Reviewed Preprint
v1 • April 28, 2026
Not revised

For correspondence:
hbfraser@stanford.edu

Competing interests: No competing interests declared

Funding: See page 22

Reviewing editor: Jenny Tung, Max Planck Institute for Evolutionary Anthropology, Germany

The causes and consequences of human-specific DNA methylation

Zhenzhen Ma[1], Alexander L Starr[1], David Gokhman[2], Hunter B Fraser[1]

[1]Department of Biology, Stanford University, Stanford, United States • [2]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

eLife Assessment

This study presents an **important** examination of the role of cis-acting versus trans-acting genetic variation on DNA methylation divergence between humans and chimpanzees, including its consequences for gene expression. By differentiating fused interspecies tetraploid cell lines into multiple cell types, the study provides **compelling** evidence for the importance of cis-acting changes, but **incomplete** evidence that these changes are of importance for adaptive trait evolution in humans. This work will be of interest to biologists and evolutionary anthropologists studying the evolution and genetics of gene regulation, particularly in primates.

https://doi.org/10.7554/eLife.110575.1.sa3

Abstract

The vast collection of human-specific traits– such as our unique morphology, cognition, behavior, and diseases– has long been attributed to gene expression divergence between us and our closest living relatives, chimpanzees. Theory suggests that changes to *cis*-regulatory elements such as promoters and enhancers may drive evolutionary adaptation, and DNA methylation is a key factor in transcriptional *cis*-regulation. However, we still lack an understanding of 1) how species-specific methylation patterns arise; 2) their downstream effects; and 3) whether they are a common target of natural selection. In this study, we investigated these three questions. By combining a novel hypothesis testing framework with DNA methylation data from six human and chimpanzee cell types, as well as fused interspecies hybrid cells, we disentangled *cis*- vs. *trans*-acting methylation divergence across the genome. Across cell types, we found that methylation divergence is primarily driven in *cis*, which can be linked in some cases to nearby sequence variants such as CpG gains and losses. Although less common, regions with *trans*-acting methylation divergence were enriched for specific transcription factor (TF) binding motifs, suggesting a role of TFs such as FOXM1 in these differences. Having established these causes of methylation divergence, we then examined the functional consequences of differential methylation. Although methylation lacks a consistent relationship with transcription, we observed that associations between methylation and gene expression are stronger for genes with *cis*-regulatory divergence. Moreover, we identified lineage-specific selection shaping promoter methylation at the level of entire pathways including those affecting human-specific traits such as speech, cognition, and susceptibility to infection with hepatitis C. Collectively, our findings provide a mechanistic framework suggesting that DNA methylation may occupy a key position, mediating the effects of both *cis*- and *trans*-acting factors on transcriptional networks, including those contributing to human-specific traits.

Introduction

While gene expression divergence has long been considered the primary driver of human evolution, identifying the molecular mechanisms underlying uniquely human traits remains challenging, particularly when they involve epigenetic modifications that control cell type-specific

expression patterns[1–4]. Theory suggests that cell type-specific *cis*-regulatory variants may fuel evolutionary adaptation through their precise effects on individual genes within specific cell types, avoiding potentially deleterious consequences of broader *trans*-acting changes that affect multiple targets simultaneously[5,6]. However, while recent advances have enabled quantification of human-specific divergence in gene expression and chromatin accessibility, the role of DNA methylation—a key regulator of spatiotemporal gene expression—in coordinating these changes and driving human-specific traits has not been fully explored[6–10]. DNA methylation, occurring primarily at CpG dinucleotides, is generally associated with transcriptional repression and provides stable, long-term regulatory control that has been linked to species-specific phenotypes in mammals[11]. Comparative methylome studies between humans and chimpanzees have revealed widespread tissue-specific methylation divergence, and linked human-specific hypomethylation to neurodevelopmental evolution and disease susceptibility[7,12–23].

Studies of allele-specific expression (ASE) and chromatin accessibility measurements in interspecies hybrids provide a powerful approach for studying regulatory evolution by eliminating confounding factors inherent to cross-species comparisons, including differences in cell type composition, environmental conditions, and developmental timing[24–33]. Analogously, allele-specific methylation (ASM) in human-chimpanzee hybrid cells enables the isolation of *cis*-regulatory effects while controlling for shared *trans*-acting factors. *Cis* regulation of methylation refers to methylation differences caused by factors such as sequence variants on the same allele, typically acting locally. In contrast, *trans* regulation of methylation involves diffusible factors such as DNA methyltransferases, transcription factors, or chromatin modifiers that can influence methylation patterns across many genomic loci (Figure 2A). Studying ASM enables direct quantification of the relative contribution of local versus diffusible factors to species-specific methylation patterns.

While previous studies of primate methylation evolution in post-mortem tissues and blood have provided valuable foundational insights[7,12–16,18–23,34,35], several important questions remain to be addressed. These include: (1) generating comprehensive genome-wide methylation maps across multiple species and cell types; (2) distinguishing between *cis*- versus *trans*-regulatory mechanisms underlying species-specific methylation differences; (3) identifying *cis*-acting genetic variants and *trans*-acting factors underlying these differences; and (4) linking natural selection on methylation to gene expression and candidate traits. Addressing these gaps would enhance our understanding of how genetic variants impact methylation and in turn shape adaptive traits.

Human-chimpanzee tetraploid hybrid induced pluripotent stem cells (iPSCs), generated through *in vitro* cell fusion, provide a controlled experimental system for quantifying *cis*- and *trans*-regulatory contributions to methylation divergence by hosting both species' alleles in an identical *trans*-environment. By measuring ASM and ASE across multiple cell types, this system generates genome-wide maps of cell type-specific regulatory differences that may contribute to evolutionary adaptation. In this study, we generated bisulfite sequencing data from human, chimpanzee, and hybrid iPSCs, cranial neural crest cells, dopaminergic neurons, skeletal myocytes, and hepatocytes, as well as from primary human and chimpanzee dental pulp stem cells. We combined these data with previously published[27] and newly generated RNA-seq data from all six cell types.

Using these data, we investigated the causes of methylation divergence and quantified contributions of *cis*- and *trans*-acting mechanisms that establish interspecies differences in methylation patterns. We developed a hypothesis testing framework tailored to count-based proportional methylation data and extended a changepoint detection algorithm to identify differentially methylated regions (DMRs) under consistent *cis*- or *trans*-regulation. We show that *cis*-acting mechanisms predominantly drive genome-wide methylation divergence, with *cis*-regulated DMRs likely affected by nearby species-specific single-nucleotide variants. In contrast, *trans*-regulated DMRs were enriched for transcription factor binding motifs, including pioneer factors and chromatin remodelers that may mediate *trans*-acting divergence. Concordance between ASM and ASE across cell types revealed species-specific and context-dependent regulatory relationships linking methylation to gene expression. Finally, coordinated methylation-expression divergence in gene sets related to speech, cognition, and viral infection susceptibility

connected lineage-specific selection to candidate phenotypes. In sum, our work provides a framework for understanding how divergent *cis*- and *trans*-acting factors converge on methylation patterns to produce human-specific traits.

Results

Measurement of DNA methylation across diverse cell types

To establish comprehensive methylation landscapes across evolutionarily relevant cell types, we differentiated human, chimpanzee, and their fused tetraploid hybrid iPSCs into four cell types (see Methods)[25,27]. We included an additional cell type, dental pulp stem cells, which are primary cells isolated from human and chimpanzee donors[25]. We then generated whole-genome bisulfite sequencing (WGBS) data from human and chimpanzee parental and hybrid cranial neural crest cells (CNCC) and dopaminergic neurons, as well as parental dental pulp stem cells (DPSC), complemented by reduced representation bisulfite sequencing (RRBS) data from induced pluripotent stem cells (iPSCs), skeletal myocytes (SKM) and hepatocytes (HEP) (Figure 1A). Principal component analysis and hierarchical clustering of methylation patterns showed that samples clustered primarily by cell type, then by species, and finally by parental versus hybrid system (Figure 1B and 1C and Figure 1—figure supplement 1). CNCC samples displayed a unique clustering pattern where parental vs. hybrid samples clustered separately, though examination of the dendrogram indicated this was due to a single human parental sample being a slight outlier (Figure 1B). Known cell type-specific gene expression markers validated the correct identity of each differentiated cell type (Figure 1D and Figure 1—figure supplement 2)[36–50]. Overall, these results are consistent with our expectation that cell type is the predominant determinant of DNA methylation patterns, while species explains a much smaller proportion[7]. In other words, a human hepatocyte is much more similar to a chimpanzee hepatocyte than it is to a human dopaminergic neuron.

Contribution of *cis* and *trans* regulation to DNA methylation divergence between human and chimpanzee

To understand the molecular mechanisms underlying species-specific methylation patterns, we first established a framework for distinguishing between *cis*- and *trans*-regulatory control of DNA methylation. By comparing methylation levels between human and chimpanzee alleles within the same cellular environment (hybrid cells), we can isolate *cis* effects from *trans* effects and determine which mechanism predominates in establishing species-specific methylation patterns (Figure 2A).

To quantify the relative contributions of *cis*- versus *trans*-acting factors in establishing human-chimpanzee methylation differences, we applied a beta-binomial model framework adapted from Hallgrímsdóttir et al. (2024) to classify differentially methylated regions based on allele-specific patterns in hybrid versus parental samples. To determine whether a CpG site reflects a difference in methylation due to *cis* or *trans* regulation, two hypothesis tests are used: one to evaluate the null hypothesis that the difference in methylation is due solely to *trans*, and the other solely to *cis* (see Methods). Each site is categorized into one of five mutually exclusive regulation groups– pure *cis*-driven, pure *trans*-driven, *cis* + *trans*, and *cis* x *trans*, and conserved (Figure 2A). We investigated genome-wide methylation divergence by fitting the model to methylation levels at each CpG site. In all cell types examined, most CpG sites (83-93%) had conserved patterns of methylation. Among divergent sites, *cis*-regulation was over twice as common as *trans*-regulation, while a smaller fraction showed both *cis* and *trans* divergence (Figure 2B, 2C, Figure 2—figure supplement 1 and Supplemental File 1). The predominance of *cis*- over *trans*-regulation was maintained after FDR correction and remained stable across a range of FDR cutoffs (FDR < 0.25, 0.20, 0.10, 0.05, 0.01; Figure 2—figure supplement 1-2 and Supplemental Files 1-2). This genome-wide predominance of *cis*-regulation indicates that local sequence differences between human

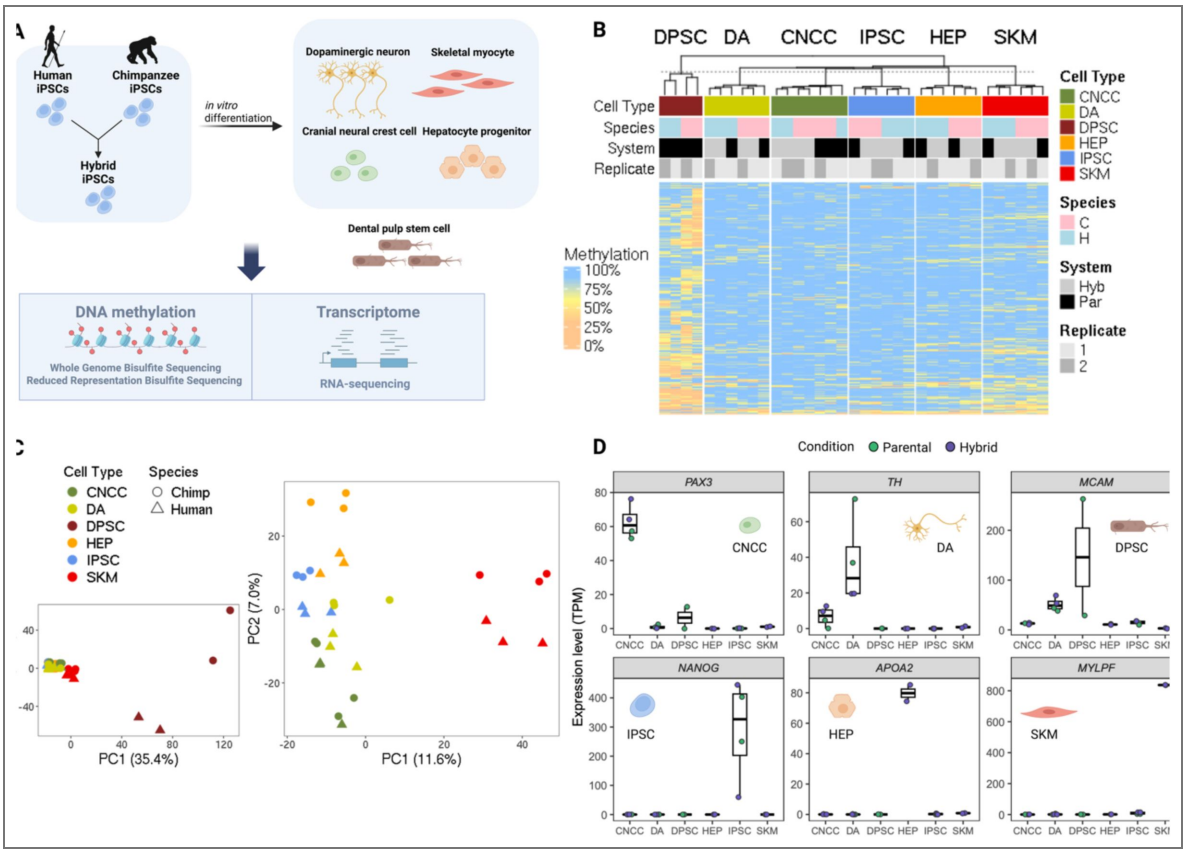

Figure 1. DNA methylation profiles across diverse human, chimpanzee, and hybrid cell types.

(A) Four cell types were differentiated from human, chimpanzee, and hybrid induced pluripotent stem cells. The cell types span diverse body systems including skeletal myocytes for the musculoskeletal system, dopaminergic neurons for the central nervous system, cranial neural crest cells for craniofacial development and hepatocyte progenitors for the liver. We collected DNA methylation and RNA-seq data from these cell types from humans, chimpanzees and their hybrids. Dental pulp stem cells were also collected from human and chimpanzee adult tissues. (B) Clustering analysis and (C) PCA between samples for all cell types (left) and for all iPSC-derived cell types (i.e. excluding DPSC, right). (D) Cell type specific markers of gene expression. Figure 1A was created using *BioRender*, and is published under a CC BY-NC-ND 4.0 license.

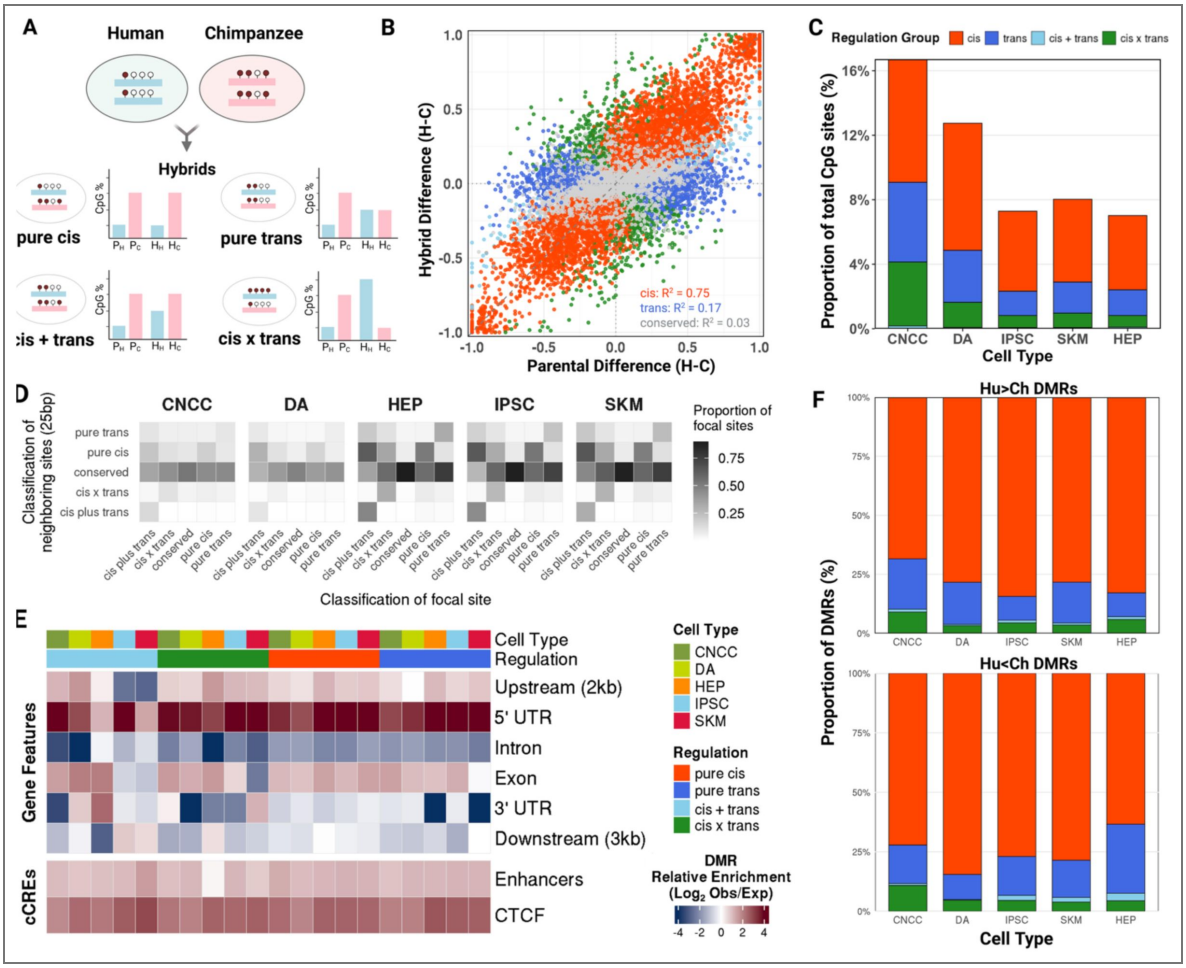

Figure 2. Contribution of *cis* and *trans* regulation to DNA methylation divergence between human and chimpanzee.

(A) Interspecies differences in methylation levels per CpG site is compared between parental and hybrid systems and assigned a regulation category based on hypothesis testing. (B) Parental and hybrid methylation ratios in dopaminergic neurons and the respective regulation category (color key in panel C). (C) Quantification of relative contribution of non-conserved regulation groups to whole-genome CpG sites across cell types. (D) Heterogeneity of the distribution of CpG sites of all regulation groups. (E) Quantification of CpG regulatory clusters by overlap with genomic features of transcripts genome-wide as well as *cis*-regulatory elements including promoters, enhancers and CTCF-bound *cis*-regulatory regions (CTCF). (F) Quantification of relative contribution of non-conserved regulation groups to regulatory clusters across cell types (color key in panel C). Figure 2A was created using *BioRender*, and is published under a CC BY-NC-ND 4.0 license.

and chimpanzee genomes are the primary drivers of methylation divergence, consistent with previous hypotheses of gene regulation that *cis*-regulatory changes may be less disruptive than broad *trans*-acting modifications[1–4,6].

The regional distribution of sites from different regulation groups reveals a consistent pattern across cell types: conserved CpG sites are most likely to be near other conserved sites, while divergent CpG sites are likely to be near conserved sites as well as sites within the same regulation group (Figure 2D). This suggests that CpG sites within the same regulation group tend to form coherent clusters in the genome, and divergent CpG clusters typically have conserved sites scattered within. When conducting regional analyses, methylation data is usually pooled from multiple sites within a pre-defined genomic region[6–10]. However, this regional heterogeneity can affect classification as well as estimates of how much each regulatory mechanism contributes to divergence between species (Figure 2—figure supplement 2-3, Supplemental File 2, see Methods). Therefore, pooling data across regions should be approached carefully, as subsets of CpG sites within the same regulatory element can be controlled by distinct mechanisms. These findings underscore the importance of studying methylation divergence at high resolution.

To identify differentially methylated regions (DMRs) exhibiting consistent inter-species methylation directionality and minimal classification heterogeneity of regulation groups, we implemented a changepoint detection method with a tolerance threshold for regulatory homogeneity (see Methods). This method systematically identifies genomic loci where co-localized CpG sites demonstrate uniform regulatory behavior punctuated by a pre-defined fraction of heterogeneous sites: non-conserved classifications showing consistent directional methylation bias, and conserved sites maintaining equal methylation levels between species. This method prioritizes DMRs that are controlled by the same mechanism and display coordinated methylation patterns (Supplemental File 3), thereby enriching for loci with interpretable regulatory mechanisms. Notably, DMRs identified through this method are enriched in 5' UTRs (just downstream of each transcription start site [TSS]) and CTCF-bound regions, supporting their functional relevance in gene regulation (Figure 2E). The quantification of *cis*- vs. *trans*-DMRs yielded the same conclusion that *cis*-regulation outweighs *trans*-regulation at establishing divergent methylation patterns (Figure 2F).

The causes of *cis*- and *trans*-acting divergence in DNA methylation

To identify the molecular mechanisms driving methylation divergence, we analyzed transcription factors (TFs) associated with *trans*-DMRs and sequence variants associated with *cis*-DMRs.

TFs represent a major class of *trans*-acting factors capable of establishing and maintaining methylation patterns across the genome, often by recruiting methyltransferase or demethylase enzymes. Interspecies variation in TF expression levels, binding activity, or target site accessibility can propagate to downstream methylation differences at their binding sites. To identify TFs potentially driving species-specific methylation divergence through such *trans*-acting mechanisms, we performed TF binding site motif enrichment analysis on *trans*-DMRs (Figure 3A, Supplemental File 4 and 5)[52]. This analysis identified several candidate factors, including a pioneer factor FOXA2 (Benjamini-Hochberg corrected FDR = 3.7×10^{-5} in DA Hu<Ch *trans*-DMRs) and a chromatin modifier FOXM1 (FDR = 0.002, 4.8×10^{-6}, and 0.021 in CNCC, DA, and HEP Hu<Ch *trans*-DMRs, respectively), both known to actively reshape DNA methylation landscapes[53–60].

If species-specific levels of a TF are responsible for the enrichment of its predicted binding sites in *trans*-DMRs, then this enrichment may be strongest in cell types with greater species-specific differential expression of the TF. Consistent with this expectation, higher relative expression of FOXM1 and FOXA2 was associated with lower methylation at their predicted binding sites in that same species and cell type. For example, FOXM1 was over 4-fold more highly expressed in human HEPs than in chimpanzee, and its predicted binding sites were highly enriched specifically for *trans*-DMRs with lower methylation in human HEPs; in contrast, three other cell types with more similar human vs. chimpanzee expression of FOXM1 showed weaker enrichment and little species-specific asymmetry in *trans*-DMR methylation directionality (Figure 3B). We observed a similar pattern for FOXA2 (Figure 3—figure supplement 1).

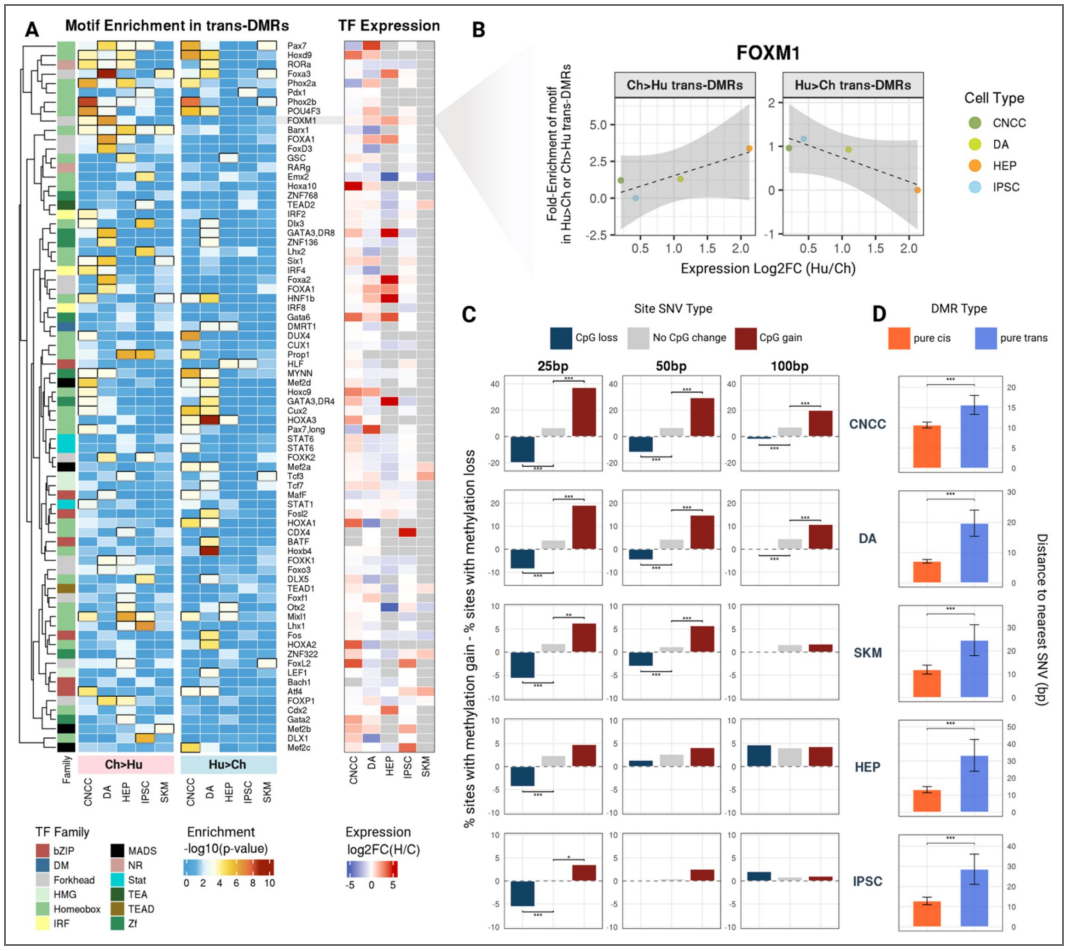

Figure 3. ***Cis*- and *trans*-factors influencing species-specific methylation patterns.**

(A) HOMER motif enrichment of *trans*-DMRs in regions Hu>Ch and Ch>Hu methylated in humans (left panel), and differential expression of corresponding transcription factors in parents (right panel). (B) Fold-enrichments against genomic background of FOXM1, one of the top motifs enriched in Hu>Ch and Ch>Hu *trans*-DMRs, plotted against Hu and Ch relative gene expression levels in parents. (C) CpG-disrupting SNV and their effects on neighboring CpG methylation within ±25, 50, and 100 bp. The differential methylation patterns around CpG-disrupting SNV variants versus non-CpG-disrupting variants are assessed using two-proportion test (***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$) (D) Distances of each *cis*-DMR vs *trans*-DMR to the nearest CpG-disrupting SNV site (Mann-Whitney U test; ***: $p < 0.001$).

Other TFs showed different patterns of association between species-specific expression and *trans*-DMR methylation. For example, FOXP1 and OTX2 were most differentially expressed in HEPs, consistent with their predicted binding sites also being most enriched in *trans*-DMRs in this same cell type (Figure 3—figure supplement 1). However, unlike the pioneer factors described above, binding sites for FOXP1 and OTX2 showed similar patterns of enrichment across cell types for both Hu>Ch and Ch>Hu methylated *trans*-DMRs. This lack of directionality may suggest that these transcription factors are not restricted to either demethylation or methylation, perhaps having a more indirect or context-specific role in regulating the epigenetic landscape[61–63], though further work is required to test this hypothesis. Altogether, our *trans*-DMR analysis nominated many TFs that may contribute to methylation divergence, including several whose *trans*-DMR binding site motif enrichment is associated with the TF's differential expression.

To identify the molecular mechanisms driving *cis*-regulatory methylation differences, we then examined whether local sequence variants could explain *cis*-acting methylation divergence. CpG dinucleotides are the primary targets of DNA methylation in mammalian genomes, and their presence, density, and spacing are known to influence local methylation patterns[64–66]. Therefore, single nucleotide variants (SNVs) that either create or disrupt CpG dinucleotides represent a plausible mechanism by which sequence divergence could drive species-specific methylation differences. To investigate this hypothesis, we systematically analyzed SNVs and their associated methylation differences in neighboring CpGs within windows of 25, 50, and 100 base pairs upstream and downstream of each SNV. Our analysis revealed that SNVs affecting CpG dinucleotides—but not other SNVs—are associated with altered methylation patterns of nearby CpG sites up to 50 base pairs from the variant site (Figure 3C). Notably, the effect diminished with increasing distance from the SNV, with the strongest methylation changes observed within 25 base pairs and progressively weaker effects at 50 base pairs. Beyond 50 base pairs, the magnitude of influence of CpG-disrupting variants on neighboring methylation patterns became negligible, suggesting a limited spatial range for these local sequence effects. This distance-dependent relationship indicates that sequence variants affecting CpG density and spacing might influence the local recruitment or activity of DNA methyltransferases or demethylases.

This model predicts that while *cis*-DMRs should often be close to CpG-disrupting SNVs, *trans*-DMRs—which arise from regulatory mechanisms independent of nearby sequence variation—should not be. To test this, we compared the genomic distance between each DMR and its nearest CpG-disrupting SNV. Consistent with our model, we found consistently shorter distances for *cis*-DMRs than *trans*-DMRs across all cell types examined (Mann-Whitney $p < 0.001$ for all comparisons)[67]. This difference suggests that *cis*-DMRs tend to occur near sequence variants and are likely directly influenced by local changes in CpG content, while *trans*-DMRs show no preferential association with CpG-disrupting variants. This genomic organization is consistent with *trans*-DMRs being regulated by diffusible factors such as transcription factors rather than by local sequence variants, supporting a model in which *cis*-and *trans*-acting mechanisms operate through fundamentally different molecular pathways to generate species-specific methylation patterns.

Cell type- and species-specific gene expression as a consequence of methylation divergence

Given the well-established repressive effect of promoter methylation on gene expression, we assessed whether methylation differences contribute to expression divergence by analyzing genome-wide correlations between allele-specific promoter methylation and expression levels across different genomic contexts (Figure 4A)[68,69]. We first examined whether *cis*-acting variants drive promoter methylation consistently across cell types or in a context-dependent manner. We found that while some promoters are allele-specifically methylated across all cell types, many more showed cell type-specific *cis*-methylation patterns, suggesting that promoter methylation regulation is often context-dependent (Figure 4B and 4C). However, while DMRs consistently detected across cell types provide strong evidence of shared regulation, DMRs

detected in only one cell type may reflect either genuine cell type-specificity or insufficient power to detect them elsewhere. Our estimate of cell type-specific regulation therefore represents an upper bound on the degree of cell type-specific methylation regulation.

Several specific examples illustrate different patterns of methylation-expression coordination (Supplemental Files 5 and 6; see Methods). The promoter of *CTSF* is a cell-type agnostic *cis*-DMR with consistent Hu>Ch methylation across cell types corresponding to coordinated Ch>Hu expression (Figure 4D and Figure 4—figure supplement 1). In contrast, a putative enhancer in *LGALS8* represents a case where methylation divergence is confined to specific cellular contexts. The candidate enhancer overlaps with a *cis*-DMR with Hu>Ch methylation and *LGALS8* exhibits allele-specific Ch>Hu expression exclusively in CNCCs, while having no divergence in methylation or expression in other cell types (Figure 4E and Figure 4—figure supplement 2). A third type of pattern is exemplified by the *HOXA9* promoter, with cell type-specific hypomethylation corresponding to cell type-specific expression only in SKM (Figure 4F). These examples illustrate how different regulatory mechanisms can produce distinct patterns of methylation-expression coordination and how epigenetic changes may contribute to both species-specific and cell type-specific regulatory evolution.

Previous studies have reported weak but significant correlations between DNA methylation and gene expression divergence in primates[7,12,14,16,70]. Consistent with these findings, our genome-wide estimates of human–chimpanzee transcriptome divergence produced similar weak but significant correlations with methylation divergence (Figure 4—figure supplement 3 and Supplemental File 6). Our comparison of CpG island (CGI) vs. non-CGI promoters also yielded trends consistent with past studies showing that CGI promoter methylation correlates more strongly with gene expression repression than does non-CGI promoter methylation[71–73] (Figure 4—figure supplement 4).

However, such genome-wide correlations may mask stronger relationships that occur within specific regulatory contexts where methylation and expression are more directly coupled. We hypothesized that our hybrid system could help identify genes exhibiting stronger methylation–expression coupling by stratifying them based on regulatory mechanisms. We reasoned that genes whose methylation and expression are both influenced by *cis*-acting variants— potentially driven by the same underlying genetic differences— should display tighter coupling than genes regulated through independent mechanisms.

To test this, we evaluated the extent to which methylation and expression show a negative relationship (i.e., higher methylation associated with lower expression) consistent with a repressive effect of methylation. In both CNCC and DA hybrid cells, we found that genes with *cis*-regulated expression were more likely to exhibit repressive patterns than those with *trans*-regulated expression (Figure 4—figure supplement 5). This is consistent with DNA methylation acting as an upstream *cis*-regulatory mechanism: genes with stronger *cis*-regulation are expected to show tighter methylation-expression coupling due to the greater contribution of local regulatory factors, including methylation itself. Further restricting the analysis to genes with concordant regulatory mechanisms revealed that genes with both methylation and expression being *cis*-regulated ("*cis*–*cis*" genes) exhibited significantly stronger negative methylation–expression correlations than "*trans*–*trans*" genes (Figure 4G, see Supplemental File 6 for detailed statistics). This supports the hypothesis that methylation–expression coupling is strongest when both are driven by local genetic variation, perhaps through shared causal variants.

The sign test identifies lineage-specific selection on DNA methylation between human and chimpanzee

To identify pathways that have evolved under lineage-specific selection, we used a two-step binomial sign test approach that tests for non-random directional bias in methylation-expression patterns in hybrids[1,24,25,27,74–78]. The use of hybrid cell lines is critical for this analysis as it ensures that different promoters represent independent observations, controlling for *trans*-acting factors and environmental variation that could otherwise confound the detection of *cis*-regulatory

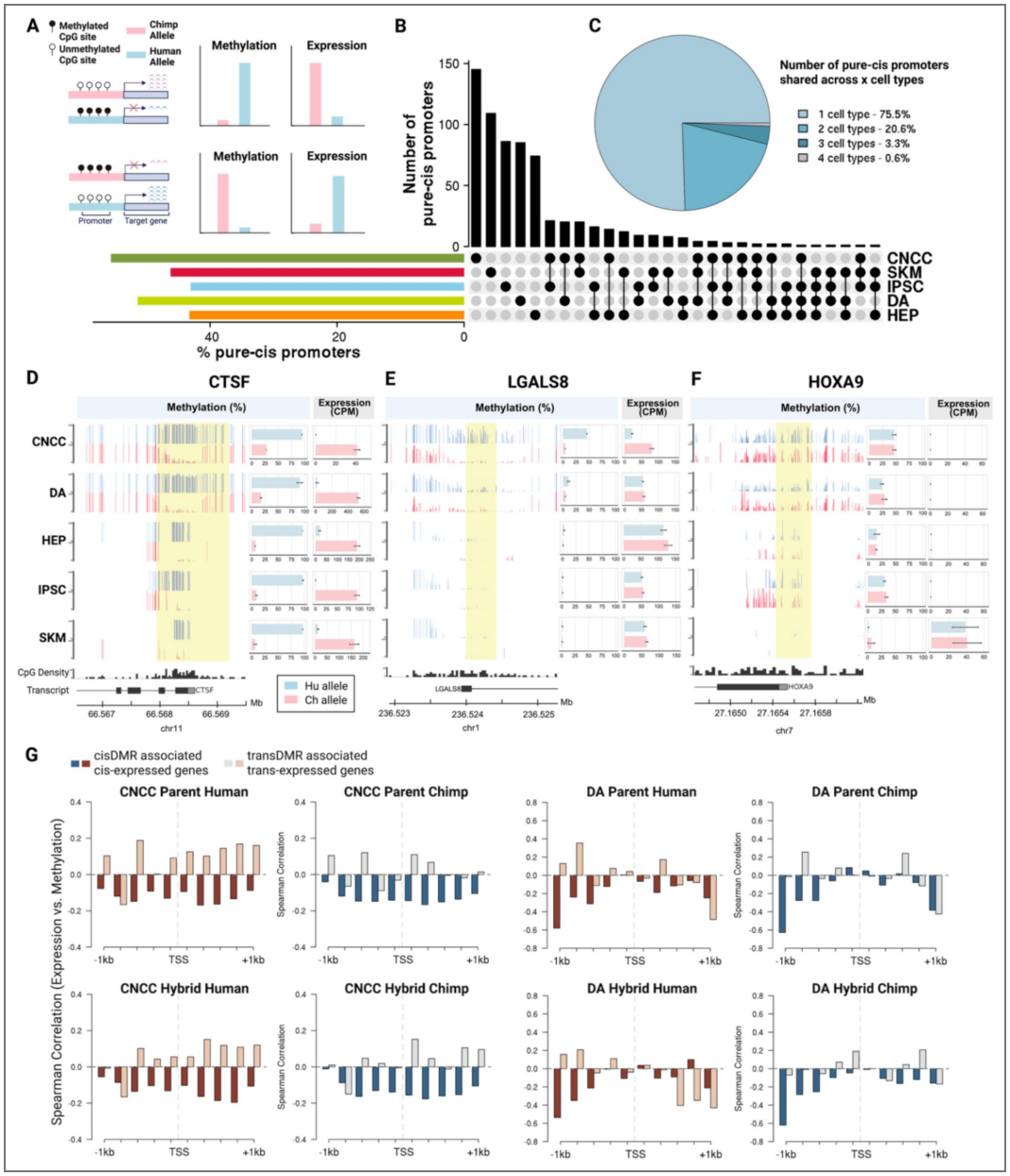

Figure 4. Human-chimpanzee allele-specific methylation can be associated with allele-specific expression.

(A) Illustration of the transcriptional consequences of DNA methylation explored in this study. DNA methylation, as a stable *cis*-acting regulatory mark, often represses transcription. In a hybrid, allele-specific methylation can lead to allele-specific gene expression, although both methylation and its effects on transcription can be context-dependent. (B) We identified many promoters with allele-specific methylation in each cell type, most of which showed allele-specific methylation in only one cell type. (C) Quantification of promoters with pure-*cis* regulation of methylation across cell types. (D, E, F) Examples of gene expression-methylation correspondence. Methylation tracks show fractional methylation (mC/coverage × 100%) of individual CpGs. Expression values are in TMM (Trimmed Mean of M-values)-normalized CPM (counts per million). (D) *CTSF* as an example of species-specific, cell type-agnostic DMR, (E) *LGALS8* as an example of cell type-specific allele-specific methylation, and (F) *HOXA9* as an example of cell type-specific but species-agnostic methylation patterns. (G) Genome-wide profile of expression-methylation correlations (binned by distance of methylated region to the TSS) for genes with both promoter methylation and expression under *cis*-regulation (dark red and dark blue) or *trans*-regulation (light brown and light gray). Detailed statistics are provided in Supplemental File 6. Figure 4A was created using *BioRender*, and is published under a CC BY-NC-ND 4.0 license.

effects. This test operates on the logic that neutral evolution should act like flipping a fair coin: if human-biased and chimpanzee-biased changes occur at equal background frequencies, then methylation and expression changes of genes within a pathway would accumulate randomly, showing no preference for human-biased versus chimpanzee-biased directionality. In contrast, lineage-specific selection would bias the coin flip, potentially causing functionally related gene sets to consistently favor one direction over the other rather than splitting evenly by chance.

To maximize our ability to find phenotypically impactful divergence, we adapted the standard sign test to account for directionality of both methylation and gene expression. In the first step, we applied a standard binomial sign test to all promoters showing ASM, counting the number of genes showing human-biased versus chimpanzee-biased methylation changes within each gene set. The background probability for this test was defined as the overall fraction of promoter ASM in each cell type that was human-biased. This initial analysis identified gene sets showing significant enrichment for directional methylation changes, establishing candidates for lineage-specific regulatory divergence. In the second step, we restricted our attention to genes showing a negative association between ASM and ASE. This subset represents cases where methylation changes are most likely to have a simple and direct repressive effect on transcription, filtering out more complex regulatory scenarios where methylation and expression changes may be uncoupled or subject to competing regulatory influences.

Gene sets showing significant deviation from neutral expectation in the ASM/ASE concordance test (nominal binomial p-value < 0.05 and FDR < 0.25 after multiple test correction; see Methods), with supporting evidence of directional bias in the methylation-only sign test (nominal binomial p-value < 0.05), are interpreted as likely targets of lineage-specific selection on promoter methylation that could directly impact gene expression, establishing a potential mechanistic connection to phenotype (Figure 5, Supplemental File 7). Compared to a standard sign test that focuses solely on directionally biased gene expression, this more stringent two-step filtering approach may identify higher-confidence candidates where regulatory divergence is reflected at both the methylation and expression levels[74].

Among gene sets from Kyoto Encyclopedia of Genes and Genomes (KEGG) and Human Phenotype Ontology (HPO)[79,80], we identified several with clear relevance to human-specific traits. In iPSC hybrids, multiple gene sets including “poor speech,” “highly arched eyebrow,” and “growth delay” all showed Hu>Ch methylation and Ch>Hu expression (Figure 5, Figure 6A and Figure 6B). In the SKM hybrids, genes related to “inability to walk” showed significant Hu>Ch methylation and Hu<Ch expression (Figure 5). In DA hybrids, genes involved in “intellectual disability” show significant Hu<Ch methylation and Hu>Ch expression (Figure 5 and Figure 6C). In HEP hybrids, we identified the “Hepatitis C” gene set with significant Hu>Ch methylation and Hu<Ch expression (Figure 5 and Figure 6D). In primary DPSC samples, we identified significant bias toward Hu<Ch methylation and Hu>Ch expression in genes related to tooth eruption timing, consistent with known differences in dental development between humans and chimpanzees, where human permanent molars emerge approximately 6 years apart compared to 3-4 years in chimpanzees (Figure 5—figure supplement 1)[81–83]. However, the DPSC sign test results should not be interpreted as evidence of lineage-specific selection due to the use of parental rather than hybrid cell lines, which may introduce dependencies among observations due to *trans*-acting effects, potentially inflating statistical significance.

We then analyzed individual genes within each candidate gene set to assess whether expression bias in associated phenotypes aligns with known human-chimpanzee trait differences, focusing on genes showing repressive methylation. Multiple gene sets demonstrated directional bias in gene expression consistent with the direction of phenotypic divergence found between humans and chimpanzees. For instance, in the “highly arched eyebrow” gene set—where most genes show human down-regulation in iPSC hybrids—*COLEC11* (which has 8-fold lower expression from the human allele) stands out as an outlier. *COLEC11* acts as a guidance cue in directing neural crest cell migration during early development, and in humans its loss-of-function (LOF) causes craniofacial dysmorphism including highly arched eyebrows and underdeveloped brow ridge, both of which are traits unique to humans (Figure 5 and Figure 6A)[84]. In the “growth delay”

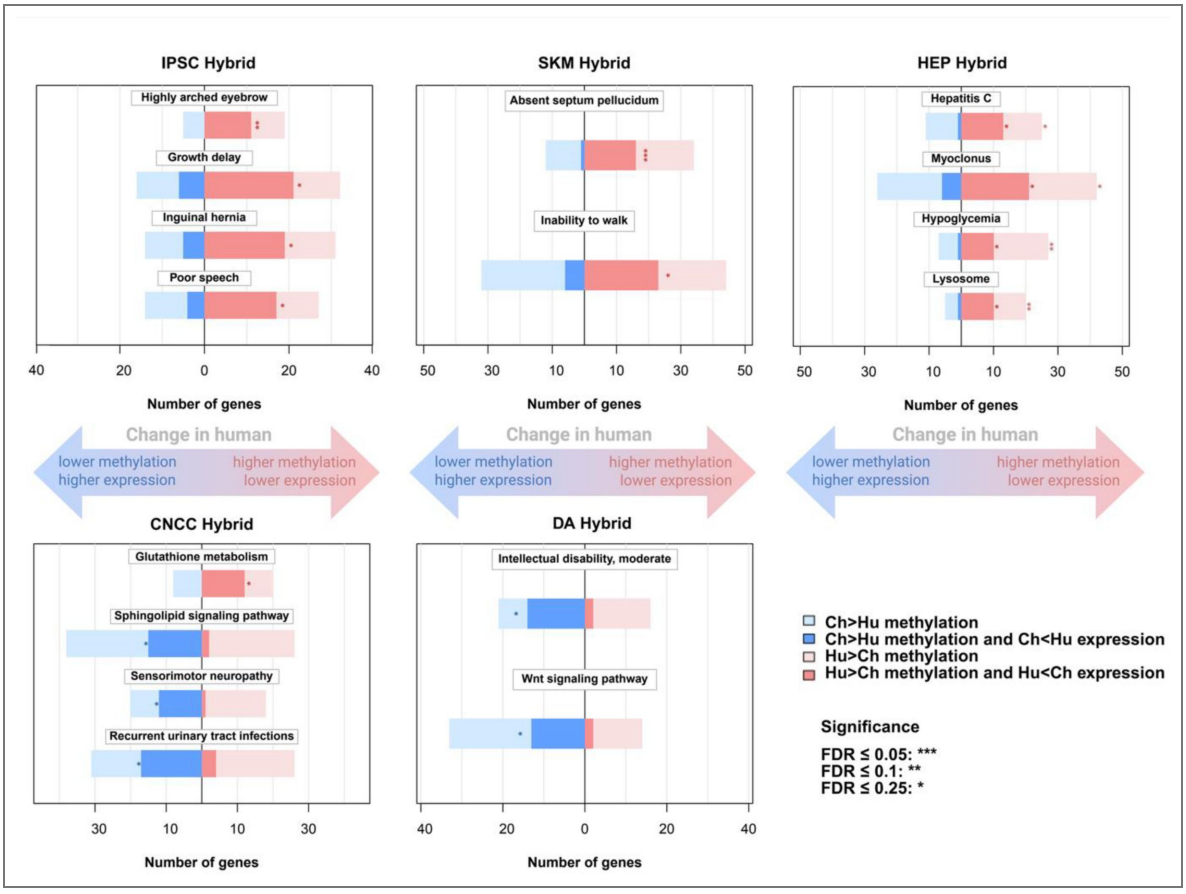

Figure 5. Gene sets with evidence of lineage-specific selection on methylation and gene expression.

The length of the bars indicates the number of genes in each gene set with allele-specific expression or methylation in each direction (human or chimpanzee-biased). The bars in darker colors represent the number of genes where higher methylation is associated with repressed gene expression, whereas lighter colors represent genes where higher methylation is associated with increased gene expression.

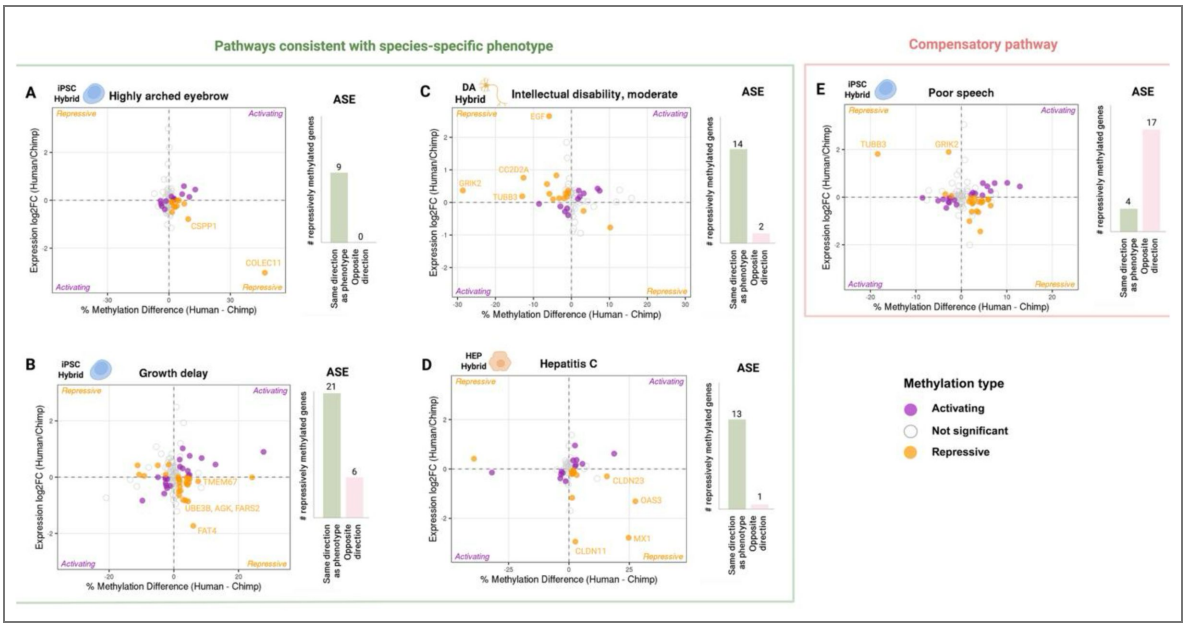

Figure 6. Directional bias in allele-specific methylation and gene expression identifies candidate genes for human-specific phenotypes.

Allele-specific expression (ASE) is plotted against allele-specific methylation (ASM). Orange dots represent genes consistent with repressive methylation (e.g. Hu>Ch ASE and Hu<Ch ASM) and purple dots represent genes consistent with activating methylation (e.g. Hu>Ch ASE and Hu>Ch ASM). Bar plots represent counts of genes with repressive methylation that exhibit expression changes in directions either consistent with (light green) or opposite to (light pink) what would be expected based on phenotypic differences between humans and chimpanzees. A) The "highly arched eyebrow" gene set in iPSC hybrids shows consistent Hu<Ch ASE and Hu>Ch ASM for genes with repressive methylation. B) The "growth delay" gene set in iPSC hybrids shows significant bias towards Hu<Ch ASE and Hu>Ch ASM for genes with repressive methylation. C) The "intellectual disability, moderate" gene set in DA shows consistent Hu>Ch ASE and Hu<Ch ASM for genes with repressive methylation, with genes including *GRIK2*, *TUBB3*, *EGF* and *CC2D2A* being among the genes with highest magnitude ASE and ASM. D) The "Hepatitis C" gene set in HEP shows consistent Hu<Ch ASE and Hu>Ch ASM for genes with repressive methylation. D) The "poor speech" gene set in iPSC hybrids represents a potentially compensatory pathway where most genes with repressive methylation show Hu<Ch ASE and Hu>Ch ASM, whereas the genes with highest magnitude of ASE and ASM (*TUBB3* and *GRIK2*) show Hu>Ch ASE and Hu<Ch ASM.

pathway, most genes with repressive methylation (whose LOF in humans and other species leads to delayed growth)[62–65] show lower expression from their human alleles, consistent with humans' substantially delayed growth pattern compared to chimpanzees (Figure 6B)[85,86]. Many of the genes in DA whose LOF causes "intellectual disability" may have contributed to the evolution of human cognitive abilities. For example, the high-magnitude genes with Hu>Ch ASE and Hu<Ch ASM such as *GRIK2*, *TUBB3*, *EGF* and *CC2D2A* are essential for excitatory synaptic plasticity and long-range axon wiring (Figure 6C)[87–91]. In addition, most genes whose inhibition or LOF increases susceptibility or severity of Hepatitis C (e.g. *MX1* and *OAS3*) show lower expression from human alleles in HEP, consistent with human's heightened vulnerability to Hepatitis C viral infections (Figure 6D)[92–94].

Interestingly, we also discovered gene sets where the direction of most gene expression differences predicted the opposite of a known human/chimpanzee phenotypic difference. For example, in the "poor speech" gene set, where loss-of-function is associated with impaired speech in humans, we counterintuitively observed that most genes were down-regulated in humans compared to chimpanzees. However, the two genes with the largest fold-changes, *TUBB3* and *GRIK2*, showed higher expression from the human allele (Figure 6E). Notably, these two genes are also among the large-magnitude genes in the predominantly human up-regulated "intellectual disability" pathway, which does align with species-specific phenotype (Figure 6C). This suggests the speculative possibility of compensatory evolution in which large, pleiotropic effects of a few genes are counterbalanced by smaller changes in many genes (Figure 6C and Figure 6E; also see Discussion). Our previous study found that genes affecting voice box show particularly extensive down-regulation in humans, consistent with our observation in the "poor speech" pathway[25].

Collectively, these results provide the first evidence for lineage-specific selection on human-specific DNA methylation. Our use of hybrid cell lines was critical to this discovery, as it allowed us to isolate *cis*-acting regulatory changes from confounding *trans*-acting and environmental effects. This approach reveals that polygenic selection has acted on entire sets of functionally related genes whose expression is modulated through DNA methylation. Importantly, by connecting lineage-specific methylation divergence to expression changes within functionally coherent pathways, we can generate plausible hypotheses linking regulatory evolution to human-specific phenotypes. These include craniofacial morphology, dental development, viral infection susceptibility, and neurodevelopmental characteristics underlying speech and cognition (Figures 5 and 6, Figure 5—figure supplement 1).

Discussion

Our investigation using human-chimpanzee hybrid tetraploid cells revealed that *cis*-regulatory mechanisms are the primary drivers of genome-wide methylation divergence between humans and chimpanzees across multiple cell types, including dopaminergic neurons, skeletal muscle, iPSCs, hepatocytes, and cranial neural crest cells. This finding aligns with previous studies highlighting the predominant role of *cis*-regulatory elements in evolutionary adaptations, as they enable precise, context-specific gene regulation while minimizing the potentially deleterious pleiotropic effects of *trans*-acting factors[6]. We demonstrated that most differentially methylated sites arise through *cis*-regulatory mechanisms, with *trans*-regulation and *cis-trans* interactions playing secondary roles. We identified clusters of CpG sites with consistent regulatory mechanisms, forming coherent differentially methylated regions (DMRs). These DMRs predominantly overlap with functionally relevant genomic features including 5' UTRs and upstream promoter regions, underscoring their regulatory significance.

Our mechanistic analysis revealed distinct pathways through which *cis*- and *trans*-regulatory factors establish species-specific methylation patterns. For *cis*-regulation, we identified CpG-altering sequence variants as a direct mechanism linking genomic divergence to epigenetic differences. These variants influence methylation patterns of neighboring CpG sites with effects diminishing with genomic distance. Importantly, *cis*-DMR clusters are significantly closer to CpG-

disrupting SNVs than *trans*-DMRs across all cell types, providing strong evidence that local sequence variation directly shapes *cis*-regulatory methylation patterns, perhaps through altered CpG density affecting recruitment of DNA methyltransferases or demethylases.

While we identified substantial numbers of allele-specifically methylated promoters in all cell types, the majority showed cell type-specific patterns. Our examples illustrate this diversity: the *CTSF* promoter exhibits consistent cross-cell-type *cis*-methylation divergence (Figure 4—figure supplement 1) with corresponding expression changes; an enhancer near *LGALS8* shows cell type-specific allele-specific methylation and overlaps with a homogeneous *cis*-DMR exclusively in cranial neural crest cells; and the *HOXA9* promoter displays cell type-specific methylation patterns with corresponding context-dependent expression effects. These findings highlight how the broader regulatory context of each cell type impacts the effects of *cis*-regulatory variants.

Our genome-wide correlation analyses confirmed the well-known pattern that while promoter methylation generally shows statistically significant inverse relationships with gene expression, these correlations are weak when examined across all genes. However, by leveraging our hybrid cell system to stratify genes based on regulatory mechanism, we uncovered significantly stronger negative correlations between methylation and expression when both are *cis*-regulated compared to genome-wide estimates or *trans*-regulated genes. This finding suggests that the repressive effects of species-specific methylation are strongest when both layers of regulation are driven by the same or closely linked genetic determinants.

Our two-step binomial sign test approach in hybrids identified multiple gene sets showing coordinated directional shifts in methylation and expression that significantly exceed neutral expectations. Many of the pathways we identified are related to known human-chimpanzee phenotypic differences: intellectual disability in hybrid DAs, growth delay and highly arched eyebrow in hybrid iPSCs, hepatitis C in hybrid HEPs, and tooth eruption timing in parental DPSCs (Figure 5, 6, Figure 5—figure supplement 1). These findings represent the first evidence for lineage-specific selection acting on human-specific DNA methylation. By restricting analysis to genes showing negative methylation-expression associations—where methylation most likely exerts direct transcriptional repression—we identified high-confidence candidates where epigenetic divergence has likely functional consequences. Our results suggest that *cis*-regulatory changes in DNA methylation represent an important substrate for adaptive evolution, complementing protein-coding and other *cis*-regulatory changes in generating human-specific traits.

We developed several methodological innovations tailored to the unique challenges of comparative epigenomics. First, our beta-binomial framework accounts for the count-based nature of bisulfite sequencing data and explicitly models allele-specific patterns in hybrid versus parental contexts. Second, our changepoint detection algorithm identifies coherent DMRs with minimal regulatory heterogeneity, prioritizing functionally organized regulatory units over isolated CpG sites. Third, the two-step sign test we introduce provides a method to identify lineage-specific selection that links methylation and corresponding changes in gene expression. These methods can be readily applied to other taxa where hybrid organisms can be generated.

However, our study also has important limitations. First, while hybrid cells provide powerful controls for *trans*-acting variation, in vitro differentiation protocols do not fully recapitulate the complexity and temporal dynamics of in vivo development. Therefore, cell lines are limited in their ability to capture species-specific methylation or expression profiles that are specific to a certain developmental stage or in vivo context. Second, pathway analysis of DPSCs used primary human and chimpanzee samples rather than hybrids, with *trans*-acting factors potentially inflating statistical significance in the sign test; these results should therefore be interpreted not as evidence for positive selection, but rather as an interesting pattern that could have important connections to phenotypic differences between humans and chimps. Third, while we identified strong associations between transcription factor binding motifs and *trans*-DMRs, definitive demonstration of causality requires targeted experimental perturbations. Fourth, our study examined a limited number of cell types, and methylation divergence in most tissues remains unexplored.

Comparative gene expression analyses between humans and closely related species offer a powerful way to explore the molecular basis of species-specific traits and disease susceptibilities. However, it remains challenging to infer phenotypic differences through the lens of human clinical ontologies. To start with, HPO terms are largely derived from pathological human data and predominantly capture phenotypic consequences of loss-of-function mutations in a clinical context. When comparing expression levels between species, one might assume that lower expression corresponds to a loss-of-function-like phenotype. However, the relationship between expression dosage and function is typically nonlinear[95], and complete gene knockouts in a model organism or loss of function mutations in patients are not equivalent to moderate downregulation. This fact may help explain cases where our sign test results suggest the opposite directionality as compared to human/chimpanzee phenotypic differences (Figure 6E). Alternatively, humans and chimpanzees may achieve a similar phenotype through distinct configurations of regulatory networks. For example, an adaptive down-regulation of a particular gene may be compensated for by up-regulation of other genes in the same pathway, maintaining overall pathway function. In other words, compensatory changes may complicate the relationship between gene expression and phenotype, as may be the case for the “poor speech” gene set (Figure 6E)[24].

Several important directions emerge from this work. First, extending this framework to additional cell types and developmental time points would reveal whether the predominance of *cis*-regulation we observe is universal or varies across developmental contexts. Longitudinal sampling during differentiation could identify critical windows when methylation patterns are established and reveal whether human-chimpanzee differences emerge gradually or through discrete developmental transitions. Second, integrating chromatin accessibility (ATAC-seq) and histone modification (ChIP-seq) data would provide a more complete picture of the regulatory landscape, clarifying how DNA methylation interacts with other chromatin features to establish gene expression patterns. Third, direct examination of transcription factor binding using techniques like CUT&RUN in both species could validate the predicted relationships between transcription factor occupancy and methylation patterns at *trans*-DMRs. Finally, extending these analyses to three-way comparisons including other great apes would provide phylogenetic resolution to distinguish human-specific changes from chimpanzee-derived changes.

Collectively, our findings provide a framework for understanding how DNA methylation divergence contributes to human-specific traits. We demonstrate that while both *cis*- and *trans*-regulatory mechanisms shape interspecies methylation differences, *cis*-acting factors predominate, thus directly linking genomic sequence variation and epigenetic divergence. The coordinated nature of many independent methylation-expression changes across phenotypically relevant pathways suggests that epigenetic modifications represent important targets of natural selection in human evolution, contributing to the regulatory innovations underlying uniquely human phenotypes.

Materials and Methods

Cell culture and differentiation

We utilized two previously characterized human-chimpanzee hybrid iPSC lines generated through in vitro cell fusion maintained on matrigel in mTeSR1 or mTeSR Plus medium (Stem Cell Technologies cat #85850 or cat #100–0276)[24]. The Columbia Stem Cell Core Facility differentiated the iPSCs into multiple cell types including dopaminergic neurons (DA), skeletal myocytes (SKM), and hepatocytes (HEP) using published protocols[27,46,96–104]. Cranial neural crest cells were differentiated as described[25]. Primary dental pulp stem cells (DPSC) samples were obtained from humans and chimpanzees as described[25].

RNA-sequencing and allele-specific expression analysis

All samples were cryopreserved in liquid nitrogen before RNA extraction[105]. Cells were gently thawed and then washed with PBS and cell pellets were collected via centrifugation at 1000 RPM for 5 min. RNA extraction was performed using the RNeasy Mini Kit (Qiagen, 74104) and DNase digestion was performed using RNase-Free DNase Set (Qiagen, 79254) following standard protocols. RNA quality was assessed using the Agilent Bioanalyzer RNA Pico assay. Only samples with an RNA integrity number (RIN) greater than or equal to 7 were used to prepare cDNA libraries. Libraries were prepared using TruSeq Stranded mRNA kits (Illumina, 20020594) and sequenced on Illumina HiSeq 4000 to generate paired-end reads. An allele-specific expression pipeline adapted from our previous work was implemented, using dual reference genome alignment to hg38 and panTro6 to eliminate mapping bias[24,25]. Reads were processed with SeqPrep, mapped using STAR, deduplicated with Picard, and corrected for allelic bias using Hornet. DESeq2 was used for differential expression analysis with FDR correction, defining significant allele-specific expression as FDR < 0.05 with consistent results across both reference genomes[24,25,27,106,107]. For details, see “Generation of allele-specific count tables” and “Identifying genes with ASE” in the Methods section in Wang et al. (2024)[27].

Bisulfite-seq library preparation, sequencing and mapping

DNA was extracted and processed, and whole-genome bisulfite sequencing (WGBS) and reduced representation bisulfite sequencing (RRBS) was performed at Admera Health Biopharma Services using Illumina Novaseq 6000. To minimize reference mapping bias in cross-species comparisons, reference genomes were first masked at previously identified human-chimpanzee SNV positions using bedtools maskfasta, reducing the potential for allelic mapping preferences[27,108]. Reference genome preparation was performed using Bismark’s bismark_genome_preparation with Bowtie2 as the aligner, which converts unmethylated cytosines in silico to enable bisulfite-aware alignment[109,110]. All samples were aligned to both masked human (hg38) and chimpanzee (pantro6) reference genomes to assess mapping quality and species-specific biases. Reads were quality-trimmed and adapter sequences removed using Cutadapt and Trim Galore with default parameters for RRBS libraries, including the removal of artificially filled-in cytosines from the 3’ end of reads[111,112]. Trimmed reads were aligned using Bismark, and methylation levels were calculated as the ratio of unconverted cytosines to total coverage at each CpG site[109]. For hybrid samples, SNPsplit was used to assign reads to either human or chimpanzee allele[113]. Strand information was collapsed for symmetric CpG sites and genomic annotations were lifted between genomic assemblies hg38 and pantro6 using UCSC LiftOver with reciprocal mapping to ensure orthologous regions[114,115].

PCA and hierarchical clustering

Fractional methylation of individual CpG sites was estimated using bismark_methylation_extractor from Bismark[109]. CpG sites shared across samples were identified using bedtools intersect[108]. PCA was performed using PCASamples from methylKit across all samples and samples within each cell

type[116]. Euclidean distance matrices were computed across all samples for hierarchical clustering analysis, with visualization performed using the pheatmap package[117]. Clustering patterns were assessed to validate cell type identity and characterize species-specific methylation signatures.

Beta-binomial regulatory classification framework

Methylation data from human (H) and chimpanzee (C) parental and hybrid samples were analyzed using a beta-binomial generalized linear modeling framework to account for overdispersion inherent to bisulfite sequencing count data. The data included measurements for parental H-genome (ParH), parental C-genome (ParC), hybrid H-allele (HybH), and hybrid C-allele (HybC) for each genomic locus. For each genomic locus (site or annotated region), methylated counts were modeled using two complementary beta-binomial GLM models to separately evaluate *cis* and *trans* regulatory effects:

Model I: *Cis* effect test with methylation data on hybrid alleles

$$\text{logit}(\pi) = \beta_0 + \beta_1(\text{Allele})$$

Model II: *Trans* effect test using System x Allele interaction term with methylation data in both hybrid and parental alleles

$$\text{logit}(\pi) = \beta_0 + \beta_1(\text{System}) + \beta_2(\text{Allele}) + \beta_3(\text{System} \times \text{Allele})$$

Where π is the probability of methylation, System distinguishes parental from hybrid samples, Allele distinguishes between human and chimpanzee methylation, and the interaction term captures any signals involving the change of allelic differences across generations. Statistical significance was assessed through likelihood ratio tests comparing nested models, testing for system effects, allele effects (*cis*-regulation), and system-by-allele interactions (*trans*-regulation). Genomic loci were classified into five categories: conserved, *trans*-only, *cis*-only, *cis* x *trans* interaction, and *cis* + *trans* effects:

Methylation

$$R_P = \log_2\left(\frac{\text{ParH}}{\text{ParC}}\right)$$

$$R_H = \log_2\left(\frac{\text{HybH}}{\text{HybC}}\right)$$

Decision Matrix

Model I FDR	Model II FDR	Regulation
Sig	Non-sig	cis only
Non-sig	Sig	trans only
Sig	Sig $(\Delta \cdot R_H \geq 0)$	cis + trans
Sig	Sig $(\Delta \cdot R_H < 0)$	cis × trans
Non-sig	Non-sig	conserved

$\text{Sig} = \text{FDR} \leq 0.05$, $\Delta = R_P - R_H$

- **Conserved**: No significant allelic differences in either generation (neither parental nor hybrid allelic tests significant)
- ***Trans* only**: Significant allelic difference in parents but not hybrids, indicating trans-acting regulation
- ***Cis* only**: Significant allelic differences in both systems with no significant interaction, indicating consistent cis-acting effects
- ***Cis* × *trans***: Significant allelic differences in both systems with significant interaction, where effects occur in opposite directions
- ***Cis* + *trans***: Significant allelic differences in both systems with significant interaction, where effects are in the same direction

Regional analyses combined CpG-level methylation across defined genomic intervals (for example, promoters). Regional methylation was quantified as pooled fractional methylation, calculated as the total number of methylated cytosines divided by the total coverage across all CpG sites within each region. Nominal p-values < 0.05 from both models were used to provide a schematic overview of site-level CpG methylation in scatterplots and bar plots, preserving the full dynamic range of the data and providing input features for downstream de novo DMR identification. For reporting significant candidate loci under each regulation category as well as formal inference of region-level regulation, Benjamini–Hochberg false discovery rate (FDR) was calculated to account for multiple testing and to evaluate the robustness of inferred relative contribution of regulatory mechanisms across significance thresholds[60]. The same classification and visualization were employed using FDR-corrected p-values on loci and regions to confirm that the relative contributions of regulatory classifications were not driven by significance cutoffs alone (Figure 2—figure supplement 1 and 2). To assess the robustness of *cis*-over-*trans* contributions to methylation with respect to the significance cutoffs, the proportion of significant loci and regions

assigned to each regulatory category was estimated across a range of FDR cutoffs (FDR < 0.05, 0.1, 0.2, 0.25), confirming that the observed *cis* dominance over other divergent regulation categories (including *trans*) was robust to the choice of cutoff (Figure 2—figure supplement 1 and 2).

Changepoint detection for identifying differentially methylated regions with homogeneous cis- or trans-acting methylation divergence

To identify clusters of CpG sites in the same regulation category, we employed the Pruned Exact Linear Time (PELT) algorithm implemented in the changepoint R package to detect shifts in mean methylation scores within each chromosome segment while accounting for a manually specified penalty against over-fitting[118–120]. We adapted this method to identify regions in the genome where neighboring CpG sites are consistently classified by the previously mentioned beta-binomial model to fall under the same regulatory classification. We first segmented the entire genome into large regions where a break is assigned whenever neighboring CpG sites are more than 5000 base pairs apart. Then, we transform all CpG sites on a segment into binary sequences, where the focal regulation category is 1 and alternative categories are 0. To identify *trans*-DMRs, we set focal pure *trans* regulation to 1, and all other regulation categories to 0. To identify *cis*-DMRs, we set focal pure *cis* regulation to 1, and all other regulation categories to 0. To identify *trans*-DMR specifically for HOMER enrichment inputs, we used a relaxed gradient scoring approach. We set focal pure *trans* regulation to 1, conserved to 0.75, additive (*cis* + *trans*) and interactive (*cis* x *trans*) regulation to 0.5 and pure *cis* regulation category to 0. To determine the optimal penalty parameter, we applied the CROPS (Changepoints for a Range Of Penalties) method, testing penalty values ranging from 0.01 to 2.0. For each penalty range giving the same number of segments, we recorded the number of detected changepoints and corresponding model fit statistics. The optimal penalty for each segment was determined using elbow point detection via the R package kneedle, which identifies the point of maximum distance from line connecting endpoints in the penalty-versus-changepoints curve[121]. We then applied a homogeneity filter requiring segments to have a minimum target proportion of 0.6, where the target proportion is the fraction of CpG sites within the segment that (i) belong to the target regulatory class under the beta-binomial cis/trans classification and (ii) share a consistent species bias in methylation (all human-biased or all chimpanzee-biased), to ensure biological relevance of identified regions. For segments where fewer than three distinct penalty values yielded different numbers of changepoints, we selected the penalty that maximized the number of segments passing the homogeneity threshold. Segments without detected changepoints were evaluated directly against the homogeneity criterion to assess their potential as single-segment DMRs.

To refine segment boundaries and remove spurious edge effects before merging overlapping DMRs, we implemented a post-processing trimming step that removes segment edges containing sites with regulatory patterns inconsistent with the dominant signal within each segment. This ensures that identified DMRs represent coherent regions of coordinated methylation changes rather than artifacts of the segmentation process.

Identification of cell-type specific DMRs, species-specific DMRs, and cell-type specific species-specific DMRs

Examples of cell-type specific DMRs (*HOXA9*), species-specific DMRs (*CTSF*), and cell type-specific species-specific DMRs (*LGALS8*) were identified in CNCC, DA, iPSC, SKM and HEP hybrid samples using the DSS-single (Dispersion Shrinkage for Sequencing data)[122–125]. For cell type-specific DMR identification, we used the formula ~species + cell type + species:cell type to test for cell type-specific methylation differences, while species-specific DMRs analysis employed ~species to identify consistent species differences across cell types. Downstream identification of cell-type specific species-specific DMRs identifies the regions where the interaction term species:cell type is significant. The DSS analysis incorporated smoothing across neighboring CpG sites with a smoothing span of 500 bp and required a minimum of 5 CpG sites per region. Statistical

significance was determined using Wald tests with FDR < 0.05, and effect size filtering retained regions with methylation difference ≥20%. Integration with gene expression data was performed through Wilcoxon rank-sum tests to identify cell type-specific gene expression patterns associated with cell-type specific DMRs, while permutations (n=1000) were used to assess the significance of coordinated methylation-expression changes in species-specific patterns, which accounts for any background correlation between neighboring genomic features.

Expression-methylation correlation analysis

Human transcript coordinates and gene body annotations were obtained from the hg38.knownGene.gtf.gz file via UCSC Table Browser[114], with transcript start positions designated as transcription start sites (TSS). CpG methylation data aggregated within ENCODE-defined promoter regions were analyzed using a beta-binomial classification framework to determine regulatory categories, while corresponding gene expression data underwent parallel hypothesis testing using a negative-binomial model (ENCODE Project Consortium, et al., 2020). Genes with the same classification for both methylation and expression—either both *cis* or both *trans*—were then analyzed separately. For genes within each regulatory category, the 2 kb surrounding each TSS was split into 10 bins of 200 bp where individual CpG sites were pooled and averaged as fractional methylation values and correlated with expression TPM values. Custom scripts used for binning and correlation are adapted from beta_profile_region.py from CpGtools (Wei et al., 2021).

Sign test selection analysis

To identify coordinated methylation-expression changes potentially under selection, we used a binomial sign test framework that accounts for background probabilities of concordant changes[24,25,27,31,128,129]. Promoter regions were obtained from the ENCODE Project Consortium[126]. Canonical transcripts were downloaded from the UCSC Table Browser hg38 assembly. For genes with multiple annotated promoters, we selected the promoter that is the closest to 5' UTR of the gene, to ensure each gene is assigned one corresponding promoter[114]. Promoter regions were mapped to canonical transcripts, and functional enrichment using pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Human Phenotype Ontology (HPO) was assessed using directional sign tests that leverage the independence of *cis*-regulatory regions to detect non-random patterns indicative of lineage-specific selection rather than neutral drift[1,24,25,27,75–79]. For the methylation-only sign test, gene sets with more than 200 genes or fewer than 20 genes with corresponding promoter methylation data were excluded from the analysis. Lineage-specific selection on a gene set was inferred if there was a statistically significant directionality bias in chimpanzee vs. human promoter ASM. For sign test on genes with putatively repressive methylation—that is, genes where increased methylation within 1 kb of their TSS correspond to decreased expression—we first classified genes as either consistent with repressive methylation (either Ch>Hu methylation paired with Hu>Ch expression, or Hu>Ch methylation paired with Ch>Hu expression) or not consistent. Gene sets with more than 200 genes or fewer than 10 genes with corresponding methylation and expression data were excluded from the analysis. We then reran the sign test using a new background probability defined as the overall fraction of genes with repressive methylation in each cell type. FDRs of the resulting binomial p-values were estimated by randomly permuting the signs of all genes and performing the binomial sign test 10^4 times. To calculate FDR, we compared each pathway's empirical binomial p-value against the distribution of permuted p-values. The expected number of false positives was estimated as the average number of pathways reaching that significance level per permutation. FDR was then defined as this expected count of false positives divided by the observed count of pathways with empirical p-values at least as extreme.

cis-DMR sequence variant analysis and *trans*-DMR motif enrichment analysis

High-confidence human-chimpanzee single nucleotide variants (SNVs) were identified using a dual-reference approach with stringent filtering criteria as described previously[27]. Allele-specific methylation within different genomic ranges (25bp, 50bp, 100bp) of CpG-disrupting variants was calculated. Proportions of regions with Hu>Ch and Hu<Ch methylation were calculated for SNVs in three categories: SNVs that disrupt CpG dinucleotides in humans, SNVs that disrupt CpG sites in chimpanzees, and SNVs that lead to no interspecies changes in CpG content between the species. The significance of the difference in relative proportions of Hu>Ch and Hu<Ch of SNVs in the three categories was assessed using a two-proportion test. For proximity tests, the Mann-Whitney U test[67] was used to compare the distribution of distances from all *cis* or all *trans*-DMRs to the closest CpG-disrupting SNV. HOMER's findMotifsGenome.pl was used to test for motif enrichment in the *trans*-DMR set[52].

Data availability

Raw and processed data generated by this study are publicly available through the Gene Expression Omnibus under accession GSE315270 (BS-seq) and GSE315269 (RNA-seq). The raw and processed bulk RNA-seq data of hybrid and parental iPSCs by Agoglia et al. is available at *https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144825* and the raw and processed bulk RNA-seq data of hybrid and parental CNCC by Gokhman et al. is available at *https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146481*. The raw and processed bulk RNA-seq data of hybrid hepatocytes and hybrid skeletal myocytes from Wang et al. are available at *https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE232949*. scRNA-seq data of parental hepatocytes and parental skeletal myocytes from Barr and Rhodes are available here: *https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE201516*. The log fold-changes and associated statistics were used directly from the supplemental material Data S1 of Barr and Rhodes. The human and chimpanzee genomes used are available here: *https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/* and here: *https://hgdownload.soe.ucsc.edu/goldenPath/panTro6/bigZips/* respectively. The human-chimp pairwise alignment used to identify SNVs and indels is available here: *https://hgdownload.soe.ucsc.edu/goldenPath/hg38/vsPanTro6/*. All scripts for performing analyses and making figures in this manuscript are publicly available at *https://github.com/zhenzhenma120/The-causes-and-consequences-of-human-specific-DNA-methylation*. All supplemental files are available on Zenodo (*https://doi.org/10.5281/zenodo.18205339*).

Acknowledgements

We thank Leslie Magtanong for cell culture work and the Columbia Stem Cell Core Facility for performing cell differentiation. We are grateful to all members of the Fraser Lab for helpful discussions and feedback on this work.

Additional files

Supplementary Figures

Additional information

Funding

Funder	Grant reference number	Author
HHS | NIH | National Human Genome Research Institute (NHGRI)	R01HG012285	Hunter B Fraser

HHS | NIH | National Institute of General Medical Sciences (NIGMS)	R35GM156526	Hunter B Fraser

Author ORCID iDs

Zhenzhen Ma: https://orcid.org/0009-0007-0107-7463
Alexander L Starr: https://orcid.org/0000-0001-6297-9742
David Gokhman: https://orcid.org/0000-0002-3536-9006
Hunter B Fraser: https://orcid.org/0000-0001-8400-8541

References

1. **Fraser H. B.** (2013) Gene expression drives local adaptation in humans. *Genome Res* **23**:1089-1096 https://doi.org/10.1101/gr.152710.112 | PubMed
2. **King M.-C.,** Wilson A. C. (1975) Evolution at Two Levels in Humans and Chimpanzees. *Science* **188**:107-116 https://doi.org/10.1126/science.1090005 | PubMed
3. **Reilly S. K.,** Noonan J. P. (2016) Evolution of Gene Regulation in Humans. *Annual Review of Genomics and Human Genetics* **17**:45-67 https://doi.org/10.1146/annurev-genom-090314-045935 | PubMed
4. **Romero I. G.,** Ruvinsky I., Gilad Y. (2012) Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics* **13**:505-516 https://doi.org/10.1038/nrg3229 | PubMed
5. **Prud'homme B.,** Gompel N., Carroll S. B. (2007) Emerging principles of regulatory evolution. *Proceedings of the National Academy of Sciences* **104**:8605-8612 https://doi.org/10.1073/pnas.0700488104 | PubMed
6. **Wittkopp P. J.,** Kalay G. (2012) Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics* **13**:59-69 https://doi.org/10.1038/nrg3095 | PubMed
7. **Blake L. E.,** et al. (2020) A comparison of gene expression and DNA methylation patterns across tissues and species. *Genome Res* **30**:250-262 https://doi.org/10.1101/gr.254904.119 | PubMed
8. **Kelley J. L.,** Gilad Y. (2020) Effective study design for comparative functional genomics. *Nature Reviews Genetics* **21**:385-386 https://doi.org/10.1038/s41576-020-0242-z | PubMed
9. **Ma S.,** et al. (2022) Molecular and cellular evolution of the primate dorsolateral prefrontal cortex. *Science* **377** https://doi.org/10.1126/science.abo7257 | PubMed
10. **Mack K. L.,** Nachman M. W. (2017) Gene Regulation and Speciation. *Trends in Genetics* **33**:68-80 https://doi.org/10.1016/j.tig.2016.11.003 | PubMed
11. **Jones P. A.** (2012) Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics* **13**:484-492 https://doi.org/10.1038/nrg3230 | PubMed
12. **Pai A. A.,** Bell J. T., Marioni J. C., Pritchard J. K., Gilad Y. (2011) A genome-wide study of DNA methylation patterns and gene expression levels in multiple human and chimpanzee tissues. *PLoS Genet* **7** https://doi.org/10.1371/journal.pgen.1001316 | PubMed
13. **Gokhman D.,** et al. (2020) Differential DNA methylation of vocal and facial anatomy genes in modern humans. *Nat Commun* **11**:1189 https://doi.org/10.1038/s41467-020-15020-6 | PubMed
14. **Zeng J.,** et al. (2012) Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. *Am J Hum Genet* **91**:455-465 https://doi.org/10.1016/j.ajhg.2012.07.024 | PubMed
15. **Hernando-Herraez I.,** Garcia-Perez R., Sharp A. J., Marques-Bonet T. (2015) DNA Methylation: Insights into Human Evolution. *PLoS Genet* **11**:e1005661 https://doi.org/10.1371/journal.pgen.1005661 | PubMed
16. **Hernando-Herraez I.,** et al. (2013) Dynamics of DNA Methylation in Recent Human and Great Ape Evolution. *PLoS Genet* **9** https://doi.org/10.1371/journal.pgen.1003763 | PubMed

17. **Hernando-Herraez I.,** et al. (2015) The interplay between DNA methylation and sequence divergence in recent human evolution. *Nucleic Acids Res* **43**:8204-14 https://doi.org/10.1093/nar/gkv693 | PubMed
18. **Mendizabal I.,** et al. (2016) Comparative Methylome Analyses Identify Epigenetic Regulatory Loci of Human Brain Evolution. *Mol Biol Evol* **33**:2947-2959 https://doi.org/10.1093/molbev/msw176 | PubMed
19. **Fukuda K.,** et al. (2013) Regional DNA methylation differences between humans and chimpanzees are associated with genetic changes, transcriptional divergence and disease genes. *J Hum Genet* **58**:446-454 https://doi.org/10.1038/jhg.2013.55 | PubMed
20. **Enard W.,** et al. (2004) Differences in DNA methylation patterns between humans and chimpanzees. *Current Biology* **14**:R148-R149 https://doi.org/10.1016/j.cub.2004.01.042 | PubMed
21. **Christoforou A.,** et al. (2012) Linkage-Disequilibrium-Based Binning Affects the Interpretation of GWASs. *The American Journal of Human Genetics* **90**:727-733 https://doi.org/10.1016/j.ajhg.2012.02.025 | PubMed
22. **Banerjee N.,** et al. (2019) Analysis of differentially methylated regions in great apes and extinct hominids provides support for the evolutionary hypothesis of schizophrenia. *Schizophr Res* **206**:209-216 https://doi.org/10.1016/j.schres.2018.11.025 | PubMed
23. **Guevara E. E.,** et al. (2020) Age-associated epigenetic change in chimpanzees and humans. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**:20190616 https://doi.org/10.1098/rstb.2019.0616 | PubMed
24. **Agoglia R. M.,** et al. (2021) Primate cell fusion disentangles gene regulatory divergence in neurodevelopment. *Nature* **592**:421-427 https://doi.org/10.1038/s41586-021-03343-3 | PubMed
25. **Gokhman D.,** et al. (2021) Human–chimpanzee fused cells reveal cis-regulatory divergence underlying skeletal evolution. *Nat Genet* **53**:467-476 https://doi.org/10.1038/s41588-021-00804-3 | PubMed
26. **Song J. H. T.,** et al. (2021) Genetic studies of human–chimpanzee divergence using stem cell fusions. *Proceedings of the National Academy of Sciences* **118** https://doi.org/10.1073/pnas.2117557118 | PubMed
27. **Wang B.,** Starr A. L., Fraser H. B. (2024) Cell-type-specific cis-regulatory divergence in gene expression and chromatin accessibility revealed by human-chimpanzee hybrid cells. *eLife* **12** https://doi.org/10.7554/elife.89594 | PubMed
28. **Starr A. L.,** Gokhman D., Fraser H. B. (2023) Accounting for cis-regulatory constraint prioritizes genes likely to affect species-specific traits. *Genome Biol* **24** https://doi.org/10.1186/s13059-023-02846-8 | PubMed
29. **Song J. H. T.,** et al. (2025) Human-chimpanzee tetraploid system defines mechanisms of species-specific neural gene regulation. *bioRxiv* https://doi.org/10.1101/2025.03.31.646367
30. **Barr K. A.,** Rhodes K. L., Gilad Y. (2023) The relationship between regulatory changes in cis and trans and the evolution of gene expression in humans and chimpanzees. *Genome Biol* **24** https://doi.org/10.1186/s13059-023-03019-3 | PubMed
31. **Starr A. L.,** Fraser H. B. (2025) A General Principle of Neuronal Evolution Reveals a Human-Accelerated Neuron Type Potentially Underlying the High Prevalence of Autism in Humans. *Mol Biol Evol* **42** https://doi.org/10.1093/molbev/msaf189 | PubMed
32. **Mack K. L.,** et al. (2023) Allele-specific expression reveals genetic drivers of tissue regeneration in mice. *Cell Stem Cell* **30**:1368-1381.e6 https://doi.org/10.1016/j.stem.2023.08.010 | PubMed
33. **Hu C. K.,** et al. (2022) cis-Regulatory changes in locomotor genes are associated with the evolution of burrowing behavior. *Cell Rep* **38** https://doi.org/10.1016/j.celrep.2022.110360 | PubMed
34. **Vilgalys T. P.,** et al. (2019) Evolution of DNA Methylation in Papio Baboons. *Mol Biol Evol* **36**:527-540 https://doi.org/10.1093/molbev/msy227 | PubMed

35. **Housman G.,** Quillen E. E., Stone A. C. (2021) An evolutionary perspective of DNA methylation patterns in skeletal tissues using a baboon model of osteoarthritis. *Journal of Orthopaedic Research* **39**:2260-2269 https://doi.org/10.1002/jor.24957 | PubMed

36. **Silva J.,** et al. (2009) Nanog Is the Gateway to the Pluripotent Ground State. *Cell* **138**:722-737 https://doi.org/10.1016/j.cell.2009.07.039 | PubMed

37. **Gerdes J.,** et al. (1984) Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67. *J Immunol* **133**:1710-5 https://doi.org/10.4049/jimmunol.133.4.1710 | PubMed

38. **Hatori A.,** et al. (2023) VCAM-1 and GFPT-2: Predictive markers of osteoblast differentiation in human dental pulp stem cells. *Bone* **166**:116575 https://doi.org/10.1016/j.bone.2022.116575 | PubMed

39. **Svandova E.,** et al. (2023) Markers of dental pulp stem cells in in vivo developmental context. *Annals of Anatomy - Anatomischer Anzeiger* **250**:152149 https://doi.org/10.1016/j.aanat.2023.152149 | PubMed

40. **Bai X.,** et al. (2023) Dental Pulp Stem Cells for Bone Tissue Engineering: A Literature Review. *Stem Cells Int* **2023**:7357179 https://doi.org/10.1155/2023/7357179 | PubMed

41. **Soldatov R.,** et al. (2019) Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364** https://doi.org/10.1126/science.aas9536 | PubMed

42. **Bronner M. E.,** Simões-Costa M. (2016) The Neural Crest Migrating into the Twenty-First Century. In: Wassarman PM (Ed). *Current topics in developmental biology* pp. 115-134 https://doi.org/10.1016/bs.ctdb.2015.12.003 | PubMed

43. **Cordero D. R.,** et al. (2011) Cranial neural crest cells on the move: Their roles in craniofacial development. *Am J Med Genet A* **155**:270-279 https://doi.org/10.1002/ajmg.a.33702 | PubMed

44. **Metzakopian E.,** et al. (2012) Genome-wide characterization of Foxa2 targets reveals upregulation of floor plate genes and repression of ventrolateral genes in midbrain dopaminergic progenitors. *Development* **139**:2625-2634 https://doi.org/10.1242/dev.081034 | PubMed

45. **Cassimeris L.,** Spittle C. (2001) Regulation of microtubule-associated proteins. *International Review of Cytology* 163-226 https://doi.org/10.1016/S0074-7696(01)10006-9 | PubMed

46. **Xu P.,** et al. (2022) Human midbrain dopaminergic neuronal differentiation markers predict cell therapy outcomes in a Parkinson's disease model. *Journal of Clinical Investigation* **132** https://doi.org/10.1172/jci156768 | PubMed

47. **Bockenhauer D.,** Jaureguiberry G. (2016) HNF1B-associated clinical phenotypes: the kidney and beyond. *Pediatric Nephrology* **31**:707-714 https://doi.org/10.1007/s00467-015-3142-2 | PubMed

48. **Grozdanov P. N.,** Yovchev M. I., Dabeva M. D. (2006) The oncofetal protein glypican-3 is a novel marker of hepatic progenitor/oval cells. *Laboratory Investigation* **86**:1272-1284 https://doi.org/10.1038/labinvest.3700479 | PubMed

49. **Graffmann N.,** Scherer B., Adjaye J. (2022) In vitro differentiation of pluripotent stem cells into hepatocyte like cells – Basic principles and current progress. *Stem Cell Res* **61**:102763 https://doi.org/10.1016/j.scr.2022.102763 | PubMed

50. **Rubenstein A. B.,** et al. (2020) Single-cell transcriptional profiles in human skeletal muscle. *Sci Rep* **10**:229 https://doi.org/10.1038/s41598-019-57110-6 | PubMed

51. **Hallgrímsdóttir I. B.,** Carilli M., Pachter L. (2024) Estimating cis and trans contributions to differences in gene regulation. *bioRxiv* https://doi.org/10.1101/2024.07.13.603403

52. **Heinz S.,** et al. (2010) Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* **38**:576-589 https://doi.org/10.1016/j.molcel.2010.05.004 | PubMed

53. **Reizel Y.,** et al. (2021) FoxA-dependent demethylation of DNA initiates epigenetic memory of cellular identity. *Dev Cell* **56**:602-612.e4 https://doi.org/10.1016/j.devcel.2021.02.005 | PubMed

54. **Li J.,** et al. (2022) TET1 dioxygenase is required for FOXA2-associated chromatin remodeling in pancreatic beta-cell differentiation. *Nat Commun* **13**:3907 https://doi.org/10.1038/s41467-022-31611-x | PubMed

55. **Stolz J. R.,** et al. (2021) Clustered mutations in the GRIK2 kainate receptor subunit gene underlie diverse neurodevelopmental disorders. *The American Journal of Human Genetics* **108**:1692-1709 https://doi.org/10.1016/j.ajhg.2021.07.007 | PubMed

56. **Gosalia N.,** Yang R., Kerschner J. L., Harris A. (2015) FOXA2 regulates a network of genes involved in critical functions of human intestinal epithelial cells. *Physiol Genomics* **47**:290-297 https://doi.org/10.1152/physiolgenomics.00024.2015 | PubMed

57. **Iwafuchi-Doi M.,** et al. (2016) The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Mol Cell* **62**:79-91 https://doi.org/10.1016/j.molcel.2016.03.001 | PubMed

58. **Lee K.,** et al. (2019) FOXA2 Is Required for Enhancer Priming during Pancreatic Differentiation. *Cell Rep* **28**:382-393.e7 https://doi.org/10.1016/j.celrep.2019.06.034 | PubMed

59. **Sérandour A. A.,** et al. (2011) Epigenetic switch involved in activation of pioneer factor FOXA1-dependent enhancers. *Genome Res* **21**:555-565 https://doi.org/10.1101/gr.111534.110 | PubMed

60. **Benjamini Y.,** Hochberg Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Series B Stat Methodol* **57**:289-300 https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

61. **Braccioli L.,** Nijboer C., Coffer P. (2018) Forkhead box protein P1, a key player in neuronal development?. *Neural Regen Res* **13**:801 https://doi.org/10.4103/1673-5374.232467 | PubMed

62. **Lalmansingh A. S.,** Karmakar S., Jin Y., Nagaich A. K. (2012) Multiple modes of chromatin remodeling by Forkhead box proteins. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1819**:707-715 https://doi.org/10.1016/j.bbagrm.2012.02.018 | PubMed

63. **Ee L.,** et al. (2025) Enhancer remodeling by OTX2 directs specification and patterning of mammalian definitive endoderm. *Dev Cell* https://doi.org/10.1016/j.devcel.2025.07.020 | PubMed

64. **Hellman A.,** Chess A. (2010) Extensive sequence-influenced DNA methylation polymorphism in the human genome. *Epigenetics Chromatin* **3**:11 https://doi.org/10.1186/1756-8935-3-11 | PubMed

65. **Gertz J.,** et al. (2011) Analysis of DNA Methylation in a Three-Generation Family Reveals Widespread Genetic Influence on Epigenetic Regulation. *PLoS Genet* **7**:e1002228 https://doi.org/10.1371/journal.pgen.1002228 | PubMed

66. **Zhi D.,** et al. (2013) SNPs located at CpG sites modulate genome-epigenome interaction. *Epigenetics* **8**:802-806 https://doi.org/10.4161/epi.25501 | PubMed

67. **Mann H. B.,** Whitney D. R. (1947) On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics* **18**:50-60 https://doi.org/10.1214/aoms/1177730491

68. **Jaenisch R.,** Bird A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **33**:245-254 https://doi.org/10.1038/ng1089 | PubMed

69. **Murrell A.,** Rakyan V. K., Beck S. (2005) From genome to epigenome. *Hum Mol Genet* **14**:R3-R10 https://doi.org/10.1093/hmg/ddi110 | PubMed

70. **Wang J.,** et al. (2024) DNA methylation and transcriptome analysis reveal epigenomic differences among three macaque species. *Evol Appl* **17**:e13604 https://doi.org/10.1111/eva.13604 | PubMed

71. **Saxonov S.,** Berg P., Brutlag D. L. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences* **103**:1412-1417 https://doi.org/10.1073/pnas.0510310103 | PubMed

72. **Boman J.,** Qvarnström A., Mugal C. F. (2024) Regulatory and evolutionary impact of DNA methylation in two songbird species and their naturally occurring F1 hybrids. *BMC Biol* **22**:124 https://doi.org/10.1186/s12915-024-01920-2 | PubMed

73. **Stein R.,** Razin A., Cedar H. (1982) In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. *Proceedings of the National Academy of Sciences* **79**:3418-3422 https://doi.org/10.1073/pnas.79.11.3418 | PubMed

74. **Fraser H. B.** (2011) Genome-wide approaches to the study of adaptive gene expression evolution. *BioEssays* **33**:469-477 https://doi.org/10.1002/bies.201000094 | PubMed

75. **Sharon E.,** et al. (2018) Functional Genetic Variants Revealed by Massively Parallel Precise Genome Editing. *Cell* **175**:544-557.e16 https://doi.org/10.1016/j.cell.2018.08.057 | PubMed

76. **Artieri C. G.,** Fraser H. B. (2014) Evolution at two levels of gene expression in yeast. *Genome Res* **24**:411-421 https://doi.org/10.1101/gr.165522.113 | PubMed

77. **Nourmohammad A.,** et al. (2017) Adaptive Evolution of Gene Expression in Drosophila. *Cell Rep* **20**:1385-1395 https://doi.org/10.1016/j.celrep.2017.07.033 | PubMed

78. **Papakostas S.,** et al. (2014) Gene pleiotropy constrains gene expression changes in fish adapted to different thermal conditions. *Nat Commun* **5**:4071 https://doi.org/10.1038/ncomms5071 | PubMed

79. **Kanehisa M.,** Goto S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**:27-30 https://doi.org/10.1093/nar/28.1.27 | PubMed

80. **Gargano M. A.,** et al. (2024) The Human Phenotype Ontology in 2024: phenotypes around the world. *Nucleic Acids Res* **52**:D1333-D1346 https://doi.org/10.1093/nar/gkad1005 | PubMed

81. **Christopher Dean M.** (2006) Tooth microstructure tracks the pace of human life-history evolution. *Proceedings of the Royal Society B: Biological Sciences* **273**:2799-2808 https://doi.org/10.1098/rspb.2006.3583 | PubMed

82. **Andrew D. D.,** David A. N. H., Olli R. (2017) *Fundamentals of Craniofacial Growth* CRC Press. https://doi.org/10.1201/9780203755327

83. **Smith T. M.,** et al. (2013) First molar eruption, weaning, and life history in living wild chimpanzees. *Proceedings of the National Academy of Sciences* **110**:2787-2791 https://doi.org/10.1073/pnas.1218746110 | PubMed

84. **Rooryck C.,** et al. (2011) Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. *Nat Genet* **43**:197-203 https://doi.org/10.1038/ng.757 | PubMed

85. **Jones J. H.** (2011) Primates and the Evolution of Long, Slow Life Histories. *Current Biology* **21**:R708-R717 https://doi.org/10.1016/j.cub.2011.08.025 | PubMed

86. **Davison R. J.,** Gurven M. D. (2021) Human uniqueness? Life history diversity among small-scale societies and chimpanzees. *PLoS One* **16**:e0239170 https://doi.org/10.1371/journal.pone.0239170 | PubMed

87. **Valbuena S.,** Lerma J. (2021) Kainate Receptors, Homeostatic Gatekeepers of Synaptic Plasticity. *Neuroscience* **456**:17-26 https://doi.org/10.1016/j.neuroscience.2019.11.050 | PubMed

88. **Tischfield M. A.,** et al. (2010) Human TUBB3 Mutations Perturb Microtubule Dynamics, Kinesin Interactions, and Axon Guidance. *Cell* **140**:74-87 https://doi.org/10.1016/j.cell.2009.12.011 | PubMed

89. **Poirier K.,** et al. (2010) Mutations in the neuronal β-tubulin subunit TUBB3 result in malformation of cortical development and neuronal migration defects. *Hum Mol Genet* **19**:4462-4473 https://doi.org/10.1093/hmg/ddq377 | PubMed

90. **Romano R.,** Bucci C. (2020) Role of EGFR in the Nervous System. *Cells* **9** https://doi.org/10.3390/cells9081887 | PubMed

91. **Noor A.,** et al. (2008) CC2D2A, encoding a coiled-coil and C2 domain protein, causes autosomal-recessive mental retardation with retinitis pigmentosa. *Am J Hum Genet* **82**:1011-8 https://doi.org/10.1016/j.ajhg.2008.01.021 | PubMed

92. **Li Y.,** et al. (2016) Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses. *Proceedings of the National Academy of Sciences* **113**:2241-2246 https://doi.org/10.1073/pnas.1519657113 | PubMed

93. **Yee L. J.,** et al. (2007) Myxovirus-1 and protein kinase haplotypes and fibrosis in chronic hepatitis C virus. *Hepatology* **46**:74-83 https://doi.org/10.1002/hep.21636 | PubMed

94. **Walker C. M.** (1997) Comparative features of hepatitis C virus infection in humans and chimpanzees. *Springer Semin Immunopathol* **19**:85-98 https://doi.org/10.1007/bf00945027 | PubMed

95. **Milind N.,** et al. (2024) Buffering and non-monotonic behavior of gene dosage response curves for human complex traits. *medRxiv* https://doi.org/10.1101/2024.11.11.24317065

96. **Burridge P. W.,** et al. (2014) Chemically defned generation of human cardiomyocytes. *Nat Methods* **11**:855-860 https://doi.org/10.1038/nmeth.2999 | PubMed

97. **Chal J.,** et al. (2016) Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat Protoc* **11**:1833-1850 https://doi.org/10.1038/nprot.2016.110 | PubMed

98. **Korytnikov R.,** Nostro M. C. (2016) Generation of polyhormonal and multipotent pancreatic progenitor lineages from human pluripotent stem cells. *Methods* **101**:56-64 https://doi.org/10.1016/j.ymeth.2015.10.017 | PubMed

99. **Mallanna S. K.,** Duncan S. A. (2013) Differentiation of hepatocytes from pluripotent stem cells. *Curr Protoc Stem Cell Biol* **1** https://doi.org/10.1002/9780470151808.sc01g04s26 | PubMed

100. **Maury Y.,** et al. (2015) Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol* **33**:89-96 https://doi.org/10.1038/nbt.3049 | PubMed

101. **Sharma R.,** et al. (2019) Clinical-Grade Stem Cell-Derived Retinal Pigment Epithelium Patch Rescues Retinal Degeneration in Rodents and Pigs. *Sci. Transl. Med* **11** https://doi.org/10.1126/scitranslmed.aat5580 | PubMed

102. **Fedele S.,** et al. (2017) Expansion of human midbrain floor plate progenitors from induced pluripotent stem cells increases dopaminergic neuron differentiation potential. *Sci Rep* **7**:6036 https://doi.org/10.1038/s41598-017-05633-1 | PubMed

103. **Studer L.** (2012) Derivation of dopaminergic neurons from pluripotent stem cells. In: Dunnett SB, Björklund A (Eds). *Progress in brain research* pp. 243-263 https://doi.org/10.1016/B978-0-444-59575-1.00011-9 | PubMed

104. **Kriks S.,** et al. (2011) Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* **480**:547-551 https://doi.org/10.1038/nature10648 | PubMed

105. **Milani P.,** et al. (2016) Cell freezing protocol suitable for ATAC-Seq on motor neurons derived from human induced pluripotent stem cells. *Sci Rep* **6**:25474 https://doi.org/10.1038/srep25474 | PubMed

106. **Dobin A.,** et al. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15-21 https://doi.org/10.1093/bioinformatics/bts635 | PubMed

107. **St. John J.** (2024) SeqPrep.

108. **Quinlan A. R.,** Hall I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841-842 https://doi.org/10.1093/bioinformatics/btq033 | PubMed

109. **Krueger F.,** Andrews S. R. (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**:1571-1572 https://doi.org/10.1093/bioinformatics/btr167 | PubMed

110. **Langmead B.,** Salzberg S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:357-359 https://doi.org/10.1038/nmeth.1923 | PubMed

111. **Martin M.** (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**:10 https://doi.org/10.14806/ej.17.1.200

112. **Krueger F.** (2021) TrimGalore.

113. **Krueger F.,** Andrews S. R. (2016) SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes. *F1000Res* **5**:1479 https://doi.org/10.12688/f1000research.9037.2 | PubMed

114. **Karolchik D.,** et al. (2004) The UCSC table browser data retrieval tool. *Nucleic Acids Res* **32** https://doi.org/10.1093/nar/gkh103 | PubMed

115. **Hinrichs A. S.** (2006) The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res* **34**:D590-D598 https://doi.org/10.1093/nar/gkj144 | PubMed

116. **Akalin A.,** et al. (2012) methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol* **13**:R87 https://doi.org/10.1186/gb-2012-13-10-r87 | PubMed

117. **Kolde R.** (2025) pheatmap: Pretty Heatmaps. R package. version: v.1.0.13

118. **Killick R.,** Fearnhead P., Eckley I. A. (2012) Optimal Detection of Changepoints With a Linear Computational Cost. *J Am Stat Assoc* **107**:1590-1598 https://doi.org/10.1080/01621459.2012.737745

119. **Haynes K.,** Eckley I. A., Fearnhead P. (2017) Computationally Efficient Changepoint Detection for a Range of Penalties. *Journal of Computational and Graphical Statistics* **26**:134-143 https://doi.org/10.1080/10618600.2015.1116445

120. **Yokoyama T.,** Miura F., Araki H., Okamura K., Ito T. (2015) Changepoint detection in base-resolution methylome data reveals a robust signature of methylated domain landscape. *BMC Genomics* **16**:594 https://doi.org/10.1186/s12864-015-1809-5 | PubMed

121. **Satopää V.,** Albrecht J., Irwin D., Raghavan B. (2011) Finding a 'kneedle' in a haystack: Detecting knee points in system behavior. In: Proceedings - International Conference on Distributed Computing Systems. pp. 166-171 https://doi.org/10.1109/ICDCSW.2011.20

122. **Wu H.,** Wang C., Wu Z. (2013) A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data. *Biostatistics* **14**:232-243 https://doi.org/10.1093/biostatistics/kxs033 | PubMed

123. **Feng H.,** Conneely K. N., Wu H. (2014) A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Res* **42** https://doi.org/10.1093/nar/gku154 | PubMed

124. **Wu H.,** et al. (2015) Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. *Nucleic Acids Res* **43** https://doi.org/10.1093/nar/gkv715 | PubMed

125. **Park Y.,** Wu H. (2016) Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics* **32**:1446-1453 https://doi.org/10.1093/bioinformatics/btw026 | PubMed

126. **ENCODE Project Consortium,** et al. (2020) Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583**:699-710 https://doi.org/10.1038/s41586-020-2493-4 | PubMed

127. **Wei T.,** et al. (2021) CpGtools: a python package for DNA methylation analysis. *Bioinformatics* **37**:1598-1599 https://doi.org/10.1093/bioinformatics/btz916 | PubMed

128. **Orr H. A.** (1998) Testing Natural Selection vs. Genetic Drift in Phenotypic Evolution Using Quantitative Trait Locus Data. *Genetics* **149**:2099-2104 https://doi.org/10.1093/genetics/149.4.2099 | PubMed

129. **Simon N. M.,** et al. (2024) Stem cell transcriptional profiles from mouse subspecies reveal cis-regulatory evolution at translation genes. *Heredity* **133**:308-316 https://doi.org/10.1038/s41437-024-00715-z | PubMed

Ma Z, Starr A, Gokhman D, Fraser H (2026) The causes and consequences of human-specific DNA methylation [WGBS; RRBS]. NCBI Gene Expression Omnibus. ID GSE315270 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE315270

Ma Z, Starr A, Gokhman D, Fraser H (2026) The causes and consequences of human-specific DNA methylation [RNA-Seq]. NCBI Gene Expression Omnibus. ID GSE315269 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE315269

Ma Z, Starr A, Gokhman D, Fraser H (2026) Supplemental Files for publication "The causes and consequences of human-specific DNA methylation". Zenodo. https://doi.org/10.5281/zenodo.18205339

Agoglia R, Fraser H, Pasca S (2021) Primate cell fusion disentangles gene regulatory divergence in neural development. NCBI Gene Expression Omnibus. ID GSE144825 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE144825

Gokhman D, Fraser H (2021) Human-chimpanzee hybrid cells reveal gene regulatory evolution underlying skeletal divergence. NCBI Gene Expression Omnibus. ID GSE146481 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE146481

Wang B, Starr A, Fraser H (2024) Cell type-specific cis-regulatory divergence in gene expression and chromatin accessibility revealed by human-chimpanzee hybrid cells. NCBI Gene Expression Omnibus. ID GSE232949 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE232949

Barr K, Rhodes K, Gilad Y (2023) The relationship between regulatory changes in cis and trans and the evolution of gene expression in humans and chimpanzees. NCBI Gene Expression Omnibus. ID GSE201516 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE201516

Peer reviews

Reviewer #1 (Public review):

Ma et al. use human-chimpanzee tetraploid cells across different cell types to identify the genetic causes and then transcriptomic consequences of divergence in DNA methylation. They conclude that the evolution of DNA methylation is driven primarily by cis-regulatory changes, and that the evolution of CpG sites contributes to cis-regulation, while transcription factor expression underlies some trans changes. They then argue that divergence in DNA methylation is associated with changes in gene expression and may contribute to human phenotypes.

The tetraploid model is able to provide compelling evidence that most regulatory evolution occurs due to cis-regulatory changes. My only concern is that the extent of trans-changes may be overstated, as almost all are eliminated by changing from a nominal p-value criterion to even a 25% false discovery rate. The follow-up analyses are incomplete with major gaps. The authors focus on single potential mechanisms for cis- and trans-changes, but it is not clear to what degree these mechanisms explain the extent of cis and trans changes. There are also other mechanisms which are not investigated, such as the importance of TF binding sites for cis-regulatory evolution. While likely beyond the scope of this work, communicating these areas for future work would have helped define the niche for this manuscript.

Next, the authors seek to show that differences in DNA methylation are functionally relevant. Consistent with previous results, they show that differences in DNA methylation are (weakly) associated with changes in gene expression. They hypothesize that genes with concordant regulatory elements should exhibit greater methylation-expression coupling than other genes and show that cis-expression/cis-methylation pairs are more strongly correlated than trans/trans pairs. However, I worry that this result could be confounded by larger effect sizes for cis-changes than trans effects. I also think that looking at cis/trans or trans/cis changes would have been useful to directly test the driving hypothesis. Another limitation is that this analysis is limited to promoter regions. It is not clear how many divergent DMRs are included and how many of those genes have differences in expression. The key question is whether differences in DNA methylation are functionally important, and the answer provided by these analyses is "sometimes".

Finally, the authors make a case for lineage-specific selection on DNA methylation that is connected to human traits. This evidence was not convincing. In fact, it is even said that these tests cannot be interpreted as evidence of lineage-specific selection (lines 399-401), so I am confused why these results are framed as testing for selection. The evidence better supports an argument connecting DNA methylation to human phenotypes.

In conclusion, I think this study provides a valuable resource for differences in DNA methylation between humans and chimpanzees across tissues, and provides important insight into the relative abundance of cis and trans regulatory evolution. Additional research is necessary to investigate the underlying regulatory mechanisms, and more care needs to be taken in exploring the functional consequences.

https://doi.org/10.7554/eLife.110575.1.sa2

Reviewer #2 (Public review):

Summary:

This manuscript investigates the causes and consequences of human-specific DNA methylation divergence relative to chimpanzees. The main aim of this study is to disentangle cis- and trans-regulatory contributions to DNA methylation differences, which the authors address using an innovative interspecies hybrid cell system differentiated into multiple cell types. This design allows them to control for trans-acting environments and directly compare allelic regulation.

The authors show that cis-regulatory mechanisms dominate DNA methylation divergence and that methylation-expression coupling is strongest when both are cis-regulated. They further explore potential mechanisms underlying these patterns, including CpG-disrupting mutations and transcription factor-associated trans effects, and identify pathways that may reflect lineage-specific regulatory evolution.

This study provides a valuable dataset and a compelling framework for understanding how local sequence variation contributes to epigenetic and transcriptional divergence, with likely broad impact in comparative and evolutionary genomics.

Strengths:

A major strength of this study is the use of human-chimpanzee hybrid cells, which provides a powerful system to disentangle cis- and trans-regulatory effects in a shared cellular environment. This experimental design allows for a more definitive assessment of regulatory mechanisms than traditional cross-species comparisons.

The study also benefits from the inclusion of multiple differentiated cell types, increasing the robustness and generality of the conclusions. The consistent observation that cis-regulatory mechanisms dominate methylation divergence across these contexts is well supported by both CpG-level and DMR-level analyses.

Another important contribution is the finding that methylation-expression coupling is strongest when both are cis-regulated. This provides a mechanistic explanation for previously observed weak global correlations between methylation and gene expression. Given that the nature of regulatory evolution is likely highly heterogeneous, this study adds valuable insights and guidelines for future investigations. I recommend that the authors provide a list of cis-cis-regulated promoters and their associated genes, which would be a valuable resource for the field.

The application of the two-step sign test identifies biologically relevant pathways, suggesting links between epigenetic divergence and human-specific traits.

The dataset itself, namely, comprehensive DNA methylation and gene expression across multiple cell types in shared cellular contexts, as well as a primary cell type, is a valuable resource for the field. Additionally, the application of the two-step sign test identifies biologically relevant pathways, suggesting links between epigenetic divergence and human-specific traits.

Weaknesses:

Although the authors identify transcription factors associated with differential methylation, it is unclear what proportion of differentially methylated CpGs or DMRs can be attributed to these factors. Providing a quantitative estimate would help assess the relative contribution of trans-acting regulation.

The analysis of CpG-disrupting mutations is interesting but raises two concerns. First, other classes of variants-such as transcription factor binding site-disrupting mutations-could also influence local methylation patterns and are not considered here. Second, the causal direction remains ambiguous: CpG-disrupting mutations may result from methylation-associated mutational processes (e.g., C→T transitions at methylated CpGs) rather than being the primary drivers of methylation divergence. While additional analyses may not be necessary, explicitly acknowledging these alternative explanations would strengthen the interpretation.

Regarding the discussion comparing the distance between CpG-disrupting SNVs and trans-DMRs, without information on the absolute or relative distance distributions, it was difficult to assess the magnitude of the observed differences. Moreover, trans-DMRs, by definition, are not driven by local (cis) variation, and the lack of proximity to CpG-disrupting SNVs is expected. Clarifying what additional insight this analysis provides beyond this expectation may improve this section.

One potential extension would be to examine whether the same cis-acting SNVs are consistently associated with methylation differences across multiple cell types. If these variants are mechanistically causal, one might expect their effects to be reproducible across contexts, or at least more frequent than expected by chance. Such an analysis could further support the proposed link between sequence variation and methylation divergence.

Regarding their two-step sign test analysis, because enrichment-based approaches can sometimes overemphasize statistical significance without reflecting effect size, I wonder if incorporating the magnitude of methylation change would provide additional information or strengthen these findings. While the authors highlight some cases, such as TUBB2 and GRIK, a more general overview and/or integration of effect size into the analysis or discussion would improve interpretability.

https://doi.org/10.7554/eLife.110575.1.sa1

Reviewer #3 (Public review):

Summary:

Ma et al. use human-chimpanzee tetraploid cells to examine species differences in DNA methylation. They identify differentially methylated regions under cis or trans regulation. Cis-DMRs are enriched near SNVs that disrupt or create CpGs, providing a plausible mechanism for cis changes in methylation. They also seek to identify transcription factors that might affect methylation in trans, as well as gene sets with evidence for consistent changes in methylation and expression between humans and chimpanzees.

Strengths:

The authors have generated a new dataset across multiple cell types, examining differences in DNA methylation between humans and chimpanzees using human diploid cells, chimpanzee diploid cells, and human-chimpanzee tetraploid cells. Using this dataset, they identify that cis-DMRs are enriched near SNVs that disrupt or create CpGs compared to trans-DMRs, and identify transcription factors as candidate trans-acting factors. Both identified SNVs and transcription factors are good candidates for future experimentation. Further, they

find that cis-DMRs are more highly correlated with cis-expressed genes than trans-DMRs with trans-expressed genes, providing evidence that methylation and expression are linked genome-wide.

Weaknesses:

The authors could greatly improve the manuscript by focusing on two issues.

(1) Strengthening their cis/trans analysis, including:

a) only showing or analyzing genomic regions that pass FDR correction;

b) clarifying how cis genes are defined (Figure 2B shows some genes labeled as cis where the direction-of-effect differs between hybrid and parent cells);

c) assessing how well powered they are to perform each analysis.

(2) Softening claims about human evolution or human specificity for several reasons:

a) Their comparison lacks tetraploid controls (e.g. human-human tetraploids and chimp-chimp tetraploids) or experimental follow-up in diploid cells, making it hard to be certain that observed effects are not due to ploidy.

b) There are no outgroup species included in the analysis.

c) The use of no or very loose FDR corrections with the sign test makes it difficult to draw conclusions.

d) Experimental data to link SNVs to changes in cis methylation or identified transcription factors to changes in trans methylation would be needed to validate the authors' predictions.

https://doi.org/10.7554/eLife.110575.1.sa0