar, David Reich, Iain Mathieson, Chris Stringer, and members of the Fraser lab for useful advice and Maayan Harel for beautiful illustrations.

AUTHOR CONTRIBUTIONS

D.G. conceived the idea and conducted all analyses. N.M. collected data for the GWAS analyses. M.d.M., D.d.J., and T.M.-B. produced the data for the nonsynonymous analysis. J.S. annotated phenotypes. Y.R. supervised comparisons to known Neanderthal and chimpanzee traits. E.M. co-supervised the generation of the original ancient DNA methylation data. L.C. supervised the study. D.G. and L.C. designed the study and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 7, 2019
Revised: May 24, 2019
Accepted: August 20, 2019
Published: September 19, 2019

REFERENCES

Aiello, L., and Dean, C. (2002). An Introduction to Human Evolutionary Anatomy (London: Elsevier).

Alisch, R.S., Barwick, B.G., Chopra, P., Myrick, L.K., Satten, G.A., Conneely, K.N., and Warren, S.T. (2012). Age-associated DNA methylation in pediatric populations. Genome Res. *22*, 623–632.

Bae, C.J. (2010). The late Middle Pleistocene hominin fossil record of eastern Asia: synthesis and review. Am. J. Phys. Anthropol. *143* (*Suppl 51*), 75–93.

Banovich, N.E., Lan, X., McVicker, G., van de Geijn, B., Degner, J.F., Blischak, J.D., Roux, J., Pritchard, J.K., and Gilad, Y. (2014). Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. PLoS Genet. *10*, e1004663.

Beall, C.M., Cavalleri, G.L., Deng, L., Elston, R.C., Gao, Y., Knight, J., Li, C., Li, J.C., Liang, Y., McCormack, M., et al. (2010). Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders. Proc. Natl. Acad. Sci. USA *107*, 11459–11464.

Been, E., Gómez-Olivencia, A., and Kramer, P.A. (2012). Lumbar lordosis of extinct hominins. Am. J. Phys. Anthropol. *147*, 64–77.

Briggs, A.W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., and Pääbo, S. (2010). Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res. *38*, e87.

Brinkley, J.F., Fisher, S., Harris, M.P., Holmes, G., Hooper, J.E., Jabs, E.W., Jones, K.L., Kesselman, C., Klein, O.D., Maas, R.L., et al.; FaceBase Consortium (2016). The FaceBase Consortium: a comprehensive resource for craniofacial researchers. Development *143*, 2677–2688.

Chen, F., Welker, F., Shen, C.-C., Bailey, S.E., Bergmann, I., Davis, S., Xia, H., Wang, H., Fischer, R., Freidline, S.E., et al. (2019). A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau. Nature *569*, 409–412.

Chirchir, H., Kivell, T.L., Ruff, C.B., Hublin, J.-J., Carlson, K.J., Zipfel, B., and Richmond, B.G. (2015). Recent origin of low trabecular bone density in modern humans. Proc. Natl. Acad. Sci. USA *112*, 366–371.

Claes, P., Liberton, D.K., Daniels, K., Rosana, K.M., Quillen, E.E., Pearson, L.N., McEvoy, B., Bauchet, M., Zaidi, A.A., Yao, W., et al. (2014). Modeling 3D facial shape from DNA. PLoS Genet. *10*, e1004224.

Claes, P., Roosenboom, J., White, J.D., Swigut, T., Sero, D., Li, J., Lee, M.K., Zaidi, A., Mattern, B.C., Liebowitz, C., et al. (2018). Genome-wide mapping of global-to-local genetic effects on human facial shape. Nat. Genet. *50*, 414–423.

Clement, A.F., Hillson, S.W., and Aiello, L.C. (2012). Tooth wear, Neanderthal facial morphology and the anterior dental loading hypothesis. J. Hum. Evol. *62*, 367–376.

Cole, J.B., Manyama, M., Kimwaga, E., Mathayo, J., Larson, J.R., Liberton, D.K., Lukowiak, K., Ferrara, T.M., Riccardi, S.L., Li, M., et al. (2016). Genome-wide Association Study of African Children Identifies Association of SCHIP1 and PDE8A with Facial Size and Shape. PLoS Genet. *12*, e1006174.

Constant, M., Nicot, R., Vieira, A.R., Raoul, G., Sciote, J.J., and Ferri, J. (2017). Condylar geometry variation is associated with ENPP1 variant in a population of patients with dento-facial deformities. J. Craniomaxillofac. Surg. *45*, 826–830.

De Groote, I.E.P.M. (2008). A comprehensive analysis of long bone curvature in Neanderthals and Modern Humans using 3D morphometrics (University College London).

De Groote, I. (2011). The Neanderthal lower arm. J. Hum. Evol. *61*, 396–410.

de Manuel, M., Kuhlwilm, M., Frandsen, P., Sousa, V.C., Desai, T., Prado-Martinez, J., Hernandez-Rodriguez, J., Dupanloup, I., Lao, O., Hallast, P., et al. (2016). Chimpanzee genomic diversity reveals ancient admixture with bonobos. Science *354*, 477–481.

Dean, M.C., Stringer, C.B., and Bromage, T.G. (1986). Age at death of the Neanderthal child from Devil's Tower, Gibraltar and the implications for studies of general growth and development in Neanderthals. Am. J. Phys. Anthropol. *70*, 301–309.

El-Brolosy, M.A., and Stainier, D.Y.R. (2017). Genetic compensation: A phenomenon in search of mechanisms. PLoS Genet. *13*, e1006780.

Erlich, Y. (2017). Major flaws in "Identification of individuals by trait using whole-genome sequencing data.". bioRxiv. https://doi.org/10.1101/185330.

Ford, E.E., Grimmer, M.R., Stolzenburg, S., Bogdanovic, O., de Mendoza, A., Farnham, P.J., Blancafort, P., and Lister, R. (2018). Frequent lack of repressive capacity of promoter DNA methylation identified through genome-wide epigenomic manipulation. bioRxiv. https://doi.org/10.1101/170506.

Fraser, H.B., Moses, A.M., and Schadt, E.E. (2010). Evidence for widespread adaptive evolution of gene expression in budding yeast. Proc. Natl. Acad. Sci. USA *107*, 2977–2982.

Gilmore, C.C., and Weaver, T.D. (2016). Comparative perspective on antemortem tooth loss in Neandertals. J. Hum. Evol. *92*, 80–90.

Gokhman, D., Lavi, E., Prüfer, K., Fraga, M.F., Riancho, J.A., Kelso, J., Pääbo, S., Meshorer, E., and Carmel, L. (2014). Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. Science *344*, 523–527.

Gokhman, D., Meshorer, E., and Carmel, L. (2016). Epigenetics: It's Getting Old. Past Meets Future in Paleoepigenetics. Trends Ecol. Evol. *31*, 290–300.

Gokhman, D., Agranat-Tamir, L., Housman, G., García-Pérez, R., Nissim-Rafinia, M., Mallick, S., Nieves-Colón, M.A., Gu, H., Ferrando-Bernal, M., Gelabert, P., et al. (2017a). Extensive Regulatory Changes in Genes Affecting Vocal and Facial Anatomy Separate Modern from Archaic Humans. bioRxiv. https://doi.org/10.1101/106955.

Gokhman, D., Malul, A., and Carmel, L. (2017b). Inferring Past Environments from Ancient Epigenomes. Mol. Biol. Evol. *34*, 2429–2438.

Gokhman, D., Kelman, G., Amartely, A., Gershon, G., Tsur, S., and Carmel, L. (2017c). Gene ORGANizer: linking genes to the organs they affect. Nucleic Acids Res. *45* (W1), W138–W145.

Gómez-Olivencia, A., Been, E., Arsuaga, J.L., and Stock, J.T. (2013). The Neandertal vertebral column 1: the cervical spine. J. Hum. Evol. *64*, 608–630.

Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A., and McKusick, V.A. (2005). Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. *33*, D514–D517.

Hon, G.C., Rajagopal, N., Shen, Y., McCleary, D.F., Yue, F., Dang, M.D., and Ren, B. (2013). Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nat. Genet. *45*, 1198–1206.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484–492.

Kapourani, C.-A., and Sanguinetti, G. (2016). Higher order methylation features for clustering and prediction in epigenomic studies. ArXiv, 1–12, 1603.08386.

Klenerman, L., and Wood, B. (2006). The Human Foot: A Companion to Clinical Studies (Springer-Verlag London).

Köhler, S., Doelken, S.C., Mungall, C.J., Bauer, S., Firth, H.V., Bailleul-Forestier, I., Black, G.C.M., Brown, D.L., Brudno, M., Campbell, J., et al. (2014). The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res. *42*, D966–D974.

Krause, J., Fu, Q., Good, J.M., Viola, B., Shunkov, M.V., Derevianko, A.P., and Pääbo, S. (2010). The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. Nature *464*, 894–897.

Kupczik, K., and Hublin, J.J. (2010). Mandibular molar root morphology in Neanderthals and Late Pleistocene and recent Homo sapiens. J. Hum. Evol. *59*, 525–541.

Lee, M.K., Shaffer, J.R., Leslie, E.J., Orlova, E., Carlson, J.C., Feingold, E., Marazita, M.L., and Weinberg, S.M. (2017). Genome-wide association study of facial morphology reveals novel associations with FREM1 and PARK2. PLoS One *12*, e0176566.

Li, A.H., Morrison, A.C., Kovar, C., Cupples, L.A., Brody, J.A., Polfus, L.M., Yu, B., Metcalf, G., Muzny, D., Veeraraghavan, N., et al. (2015). Analysis of loss-of-function variants and 20 risk factor phenotypes in 8,554 individuals identifies loci influencing chronic disease. Nat. Genet. *47*, 640–642.

Li, Z.-Y., Wu, X.-J., Zhou, L.-P., Liu, W., Gao, X., Nian, X.-M., and Trinkaus, E. (2017). Late Pleistocene archaic human crania from Xuchang, China. Science *355*, 969–972.

Lieberman, D.E. (1998). Sphenoid shortening and the evolution of modern human cranial shape. Nature *393*, 158–162.

Lieberman, D.E., Pearson, O.M., and Mowbray, K.M. (2000). Basicranial influence on overall cranial shape. J. Hum. Evol. *38*, 291–315.

Lippert, C., Sabatini, R., Maher, M.C., Kang, E.Y., Lee, S., Arikan, O., Harley, A., Bernal, A., Garst, P., Lavrenko, V., et al. (2017). Identification of individuals by trait prediction using whole-genome sequencing data. Proc. Natl. Acad. Sci. USA *114*, 10166–10171.

Liu, F., van der Lijn, F., Schurmann, C., Zhu, G., Chakravarty, M.M., Hysi, P.G., Wollstein, A., Lao, O., de Bruijne, M., Ikram, M.A., et al. (2012). A genome-wide association study identifies five loci influencing facial morphology in Europeans. PLoS Genet. *8*, e1002932.

MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J.K., Montgomery, S.B., et al. (2012). A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding Genes. Science *335*, 823–828.

MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. *45* (D1), D896–D901.

Mafessoni, F., and Prüfer, K. (2017). Better support for a small effective population size of Neandertals and a long shared history of Neandertals and Denisovans. Proc. Natl. Acad. Sci. USA *114*, E10256–E10257.

Martin, A.R., Gignoux, C.R., Walters, R.K., Wojcik, G.L., Neale, B.M., Gravel, S., Daly, M.J., Bustamante, C.D., and Kenny, E.E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. Am. J. Hum. Genet. *100*, 635–649.

Maureille, B., and Bar, D. (1999). The premaxilla in Neandertal and early modern children: ontogeny and morphology. J. Hum. Evol. *37*, 137–152.

Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., de Filippo, C., et al. (2012). A high-coverage genome sequence from an archaic Denisovan individual. Science *338*, 222–226.

Numata, S., Ye, T., Hyde, T.M., Guitart-Navarro, X., Tao, R., Wininger, M., Colantuoni, C., Weinberger, D.R., Kleinman, J.E., and Lipska, B.K. (2012). DNA methylation signatures in development and aging of the human prefrontal cortex. Am. J. Hum. Genet. *90*, 260–272.

Pedersen, J.S., Valen, E., Velazquez, A.M.V., Parker, B.J., Rasmussen, M., Lindgreen, S., Lilje, B., Tobin, D.J., Kelly, T.K., Vang, S., et al. (2014). Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. Genome Res. *24*, 454–466.

Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., and Bejerano, G. (2013). Enhancers: five essential questions. Nat. Rev. Genet. *14*, 288–295.

Pickering, T.R., Schick, K.D., and Toth, N.P. (2007). Breathing Life into Fossils: Taphonomic Studies in Honor of C.K. (Bob) Brain (Stone Age Insitute Press).

Prescott, S.L., Srinivasan, R., Marchetto, M.C., Grishina, I., Narvaiza, I., Selleri, L., Gage, F.H., Swigut, T., and Wysocka, J. (2015). Enhancer divergence and cis-regulatory evolution in the human and chimp neural crest. Cell *163*, 68–83.

Price, A.L., Spencer, C.C.A., and Donnelly, P. (2015). Progress and promise in understanding the genetic basis of common diseases. Proc. Biol. Sci. *282*, 20151684.

Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P.H., de Filippo, C., et al. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. Nature *505*, 43–49.

Prüfer, K., De Filippo, C., Grote, S., Mafessoni, F., Korlević, P., Hajdinjak, M., Vernot, B., Skov, L., Hsieh, P., Peyrégne, S., et al. (2017). A high-coverage Neandertal genome from Vindija Cave in Croatia. Science *358*, 655–658.

Racimo, F., Sankararaman, S., Nielsen, R., and Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. Nat. Rev. Genet. *16*, 359–371.

Racimo, F., Gokhman, D., Fumagalli, M., Ko, A., Hansen, T., Moltke, I., Albrechtsen, A., Carmel, L., Huerta-Sánchez, E., and Nielsen, R. (2017). Archaic Adaptive Introgression in TBX15/WARS2. Mol. Biol. Evol. *34*, 509–524.

Raichlen, D.A., Armstrong, H., and Lieberman, D.E. (2011). Calcaneus length determines running economy: implications for endurance running performance in modern humans and Neandertals. J. Hum. Evol. *60*, 299–308.

Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., et al. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature *468*, 1053–1060.

Rogers, A.R., Bohlender, R.J., and Huff, C.D. (2017). Early history of Neanderthals and Denisovans. Proc. Natl. Acad. Sci. USA *114*, 9859–9863.

Sawyer, S., Renaud, G., Viola, B., Hublin, J.J., Gansauge, M.T., Shunkov, M.V., Derevianko, A.P., Prüfer, K., Kelso, J., and Pääbo, S. (2015). Nuclear and mitochondrial DNA sequences from two Denisovan individuals. Proc. Natl. Acad. Sci. USA *112*, 15696–15700.

Schultz, M.D., He, Y., Whitaker, J.W., Hariharan, M., Mukamel, E.A., Leung, D., Rajagopal, N., Nery, J.R., Urich, M.A., Chen, H., et al. (2015). Human body epigenome maps reveal noncanonical DNA methylation variation. Nature *523*, 212–216.

Shaffer, J.R., Orlova, E., Lee, M.K., Leslie, E.J., Raffensperger, Z.D., Heike, C.L., Cunningham, M.L., Hecht, J.T., Kau, C.H., Nidey, N.L., et al. (2016). Genome-Wide Association Study Reveals Multiple Loci Influencing Normal Human Facial Morphology. PLoS Genet. *12*, e1006149.

Skoglund, P., and Jakobsson, M. (2011). Archaic human ancestry in East Asia. Proc. Natl. Acad. Sci. USA *108*, 18301–18306.

Slon, V., Viola, B., Renaud, G., Gansauge, M.-T., Benazzi, S., Sawyer, S., Hublin, J.-J., Shunkov, M.V., Derevianko, A.P., Kelso, J., et al. (2017). A fourth Denisovan individual. Sci. Adv. *3*, e1700186.

Smith, T.M., Toussaint, M., Reid, D.J., Olejniczak, A.J., and Hublin, J.-J. (2007). Rapid dental development in a Middle Paleolithic Belgian Neanderthal. Proc. Natl. Acad. Sci. USA *104*, 20220–20225.

Smith, T.M., Tafforeau, P., Reid, D.J., Pouech, J., Lazzari, V., Zermeno, J.P., Guatelli-Steinberg, D., Olejniczak, A.J., Hoffman, A., Radovcic, J., et al. (2010). Dental evidence for ontogenetic differences between modern humans and Neanderthals. Proc. Natl. Acad. Sci. USA *107*, 20923–20928.

Suzuki, H., and Takai, F. (1970). The Amud man and his cave site (Tokyo: Tokyo University Press, Keigaku Publishing).

Teschendorff, A.E., and Relton, C.L. (2018). Statistical and integrative system-level analysis of DNA methylation data. Nat. Rev. Genet. *19*, 129–147.

Trinkaus, E. (2003). Neandertal faces were not long; modern human faces are short. Proc. Natl. Acad. Sci. USA *100*, 8142–8145.

Wagner, J.R., Busche, S., Ge, B., Kwan, T., Pastinen, T., and Blanchette, M. (2014). The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. Genome Biol. *15*, R37.

Walsh, S., Liu, F., Wollstein, A., Kovatsi, L., Ralf, A., Kosiniak-Kamysz, A., Branicki, W., and Kayser, M. (2013). The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci. Int. Genet. *7*, 98–115.

Weaver, T.D. (2009). The meaning of Neandertal skeletal morphology. Proc. Natl. Acad. Sci. USA *106*, 16028–16033.

Weber, J., and Pusch, C.M. (2008). The lumbar spine in Neanderthals shows natural kyphosis. Eur. Spine J. *17* (*Suppl 2*), S327–S330.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science *356*, eaaj2239.

Zanolli, C., Hourset, M., Esclassan, R., and Mollereau, C. (2017). Neanderthal and Denisova tooth protein variants in present-day humans. PLoS ONE *12*, e0183802.

Zilberman, U., and Smith, P. (1992). A comparison of tooth structure in Neanderthals and early Homo sapiens sapiens: a radiographic study. J. Anat. *180*, 387–393.

Ziller, M.J., Gu, H., Müller, F., Donaghey, J., Tsai, L.T.-Y., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E., et al. (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477–481.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data used and generated		
Reconstructed DNA methylation maps	Gokhman et al. (2017a)	GEO: GSE96833
Human Phenotype Ontology database (build 110)	https://hpo.jax.org/app/	N/A
Software and Algorithms		
Gene ORGANizer	Gokhman et al. (2017c)	http://geneorganizer.huji.ac.il

LEAD CONTACT AND MATERIALS AVAILABILITY

All data generated for this study appear in the supplementary material. Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Liran Carmel (liran.carmel@huji.ac.il).

METHOD DETAILS

Anatomically-related protein-coding variants

The vast majority of protein sequence changes between MHs and archaic humans stem from nonsynonymous substitutions, whose functional consequences are usually hard to predict, except in the rare cases where the same substitution exists in modern populations (Zanolli et al., 2017). Nevertheless, disease phenotypes associated with these genes could offer some clues to their effect. Out of 45 genes with fixed MH-derived substitutions (de Manuel et al., 2016; Prüfer et al., 2014, 2017), particularly noteworthy are four genes, wherein mutations that reduce protein function were shown to underlie disease phenotypes which are similar to traits that separate MHs and Neanderthals. These genes include *SETD5*, which affects nasal bridge depression, jaw size and cranial length; *ZBTB24*, which affects maxillary and mandibular projection, nose length, nasal bridge depression, palate height, and chin length; *PGM1*, which affects stature; and *ATRX*, which affects skeletal maturation rate, jaw size, nasal bridge width and depression, palate height, nose length, femoral head-to-neck angle, and more (Hamosh et al., 2005). Because the substitutions in these genes separate MHs from Denisovans too, they could potentially be used to reconstruct Denisovan morphology. However, in order to infer whether these changes underlie phenotypic divergence, one should first determine how they affect protein function, if at all. As of today, the ability to do so is very restricted, thus limiting protein-based phenotypic inference.

Lineage-specific methylation changes

Differentially methylated regions (DMRs) separating the human groups were previously identified through the comparison of 59 modern human (MH) DNA methylation maps, two Neanderthal maps, one Denisovan map and 5 chimpanzee maps (Gokhman et al., 2017a). This comparison yielded 873 MH-derived DMRs, 939 archaic-derived DMRs (i.e., the lineage leading to the last common ancestor of Neanderthals and Denisovans), 570 Neanderthal-derived DMRs, 443 Denisovan-derived DMRs, and 2,031 DMRs that separate chimpanzees and humans. The filters used in order to detect these DMRs were conservative, requiring at least 50% difference in methylation, a minimum of 50 CpGs, and that all individuals within a group cluster completely outside the range of methylation across the other groups (Gokhman et al., 2017a). In addition to evolutionary dynamics, the number of DMRs detected along each lineage is determined by two technical factors: 1. higher coverage and deamination rates lead to increased power to detect DMRs. 2. more samples within a human group lead to increased power in identifying fixed changes, hence leading to the detection of less, but more robust DMRs. For further discussion of the DMR identification process, and on removing noise variability that is not relevant to differences between the human groups, see Gokhman et al. (Gokhman et al., 2017a).

We defined promoter DMRs as DMRs that overlap at least 1 bp of the region 5 kb upstream to 1 kb downstream a transcription start site (TSS). This included 154 MH-derived DMRs, 171 archaic-derived DMRs, 113 Neanderthal-derived DMRs, 55 Denisovan-derived DMRs, and 415 chimpanzee-human DMRs (Table S1). Differentially methylated genes (DMGs) were defined as genes that have a DMR in their promoter.

Of note, the source of these methylation maps has been previously shown to be osteogenic (rather than hematopoietic or mesenchymal), but the precise cell type (i.e., osteoblasts/osteocytes/osteoclasts) has not been determined (Gokhman et al., 2014).

Associating genes with phenotypes

We used the Human Phenotype Ontology (HPO) database (build 110) to link DMGs to the phenotypes they might underlie (Köhler et al., 2014). HPO is based on ~4,000 human disorders, which were translated to over 100,000 gene-phenotype associations.

The monogenic diseases are taken from highly curated databases (OMIM, Orphanet and DECIPHER). Thus, HPO provides direct information on the phenotypes that are incurred by alterations of each gene in the database. HPO also includes information on the frequency of each phenotype (i.e., *frequent* or *non-frequent*) (Köhler et al., 2014). We took only *frequent* HPO phenotypes, as they represent the phenotypes that are most commonly associated with alterations to the gene (appear in > 50% of patients). Many of the phenotypes in HPO are driven by loss-of-function mutations, where one or both copies of a gene are dysfunctional due to mutations that cause frameshifts, gain of stop codons, splice site alterations, or partial or complete deletions (Hamosh et al., 2005; Li et al., 2015; MacArthur et al., 2012). Therefore, loss of one or both copies of a gene could be roughly paralleled to a decrease in its activity. This potentially provides information on the organ that is affected by gene silencing, as well as the directionality of the phenotypic change, but not on the extent of phenotypic change. Based on this logic, for each hypermethylated promoter we assigned putative phenotypes based on the diseases that the gene was shown to underlie. Gene-phenotype associations that were shown to be driven by gain-of-function mutations were removed from the analysis (Table S3). This is because gain-of-function mutations usually do not represent increased activity, but rather a newly acquired gene function, and this novel function could not be used to predict what phenotype would arise from gene silencing (Table S4).

The DNA methylation maps used for DMR detection come from bone and tooth tissues (Gokhman et al., 2017a). Therefore, we used Gene ORGANizer (Gokhman et al., 2017c) to discard phenotypes that do not affect these body parts (e.g., *Coarse hair*, HP:0002208). This left 1,627 out of 6,037 HPO phenotypes. We then divided HPO phenotypes into two groups: directional and non-directional. Directional phenotypes were defined as those that could be described on a one-dimensional axis (e.g., higher or lower, robust or gracile, accelerated or delayed). Examples of such phenotypes include *Delayed skeletal maturation* (HP:0002750), *Biparietal narrowing* (HP:0004422), and *Hypoplasia of dental enamel* (HP:0006297). Non-directional phenotypes are phenotypes that could not be described on a one-dimensional scale. Such phenotypes were discarded from our analysis, as changes in the level of gene activity could not be paired with a directional phenotypic change, and thus, a specific phenotypic effect could not be predicted. Examples of such phenotypes include *Abnormality of the forehead* (HP:0000290), *Dental malocclusion* (HP:0000689), and *Lateral clavicle hook* (HP:0000895). This left 815 HPO phenotypes (Table S2). Next, we determined for each of the remaining phenotypes whether the trait is indeed divergent between Neanderthals and MHs, or between chimpanzees and MHs, and whether the HPO phenotype matches the direction in which the trait is divergent. This was achieved for 644 of the 815 phenotypes for the Neanderthal, and 684 for the chimpanzee, by reviewing key sources that surveyed Neanderthal and chimpanzee anatomy (Aiello and Dean, 2002; Been et al., 2012; Chirchir et al., 2015; Clement et al., 2012; Dean et al., 1986; Gilmore and Weaver, 2016; Gómez-Olivencia et al., 2013; De Groote, 2011, 2008; Kupczik and Hublin, 2010; Lieberman, 1998; Lieberman et al., 2000; Maureille and Bar, 1999; Pickering et al., 2007; Raichlen et al., 2011; Reich et al., 2010; Smith et al., 2007, 2010; Trinkaus, 2003; Weaver, 2009; Weber and Pusch, 2008; Zilberman and Smith, 1992). The rest were discarded due to (a) insufficient fossil or morphological information that did not enable us to determine whether the phenotype is divergent (e.g., *Delayed calcaneal ossification*, HP:0008142, and *Triangular nasal tip*, HP:0000451 in the Neanderthal), (b) ambiguity of the definition of the phenotype (e.g., *Skeletal dysplasia*, HP:0002652, and *Chronic pain*, HP:0012532 in both), or (c) contradicting evidence (*Long phalanx of finger*, HP:0006155, and *Hypoplastic pubic rami*, HP:0008830 in the Neanderthal, Table S2).

Next, we examined for each HPO phenotype whether genes associated with it contain promoter DMRs. We determined the lineage in which the DMRs associated with it emerged (e.g., Neanderthal-specific), and the predicted direction of phenotypic change based on the lineage in which hypermethylation was observed (i.e., the lineage in which hypermethylation is observed was associated with the disease phenotype). Phenotypes associated with promoter DMRs were predicted to have changed. However, in order to determine the direction of change, an analysis of gene activity levels is required.

Notably, none of the relevant DMGs showed a significant positive correlation between methylation and gene expression. Because a trait that separates two groups could have arisen on either of their lineages, reconstructing the skeletal profile of a hominin group included also DMGs derived along the other lineages. Specifically, MH-, archaic- and Neanderthal-derived DMGs were used to construct the Neanderthal profile, MH-derived and chimpanzee-human DMGs were used to construct the chimpanzee profile, and MH-, archaic- and Denisovan-derived DMGs were used to construct the Denisovan profile. Additionally, Neanderthal-derived DMGs were used to identify Neanderthal-derived traits that are not shared with Denisovans.

One potential limitation of this approach is that our method is based on the conjecture that even partial, rather than complete, change in the activity of a gene could potentially result in a phenotypic change. This is certainly not always the case, and therefore, cases where a gene activity partially decreased along a lineage, but only complete silencing results in a phenotype, would result in incorrect predictions of divergence. We estimate such instances to be relatively rare. Their extent could be estimated by investigating whether the symptoms of the monogenic diseases used in our study were shown to be driven by homozygous (recessive) or heterozygous (dominant) mutations. In 12 out of the 27 genes used for the Denisovan profile, heterozygous mutations have been shown to underlie the HPO phenotypes. At first glance, this suggests that 15 of the genes present a phenotype only in homozygous loss-of-function states and are therefore unsuitable for inference about their phenotypic effect in partial silencing. However, such a conclusion would be inaccurate because heterozygous mutations are vastly under-represented in genetic screenings: First, many of the studies that have tried to identify the underlying genetic mechanism of a disease specifically searched for stretches of homozygosity, often in consanguineous families, while ignoring heterozygous candidates (Hamosh et al., 2005). Second, many of these mutations do not result in a complete loss-of-function and the gene remains partially expressed and active, even in homozygous states. This suggests that the phenotype is, to some extent, dosage-dependent, and that it manifests even when the gene is still

active. Finally, heterozygous mutations regularly drive more subtle phenotypes which are sometimes classified as "normal variation" or more often - ignored. An example of such a gene in our reconstruction is *PDE6A*. One of the HPO phenotypes associated with this gene is *Wide nasal bridge* (HP:0000431). This phenotype was shown to be driven by homozygous mutations, ostensibly suggesting that the mechanism is recessive. However, nasal morphology GWAS show that a SNP in this gene (rs77409096) is significantly associated with nasal morphology (Lee et al., 2017). An additional example is *ENPP1*, which was linked to the *high palate* (HP:0000218) phenotype only in homozygous individuals, but contains a SNP (rs9373000) which was shown to significantly affect jaw morphology (and specifically its height) in heterozygous as well as homozygous individuals (Constant et al., 2017). This suggests that while the homozygous variants in these genes cause a severe set of phenotypes, heterozygous variants underlie milder phenotypes and are often classified as normal variation. This is further exemplified in the fact that six out of the 15 genes that supposedly drive phenotypes only through recessive inheritance were shown to be associated with skeletal phenotypes through GWAS, including in heterozygous individuals (MacArthur et al., 2017). Six additional genes out of the 15 were shown to have partial activity even in homozygous states (Hamosh et al., 2005). This implies that only three out of the 27 genes used in the Denisovan profile do not manifest in skeletal phenotypes when gene activity is partially altered. We believe that the true number is probably even lower than that due to the limited number of skeletal GWAS.

In summary, instances where only complete (homozygous) loss-of-function of a gene results in a phenotype could potentially cause an over-estimation of divergent traits. This is because a trait would be classified as divergent due to its associated gene being partially silenced, whereas only complete silencing of this gene results in a phenotype. We estimate such instances to be relatively rare, and they probably underlie some (or even many) of the ~15% of cases where we predicted a trait to be divergent but no difference in morphology was reported between the groups.

Of note, the fact that most HPO phenotypes are observed as a result of loss-of-function events suggests that instances where the archaic promoter is hypermethylated would better match patient's phenotypes (compared to instances where the MH promoter is hypermethylated). This is because in instances where the MH promoter is hypermethylated, the phenotypic prediction in our profile is a mirror image of the phenotypic effect in patients, which would not necessarily be correct, whereas in instances where the MH promoter is hypomethylated, the downregulation in patients and archaic humans is similar, and therefore the phenotype is more likely to be so too. Such incorrect predictions are encompassed in our ~15% incorrect predictions of divergence or directionality.

It is also important to test whether these genes are expressed in MH skeletal tissues to begin with, otherwise their hypermethylation in archaic humans is not expected to have a functional effect. To test this, we used four RNA-seq datasets in skeletal tissues: osteoblasts (GEO: GSE57925), skull and long bone (GEO: GSE55282), fetal cartilage (GEO: GSE107649), and adult chondrocytes (GEO: GSE74220). For each gene used in any of the reconstructed profiles, we tested whether it is expressed (> 1 FPKM) in any of the datasets. Three gene used in the Denisovan profile (*BLNK, MYO7A* and *PDE6A*) showed no expression in skeletal tissues. Including or excluding these genes from the analyses does not alter the reconstructed profile as all traits linked to these genes (*high palate* (HP:0000218), *Enlarged vestibular aqueduct* (HP:0011387), *Hypoplasia of the nasal bone* (HP:0004646), and *Wide nasal bridge* (HP:0000431) are affected by additional genes with the same directionality or contained within other traits (Table S9). In the Neanderthal and chimpanzee profiles there are four genes each which are not expressed in skeletal tissues (*MYH11, MYO7A, KCNT1* and *PDE6A* in the Neanderthal profile, and *TNNT3, CNTN1, GJA5*, and *SLC4A1* in the chimpanzee profile). Here too, all associated traits are affected by other genes pointing to the same direction and therefore, removing these genes from the analyses does not affect the results. The fact that almost all genes used in this study, including those with hypermethylation patterns in MHs, are expressed in skeletal tissues is perhaps unsurprising considering that loss-of-function of these genes was shown to have a phenotypic effect on skeletal tissues in patients, suggesting they are active in healthy individuals.

Unidirectionality filtering

Because precise expression levels and the relative contribution of each individual DMR to the phenotype could not be determined, we focused on phenotypes where all methylation changes point to the same direction with no contradiction. This was achieved by a series of three filter, together denoted *unidirectionality* filters (Figure 1, Table S4, Figure S2).

This included unidirectional changes of (a) DMRs within the same gene (i.e., discarding genes with a hypermethylated and a hypomethylated DMR within the same promoter). No such DMGs were found in the Neanderthal analysis, and 29 such DMGs were removed from the chimpanzee analysis. (b) Genes linked to the same phenotype (e.g., if a phenotype was associated with two DMGs, where one is hypermethylated and the other is hypomethylated, this phenotype was discarded). In the Neanderthal, DMGs were linked to 193 unidirectional skeletal HPO phenotypes. On average, each phenotype was associated with 2.24 DMGs, with 98 out of 193 phenotypes associated with one DMG, and a maximum of 17 DMGs associated with the phenotype *Microcephaly* (HP:0000252) (Figure S3). In the chimpanzee, DMGs were linked to 192 unidirectional skeletal HPO phenotypes. On average, each phenotype was associated with 2.48 DMGs, with 105 out of 192 phenotypes associated with one DMG, and a maximum of 21 DMGs associated with the phenotype *Scoliosis* (HP:0002650) (Figure S3, Table S4). (c) Phenotypes linked to the same trait. If two phenotypes refer to the same trait (e.g., finger length), and one points to longer fingers while the other points to shorter fingers, the two phenotypes were discarded. In the Neanderthal, 193 HPO phenotypes passed the 1st filter, all of them passed the 2nd filter and 85 passed the 3rd filter. In the chimpanzee, 175 HPO phenotypes passed the 1st filter, 140 passed the 2nd filter, and 76 passed the 3rd filter.

To minimize dependency between predictions, phenotypes that are contained within other phenotypes were merged into a single morphological unit. For example, *Coronal craniosynostosis* (HP:0004440) and *Lambdoidal craniosynostosis* (HP:0004443) are contained within *Craniosynostosis* (HP:0001363), and therefore, these three phenotypes were counted as one unit. Some phenotypes partially overlap (e.g., *Small face*, HP:0000274 and *Narrow face*, HP:0000275). In such cases, merging was performed as follows: the overlapping part of the phenotypes (i.e., narrow face) was counted as one phenotype, while the non-overlapping part of the phenotype (i.e., short face) was counted as another phenotype. Phenotypes that were linked to contradicting methylation changes (i.e., did not pass the 1st or 2nd filters), were annotated both as an opposite phenotype and a containing phenotype, as their directionality could not be determined. In the Neanderthal, this clustering grouped the 193 phenotypes into 64 clusters of traits that are predicted to have changed, and which we could examine against known Neanderthal morphology. 53 of which are indeed divergent between Neanderthals and MHs, and were thus tested for directionality too. 33 passed all unidirectionality filters, and 29 of which were predicted in the right direction (Figure 2, Table S4). In the chimpanzee reconstruction, this clustering grouped the 192 phenotypes into 42 clusters of traits that are predicted to have changed, and which we could examine against known chimpanzee morphology. 38 of which are indeed divergent between chimpanzees and humans, and were thus tested for directionality too. 22 passed the unidirectionality filters, and 20 of which were predicted in the right direction (Table S4).

Importantly, some HPO phenotypes are defined in absolute terms, whereas some are defined in relative terms. For example, the term *triangular face* (HP:0000325), which refers to the width of the temples compared to the width of the chin, is a term that is defined in relative terms, whereas the term *narrow face* (HP:0000275) is defined in absolute terms. In the case of relative terms, if the reference to which these phenotypes are compared is defined absolutely, then the relative term was treated as absolute as well. For example, we predict a widening of the Neanderthal and Denisovan chin compared to MHs. We also predict that the Denisovan had a more triangular face compared to the Neanderthal. Therefore, it is expected that the Denisovan bitemporal distance would be absolutely longer compared to Neanderthals.

Validation and accuracy estimation

We tested whether some features associated with the way DMRs were defined could potentially introduce a bias. We replaced each DMG along each of the human lineages with its closest downstream gene and ran the entire reconstruction pipeline as described above resulting in a neighbor-based Neanderthal profile. We then matched this profile against known Neanderthal morphology. This resulted in a prediction of 25 traits, 14 of which are known to be divergent, not different to what is expected by chance (0.998x, p = 0.576, Hypergeometric test). We then filtered for unidirectionality, which resulted in 11 directional traits being predicted, 6 of which were predicted in the correct direction. Here too, the results match random expectation (1.092x, p = 0.545, Binomial test). Importantly, this control also predicted considerably fewer traits overall (25 compared to 64 divergent traits, and 11 compared to 33).

Quantifying precision

To examine the precision (*PRE*) of our predictions, TP/(TP+FP), where TP are true positives and FP are false positives, we break the total of N predictions into four components:

1. n_0 - The total number of times that we correctly predicted a divergent phenotype for which the direction of change could not be determined.
2. n_1 - The total number of times that we correctly predicted a divergent phenotype for which direction of change could be determined, and where we also correctly predicted the direction of change.
3. n_2 - The total number of times that we correctly predicted a divergent phenotype for which direction of change could be determined, and where we did not correctly predict the direction of change.
4. n_3 - The total number of times that we predicted a divergent phenotype that had not changed.

These components are non-overlapping, and obey $n_0 + n_1 + n_2 + n_3 = N$. The probability to correctly predict the direction of the phenotypic change is

$$PRE_{direction} = \frac{n_1}{n_1 + n_2}$$

Under the null model, which is simply picking everything by random, $PRE_{direction} = 0.5$. Therefore, the p value of our estimator is

$$\alpha = \sum_{i=n_1}^{n_1+n_2} \binom{n_1+n_2}{i} \left(\frac{1}{2}\right)^{i} \left(\frac{1}{2}\right)^{n_1+n_2-i} = \left(\frac{1}{2}\right)^{n_1+n_2} \sum_{i=n_1}^{n_1+n_2} \binom{n_1+n_2}{i},$$

Similarly, the probability to correctly predict a diverge phenotype is

$$PRE_{divergence} = \frac{n_0 + n_1 + n_2}{N}.$$

In order to compute the null for $PRE_{divergence}$, we need to know N_0 – the total number of directional phenotypes associated with Gene ORGANizer (Gokhman et al., 2017c), and we also need to know K_0 – how many of the N_0 phenotypes had changed in the

Neanderthal. Then, if we just pick by random N phenotypes out of the N_0, the probability that $n_0 + n_1 + n_2$ of them had changed in the Neanderthal is hypergeometric. Thus, the p-value for this parameter would be

$$\alpha = 1 - hygecdf(n_1 + n_2 - 1, N_0, K_0, N).$$

Quantifying sensitivity

To measure the sensitivity (*SEN*) of our method, TP/(TP+FN), where FN are false negatives, we first compiled a list of all known skeletal traits in which Neanderthals and MHs differ, as well as traits in which chimpanzees and MHs differ. We reviewed key comparative anatomy studies that surveyed Neanderthal and chimpanzee morphology (Aiello and Dean, 2002; Been et al., 2012; Chirchir et al., 2015; Clement et al., 2012; Dean et al., 1986; Gilmore and Weaver, 2016; Gómez-Olivencia et al., 2013; De Groote, 2011, 2008; Kupczik and Hublin, 2010; Lieberman, 1998; Lieberman et al., 2000; Maureille and Bar, 1999; Pickering et al., 2007; Raichlen et al., 2011; Reich et al., 2010; Smith et al., 2007, 2010; Trinkaus, 2003; Weaver, 2009; Weber and Pusch, 2008; Zilberman and Smith, 1992), identifying traits in which Neanderthals and/or chimpanzees are found completely outside MH variation, or traits where Neanderthals or chimpanzees are significantly different than MHs, but the distribution of observed measurements partially overlap. Here too, non-directional traits were discarded due to their low information content. For the Neanderthal, the compiled list included 107 phenotypes, 75 of which have at least one equivalent HPO phenotype (4.8 on average) and could thus be further tested. For example, the derived trait of *Taurodontia* in Neanderthals was linked to a single HPO phenotype: *Taurodontia* (HP:0000679), whereas the trait *Rounded and robust rib shafts*, was linked to the following HPO phenotypes: *Broad ribs* (HP:0000885), *Hypoplasia of first ribs* (HP:0006657), *Short ribs* (HP:0000773), *Thickened cortex of long bones* (HP:0000935), *Thickened ribs* (HP:0000900), *Thin ribs* (HP:0000883), *Thoracic hypoplasia* (HP:0005257). Each of the associated HPO phenotypes was also assigned a direction, i.e., whether it matches the direction of the derived trait, or whether it represents the opposite phenotype. For example, *Broad ribs* (HP:0000885) represents an HPO phenotype that matches the direction of the trait in Neanderthals, whereas *Thin ribs* (HP:0000883) represents the opposite direction (Table S5). For the chimpanzee, this list included 201 traits, 83 of which have at least one parallel phenotype on HPO.

For each trait we examined whether at least one of the HPO phenotypes associated with it is linked to a gene containing any of the following evolutionary changes that separate MHs from Neanderthals or from chimpanzees: 1. Fixed protein changes (e.g., nonsynonymous, frameshifts). 2. DMRs. 3. Promoter DMRs. 4. Unidirectional promoter methylation changes (1^{st} + 2^{nd} + 3^{rd} filters). For the first group (fixed protein changes) we took the previously identified 294 nonsynonymous changes and indels across 268 genes, which separate MHs and archaic humans (Prüfer et al., 2014). We discarded changes where the Vindija Neanderthal high-coverage sequence did not match the Altai Neanderthal sequence (Prüfer et al., 2017). The Vindija Neanderthal VCF files were downloaded from the Max Planck Institute for Evolutionary Anthropology website: http://cdna.eva.mpg.de/neandertal/Vindija/VCF/. Along the MH branch, this left 45 nonsynonymous substitutions that are fixed across genomes in the 1000 genome project and do not appear in dbSNP, and 7 such indels within reading frames. Along the archaic human branch, this included 191 nonsynonymous changes, and 17 indels. The higher number of changes along the archaic human lineage is probably a result of the lower sample size, which minimizes the ability to discard low frequency variants. Along the chimpanzee, bonobo and human branches, we found 12,291 nonsynonymous changes in 6,009 genes.

In the Neanderthal, we found that 34 of the 75 traits (45.9%) were linked to a protein sequence change, 71 (94.7%) were linked to a DMR (either variable or non-variable within human population), 70 (93.3%) were linked to a non-variable DMR, and 62 (82.7%) were linked to promoter DMRs. For 46 of these 62 traits we were able to assign a direction, which was correct for 36 (78.3% of predictions, Figure 3B). In the chimpanzee, we found that 66 of the 83 traits (79.5%) were linked to a protein sequence change, 74 (89.2%) were linked to a DMR (either variable or non-variable within human population), 67 (80.7%) were linked to a non-variable DMR, and 51 (61.4%) were linked to promoter DMRs. For 32 of these 51 traits we were able to assign a direction, which was correct for 23 (71.9% of predictions, Figure S1D). To test if these results are statistically significant, we examined for each of the four groups whether they are associated with divergent traits more frequently than expected by chance. We did this by testing two aspects: 1. How many genes are expected to be associated by random with divergent traits. 2. How many divergent traits are expected to be identified by chance. The comparison of DMRs that were detected in skeletal tissues with protein changes that exist across all tissues could potentially introduce a bias when examining skeleton-related phenotypes. While both groups probably contain changes that affect non-skeletal tissues, with some affecting only non-skeletal parts, this is more plausible within protein changes. To minimize this bias, we took in all groups only genes which are linked to skeletal phenotypes on Gene ORGANizer (Gokhman et al., 2017c). In the Neanderthal analysis, this left 36 genes with protein changes, 713 genes with a DMR in their gene body or promoter, 223 genes with gene body or promoter non-variable DMRs, 48 genes with promoter DMRs, and 13 with unidirectional changes. In the chimpanzee analysis, this left 728 genes with protein changes, 709 genes with a DMR in their gene body or promoter, 308 genes with gene body or promoter non-variable DMRs, 68 genes with promoter DMRs, and 9 with unidirectional changes. For each group, we randomly replaced the genes with skeleton-related genes, and examined with how many divergent traits they are associated. We repeated this 10,000 times for each group. When examining protein changes, we found that the number of divergent traits identified is not higher than expected by chance (0.81x and 0.98x compared to expected, $p = 0.91$ and $p = 0.80$, for the Neanderthal and chimpanzee, respectively). DMRs, on the other hand, allow for a slightly higher rate of detection of divergent traits than expected by chance, at least for chimpanzees (1.02x and 1.10x, $p = 0.29$ and $p = 1.5 \times 10^{-2}$). This enrichment increases when looking at

non-variable DMRs (1.15x and 1.11x, $p = 2.0 \times 10^{-3}$ and $p = 0.145$), increases further still when looking at promoter DMRs (1.36x and 1.15x, $p = 7.0 \times 10^{-4}$ and $p = 0.079$), and peaks when looking at unidirectional changes (1.68x and 1.59x, $p = 9.5 \times 10^{-3}$ and $p = 0.04$).

Testing the method on other types of data

Finally, to examine the performance of our method when applied to other types of data, we tested it on two additional datasets. These datasets measured regulatory differences between three human and two chimpanzee induced pluripotent stem cell lines differentiated into cranial neural crest cells (CNCCs, cells that contribute to the formation of the pharyngeal arches and facial skeleton (Prescott et al., 2015)). The first dataset included chromatin immunoprecipitation sequencing (ChIP-seq) of histone modifications and enhancer-related transcription factors. Using these data, Prescott et al. identified 66 chimpanzee and 33 human putative enhancer clusters that exhibit upregulation marks. These regions were associated with 410 and 182 genes, respectively, based on regional overlap. Using these genes, we identified 72 potentially divergent traits, 51 of which (70.8%) are known to be divergent (0.94x, $p = 0.84$). For 23 of the divergent traits we were able to predict a direction of change, which was correct for 13 (56.5%, 1.13x, $p = 0.34$, Table S7). The relatively low performance of this dataset could possibly be attributed to two factors: (1) the regulatory effect of these loci on gene expression is unknown; enrichment of active chromatin marks in these regions is often linked to down- rather than upregulation of the gene associated with them (Prescott et al., 2015); (2) the genes associated with each locus are not necessarily its target genes, as distant regulatory regions often activate genes hundreds of kb away, frequently bypassing other genes (Pennacchio et al., 2013). The second dataset on which we tested our method included 1,039 human and 852 chimpanzee upregulated genes in the same CNCCs (> 2x change, FDR < 0.05, t test). Based on these genes, we identified 40 potentially divergent traits. Of these traits, 35 (87.5%) are indeed known to be divergent between the species (1.16x compared to random expectation, $p = 0.039$, hypergeometric test). For 13 of the divergent traits we were also able to predict a direction of change, which was correct for 10 traits (76.9%, 1.54x, $p = 0.047$, Table S7).

Whereas this second dataset shows significantly better performance, we conjecture that the results for both datasets are still an under-estimation of the predictive power of such data, as they were based on merely three human and two chimpanzee individuals, and thus the identified segregating changes are not necessarily fixed between the species. We tested this by applying our method to a down-sampled list of DMRs, where only five humans (three MHs, a Denisovan and a Neanderthal) and one chimpanzee were used to detect DMRs (Gokhman et al., 2017a), and thus these DMRs are not necessarily fixed between the lineages. Indeed, we found that the ability to detect divergent traits that separate MHs and Neanderthals decreased by 7.4% ($PRE_{divergence} = 75.4\%$, $p = 2.0 \times 10^{-6}$). This also suggests that with a similar number of samples, expression-based analysis provides better predictions of divergent traits than methylation-based analysis (87.5% compared to 75.4%). This is unsurprising given that in our method methylation serves as a proxy for expression. Alternatively, these results could suggest that CNCCs have a higher prediction power of skeletal morphology than bone samples, at least for the skull where many of the divergent traits are observed. Indeed, compared to the bone methylation-based data, the CNCC-based reconstructed phenotypes tend to be more often related to the skull (47.5% in the expression-based data compared to 30.9% in the methylation-based data, $p = 0.031$, Fisher's exact test). This observation supports the notion that regulatory changes in an investigated tissue tend to be associated with phenotypes of the same tissue.

In the above analyses, we have used DMRs which were detected using very stringent criteria of a minimum methylation change of 50% and a minimum span of 50 CpGs (Gokhman et al., 2017a). Next, we tested our prediction accuracy using relaxed criteria. To this end, we re-ran the DMR-detection analyses with thresholds of (1) 33% methylation change and 33 CpGs, and (2) 25% methylation change and 25 CpGs (while keeping all other parameters unchanged). As expected, these new criteria result in more DMRs being detected: 1.4x more DMRs (3860 in total) for criterion 1, and 2.1x more (5957) for criterion 2 (Table S6). The results show a slightly lower prediction power, but nevertheless significantly better than expected by chance. For criterion 1 we predicted 60 divergent traits, of which 46 matched the Neanderthal profile (76.7%, $p = 1.5 \times 10^{-4}$, Hypergeometric test, Table S6). For 22 traits we could predict directionality, which was correct in 18 (81.8%, $p = 4.3 \times 10^{-3}$, binomial test). For criterion 2 we predicted 56 divergent traits, of which 42 were correct (75.0%, $p = 6.5 \times 10^{-4}$). 26 traits passed the unidirectionality filter, of which 19 were correct (73.1%, $p = 0.029$). These analyses demonstrate that even with more relaxed criteria, prediction power remains high.

Comparing the profile to the jawbone

Comparison of the reconstructed Denisovan profile to the reported Denisovan mandible was done using the information provided in Table S2 of (Chen et al., 2019). The authors report a dental arch length of 55.7mm, compared to an average of 54.6 in Neanderthals (Asian and European) and 52.01 in MHs. In our reconstructed profile we predicted the Denisovan dental arch to be longer than both that of Neanderthals and MHs (Figure 4), as was indeed observed in the jawbone sample. For mandibular protrusion, the authors do not provide precise measurements, but report a mental foramen that is located under P4 and low on the body, similarly to Neanderthals and more protruding than MHs (Aiello and Dean, 2002), as predicted by our profile. The authors also report an anterior expansion of the mandibular corpus height compared to its posterior height. We used the numbers provided in Table S2 to compute this ratio: 1.030 for MHs, 1.088 for Neanderthals (shorter in Asian Neanderthals: 1.065), and 1.062 for the Denisovan. Finally, the fourth mandibular trait we predicted was large anterior width of the Denisovan mandible compared to both MHs, but similar to Neanderthals. With regard to anterior mandible width (measured by Bicanine distance) the authors reported a substantially wider mandible in the Denisovan compared to MHs (42.6mm compared to an average of 32.64mm), thus matching our profile, but this feature is also wider compared to Neanderthals (36.47mm on average), whereas we predicted it to be similar.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical methods for DNA methylation reconstruction, DMR-detection and lineage-assignment are described in (Gokhman et al., 2017a). For completeness, we provide here a brief overview of these methods.

DNA methylation reconstruction

We developed an algorithm to identify DMRs between a deamination map and a full methylome reference. This algorithm works in two steps. First, we compute a statistic for each CpG position, measuring the level of differential methylation between the two samples. Second, we used scanning statistics to identify differentially methylated regions.

The statistic is defined as

$$\ell_i^+ = t_i\left[\ln\left(1+\frac{\Delta}{\varphi_i}\right)-\ln\frac{1-\pi(\varphi_i+\Delta)}{1-\pi\varphi_i}\right]+n_i\ln\frac{1-\pi(\varphi_i+\Delta)}{1-\pi\varphi_i}$$

where i is the index of the CpG position, t_i is the number of T's at this position, c_i is the total number of C's at that position, and $n_i = t_i + c_i$; φ_i and ψ_i measure the methylation level at that position in the reference map and the reconstructed map, respectively; π is the deamination rate. This statistics measures a likelihood ratio, assuming that t_i is a result of a binomial process $t_i \sim B(n_i, \pi\psi_i)$, and comparing the null hypothesis $\psi_i = \varphi_i$ to the alternative $\psi_i - \varphi_i \geq \Delta$, where Δ is a pre-specified parameter. An analogous statistic was used to the reverse alternative $\varphi_i - \psi_i \geq \Delta$.

DMR detection

This part is accomplished using three steps. (i) Two-way comparisons. The three main ancient samples that represent the different human groups – the Altai Neanderthal, Denisovan, and Ust'-Ishim – were compared to each other in a pairwise manner. For each pair of samples, two comparisons were carried out – we used raw C→T ratio data for one sample and a reconstructed methylation map for the other, and then reversed this. All comparisons used minimum methylation difference threshold of 50%, and a minimum CpG span of 50 CpGs. (ii) Three-way comparisons. We intersected the pairwise DMRs produced in step (i) to identify DMRs that may be attributed to specific hominins. (iii) FDR filtering. To eliminate DMRs that reflect factors other than evolutionary differences between hominins, such as sliding window size and read depth, we repeated the analyses using simulation that mimicked the post-mortem degradation processes of ancient DNA, and then selected the parameters to obtain FDR < 0.05.

Lineage assignment

Association of methylation changes to specific lineages was performed using the chimpanzee methylation maps as outgroups. Then, to determine whether each DMR represents entire groups, we used a total of 67 AMH, archaic and chimpanzee methylation maps. Each DMR where the methylation level in one group overlapped (even partially) the methylation levels in another group was discarded. The within-group variability of our samples was utilized to ascertain that a DMR along a lineage does not represent a sex-, bone-, age-, technology or disease-specific DMR.

DATA AND CODE AVAILABILITY

The code used to generate the anatomical profile is available upon request from Lead Contact, Liran Carmel (liran.carmel@huji.ac.il) or from David Gokhman (davidgokhman@gmail.com).

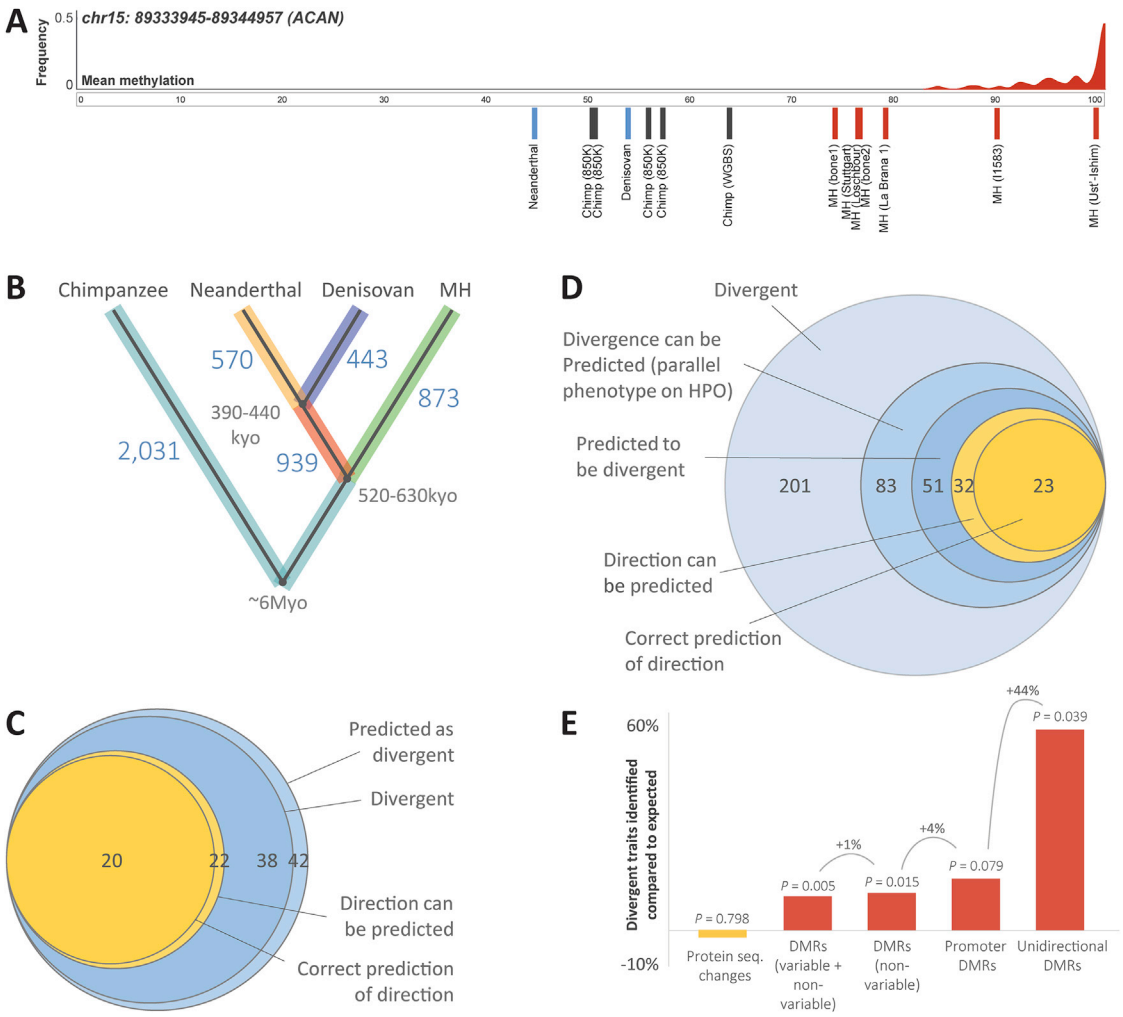


Figure S1. Unidirectional Promoter Methylation Changes Are Predictive of Chimpanzee Anatomy, Related to Figures 1 and 3 and STAR Methods, Validation and Accuracy Estimation

A. An example of methylation levels in a MH-derived DMR in *ACAN* (adapted from (Gokhman et al., 2017a)). MH samples are marked with red lines, archaic human samples are marked with blue lines and chimpanzee samples are marked with gray lines. The distribution of methylation across 52 MH samples (450K methylation arrays) is presented in red. B. The number of DMRs previously identified along each of the hominin branches. C. Out of 42 traits that are known to be derived in chimpanzee-human comparative anatomy, we predicted 38 as derived. Out of the 22 for which a direction of change could be assigned, the direction is predicted correctly in 20. D. The fraction of traits that separate chimpanzees and humans, which could be identified using DNA methylation. Out of 201 known derived traits, only 83 have an equivalent HPO phenotype, 51 of which are identified as divergent. For 32 a prediction of direction could be assigned, 23 of which are predicted in the direction that is observed in the chimpanzee. E. The ability to detect divergent traits was tested for five groups of genes: (i) genes containing protein sequence changes that separate chimpanzees and bonobos from humans, (ii) genes overlapping any chimpanzee-human DMR in their gene body or promoter, (iii) genes overlapping a DMR in their gene body or promoter, for DMRs that show little to no variability within human and chimpanzee population, (iv) genes whose promoter overlaps a DMR with little variability, and (v) genes which passed the three unidirectionality filters. Genes in each of the five groups were randomly replaced with skeleton-related genes, and the number of divergent traits linked to each group was examined. This test was repeated 10,000 times for each of the five groups. In each group, we computed the ratio (*obs* – *exp*) / *exp* for the number of divergent traits that are linked to the genes.

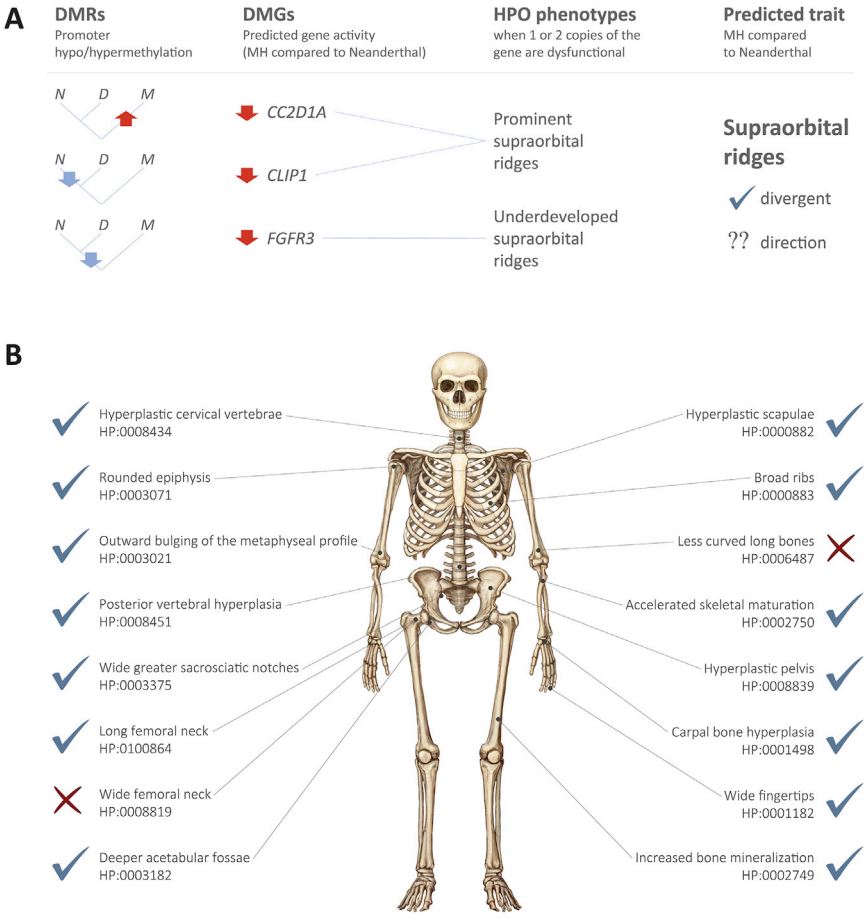


Figure S2. Unidirectional Promoter Methylation Changes Are Predictive of Neanderthal Anatomy, Related to Figure 2 and STAR Methods, Unidirectionality Filtering

A. An example of a trait where promoter methylation changes point to divergence, but the direction of change could not be determined. See Figure 2A for an example of a unidirectional trait. The DMRs column shows promoter DMRs along the different lineages (N – Neanderthal, D – Denisovan, M – MH). Up/down arrows mark hyper/hypomethylation, respectively. The DMGs column shows predicted gene activity change for each of the DMGs. Increased/decreased activity is marked with up/down arrows, respectively. All three genes show patterns of decreased activity in MHs compared to Neanderthals. The HPO phenotypes column shows phenotypes associated with each of the genes. These phenotypes are contradicting, suggesting that suborbital ridges are divergent between MHs and Neanderthals, but the direction of change cannot be determined. B. The 16 non-craniofacial divergent phenotypes for which a direction of change could be assigned. See Figure 2B for skull phenotypes. Whenever overlapping phenotypes were merged, the displayed HPO ID belongs to the most general phenotype. Check marks represent correct predictions (trait is divergent and was predicted in the direction that matches known morphology), X marks predictions where the known Neanderthal phenotype is opposite to the prediction.

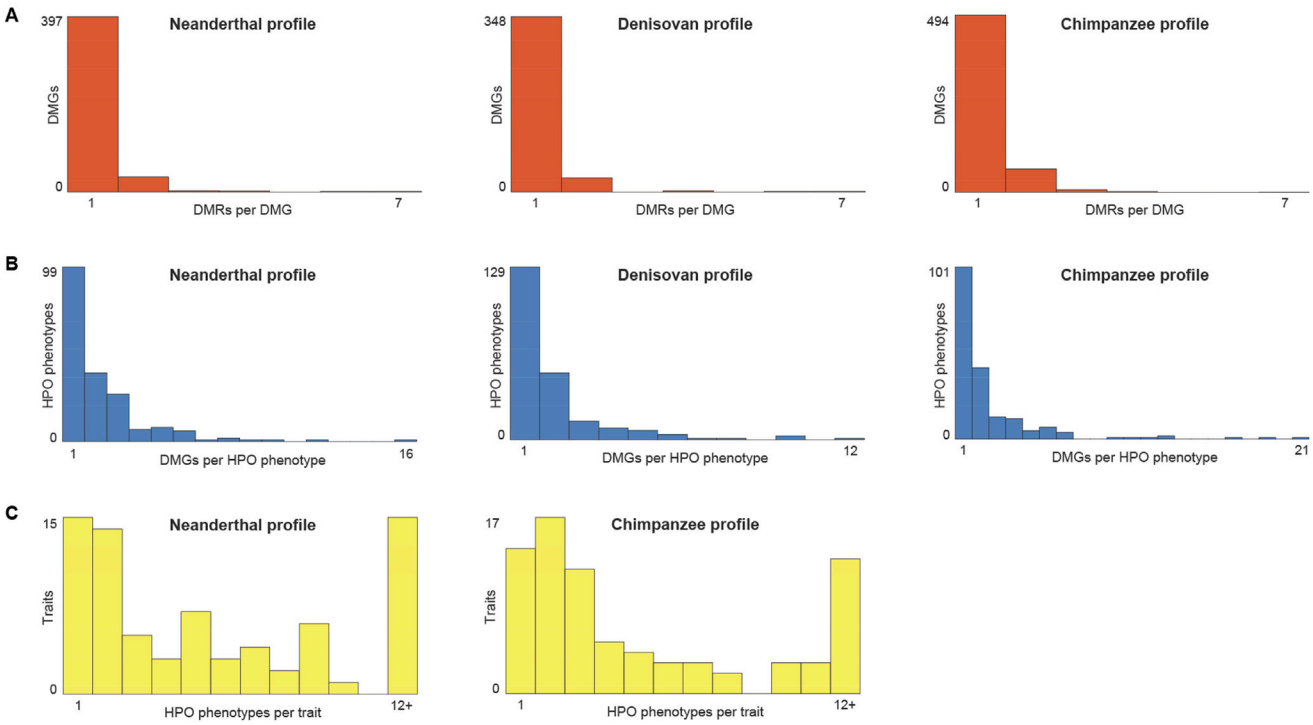


Figure S3. Descriptive Statistics of DMRs, DMGs, and HPO phenotypes, Related to STAR Methods, Unidirectionality Filtering

A. Distribution of DMRs per DMG. B. Distribution of DMGs per HPO phenotype. C. Distribution of HPO phenotypes per traits. All distributions are given for the Neanderthal, Denisovan and chimpanzee reconstructions.