ARTICLE

www.the-innovation.org

Comparative analysis reveals epigenomic evolution related to species traits and genomic imprinting in mammals

Yisi Hu,[1,2,3] Shenli Yuan,[4] Xin Du,[1,2] Jiang Liu,[4] Wenliang Zhou,[3] and Fuwen Wei[1,2,3,*]

*Correspondence: weifw@ioz.ac.cn

Received: February 26, 2023; Accepted: April 25, 2023; Published Online: April 28, 2023; https://doi.org/10.1016/j.xinn.2023.100434

GRAPHICAL ABSTRACT

PUBLIC SUMMARY

- Changes in promoter methylation are related to trait evolution.
- Genomic imprinting accumulated in a stepwise fashion during mammalian evolution and mostly functions in embryonic development.
- Epigenomic evolution should be incorporated to develop a unified evolutionary theory.

ARTICLE

Comparative analysis reveals epigenomic evolution related to species traits and genomic imprinting in mammals

Yisi Hu,[1,2,3] Shenli Yuan,[4] Xin Du,[1,2] Jiang Liu,[4] Wenliang Zhou,[3] and Fuwen Wei[1,2,3,*]

[1]CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

[2]University of Chinese Academy of Sciences, Beijing 100049, China

[3]Center for Evolution and Conservation Biology, Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China

[4]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China

*Correspondence: weifw@ioz.ac.cn

Received: February 26, 2023; Accepted: April 25, 2023; Published Online: April 28, 2023; https://doi.org/10.1016/j.xinn.2023.100434

Citation: Hu Y., Yuan S., Du X., et al., (2023). Comparative analysis reveals epigenomic evolution related to species traits and genomic imprinting in mammals. The Innovation **4(3)**, 100434.

DNA methylation is an epigenetic modification that plays a crucial role in various regulatory processes, including gene expression regulation, transposable element repression, and genomic imprinting. However, most studies on DNA methylation have been conducted in humans and other model species, whereas the dynamics of DNA methylation across mammals remain poorly explored, limiting our understanding of epigenomic evolution in mammals and the evolutionary impacts of conserved and lineage-specific DNA methylation. Here, we generated and gathered comparative epigenomic data from 13 mammalian species, including two marsupial species, to demonstrate that DNA methylation plays critical roles in several aspects of gene evolution and species trait evolution. We found that the species-specific DNA methylation of promoters and noncoding elements correlates with species-specific traits such as body patterning, indicating that DNA methylation might help establish or maintain interspecies differences in gene regulation that shape phenotypes. For a broader view, we investigated the evolutionary histories of 88 known imprinting control regions across mammals to identify their evolutionary origins. By analyzing the features of known and newly identified potential imprints in all studied mammals, we found that genomic imprinting may function in embryonic development through the binding of specific transcription factors. Our findings show that DNA methylation and the complex interaction between the genome and epigenome have a significant impact on mammalian evolution, suggesting that evolutionary epigenomics should be incorporated to develop a unified evolutionary theory.

INTRODUCTION

DNA methylation at cytosines (5-methylcytosine, 5mC) occurs in most eukaryotic groups, with evolutionarily variable levels and genomic distributions.[1,2] In mammals, it is a key epigenetic modification involved in a variety of biological processes, including the regulation of gene expression, regulation of chromatin structure, repression of transposable elements, X chromosome inactivation, and genomic imprinting.[3] Numerous studies have demonstrated that hypermethylation in the promoter typically results in decreased transcription of downstream genes by blocking transcription factor binding and altering chromatin states.[4,5] In addition, 5mC is epigenetically heritable through mitotic cell divisions and is thus conserved in somatic tissues[6,7]; it is also transgenerationally heritable at certain loci and may persist over evolutionary time.[8,9]

However, studies on DNA methylation in mammals have focused mainly on mice and primates. They have revealed the basic pattern of extensively methylated cytosines (~70%) in the context of CpG and epigenetic dynamics during human evolution.[10–16] Less is known about the roles of DNA methylation in gene evolution under selective forces and in the development of mammalian traits, especially in root lineages such as marsupials.

Moreover, genomic imprinting in mammals, a phenomenon in which genes are expressed according to parental origins, is also intimately associated with differential DNA methylation in imprinting control regions (ICRs) in the two parental genomes.[17] Parental epigenetic memory in ICRs can withstand genomic reprogramming during early embryogenesis and maintain allele-specific methylation in somatic cells, which controls the allele-specific expression of a single or multiple neighboring imprinted genes.[18] Due to the relative scarcity of genomes and epigenomes in non-model animals, examining these allelic DNA methylation marks at whole-genome scales and across the mammalian phylogeny remains challenging, and many questions persist about the evolutionary origins and functions of genomic imprinting in mammals.

In recent years, omics advances have provided a strong foundation for comparative epigenomics research, enabling the comprehensive investigation of epigenomic evolution in mammals.[19] In this study, we utilized comparative epigenomic data from 13 mammalian species to explore the epigenetic regulation of mammals from an evolutionary standpoint, while accounting for mammalian phylogenetics. Our aim was to investigate the relationship between promoter methylation and coding sequence molecular changes, as well as explore the evolutionary significance of DNA methylation in promoters and noncoding sequences for species trait evolution, to draw connections between regulatory adaptations and physiological or morphological phenotypes. In addition, we explored the evolutionary conservation of allelic methylation associated with genomic imprinting, specifically for the stepwise accumulation of 88 known ICRs. We also conducted the *de novo* identification of allelic methylation regions (AMRs) in each species and analyzed their genomic distribution, gene functional enrichment, and the motif enrichment of potential imprints.

Based on a combination of comparative genomics and comparative epigenomics approaches, we could investigate a relatively unexplored research avenue in the evolution of mammalian epigenomics: the conservation and divergence of mammalian epigenomes and the potential connections between these epigenomic patterns and gene evolution, species trait differentiation, and genomic imprinting. Our findings could prove foundational for a better understanding of systematic evolutionary theories that incorporate both genomic and epigenomic evolution.

RESULTS

WGBS of 13 mammalian species

We generated and collected single-base-resolution methylome data by whole-genome bisulfite sequencing (WGBS) for 13 mammal species spanning the mammalian phylogenetic tree, representing eight orders and 13 families (human [*Homo sapiens*], macaque [*Macaca mulatta*], mouse [*Mus musculus*], cow [*Bos taurus*], bottlenose dolphin [*Tursiops truncatus*], pig [*Sus scrofa*], horse [*Equus caballus*], giant panda [*Ailuropoda melanoleuca*], dog [*Canis lupus familiaris*], greater horseshoe bat [*Rhinolophus ferrumequinum*], Linnaeus's two-toed sloth [*Choloepus didactylus*], red kangaroo [*Osphranter rufus*], and sugar glider [*Petaurus breviceps*]; see Figure 1A for phylogenetic relationships, Tables S1 and S2 for detailed information on the WGBS data). Among these species, five were sequenced to examine base-resolution DNA methylation for the first time. Liver and skeletal muscle samples from two different individuals were utilized as replicates in all species with the exception of the bottlenose dolphin, for which only one muscle sample could be obtained. The CpG methylation levels of the two replicates were significantly correlated across all species ($p < 2.2 \times 10^{-16}$ for all species, Pearson correlation coefficients from 0.64 to 0.82, Figure S1). The genome tracks of methylation profiles in the two replicates are provided in Figure S2.

Although previous studies have reported tissue differences in DNA methylation both within the same species and between closely related species,[10,15,20,21] our analysis showed that the hierarchical clustering of samples based on the methylation levels of syntenic CpG sites or syntenic regions longer than 100 base pairs (bp) both grouped the same species tightly together (Figure S3). The only exception was the clustering of human and macaque samples, where DNA methylation was more consistent between the same tissue across the two species, in line

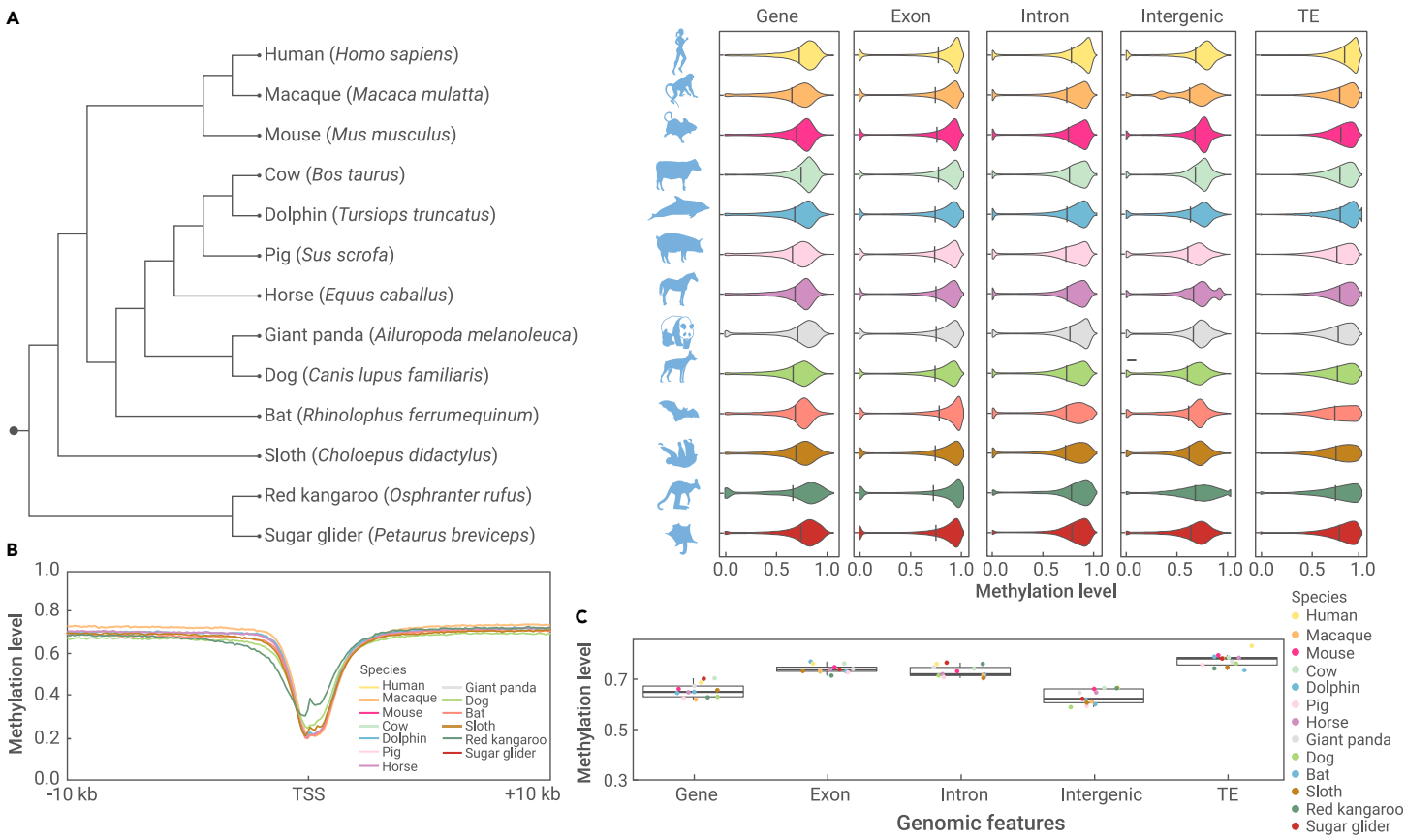

Figure 1. Interspecies conservation of DNA methylation among various genomic features (A) Each violin plot depicts the distribution of the DNA methylation levels in the genomic regions of gene bodies, exons, introns, intergenic regions, and TEs in the species that we analyzed, as shown in the phylogenetic tree. (B) This figure shows the methylation levels of regions 10 kb upstream and 10 kb downstream of the TSS of genes in each species calculated using 100 bp bins, indicating that the regions around the TSS have the lowest methylation levels. (C) This figure shows the average methylation levels of each genomic feature in each species.

with previous research.[10,15] The closely clustered epigenomic structure of tissues within most species suggests that species characteristics take precedence over tissue specificity on the wider scale of mammalian evolution. Intriguingly, the clustering dendrograms based on the methylation levels of CpG sites or syntenic regions both resembled the phylogenetic tree to some extent (species from Primates, Cetartiodactyla, Carnivora, and Marsupialia were clustered together, respectively) but also differed significantly for distantly related species, indicating that changes in DNA methylation may follow phylogeny for closely related species but evolve faster than genetic variations after species diverge substantially. This also implies that the epigenomic landscape is influenced by other factors in addition to genetic inheritance.

DNA methylation patterns of genomic features across mammals

The DNA methylation patterns of gene bodies, exons, introns, intergenic regions, and transposable elements (TEs) were largely comparable across all species examined (Figure 1A) and between replicates from the same species (Figure S4). TEs were the most heavily methylated among these genomic features, with an average methylation level of 77% (Figure 1C). We found that Alu elements, a family of short interspersed nuclear elements that arose recently in primate ancestors, were significantly more methylated than older TE families ($p < 0.05$, Kruskal-Wallis test, Figure S5). This supports the importance of DNA methylation in the epigenetic regulation of TE activity. In contrast, the methylation levels around transcription start sites (TSSs) were the lowest, compared with other genomic regions (Figure 1B), with promoter regions showing a bimodal distribution pattern of methylation levels (Figure S6), in accordance with their function in controlling gene expression. Overall, the genomic distribution of DNA methylation was relatively conserved in mammals.

Promoter methylation evolves in synchrony with gene evolution

We identified 10,219 single-copy orthologous genes in at least 11 eutherian mammals. For each pair of orthologous genes from two species, promoter methylation divergence (see material and methods for details) was significantly positively correlated with synonymous substitution rates (dS) of the coding regions, which represents the evolutionary distance of each gene pair ($p < 2.2 \times 10^{-16}$, Figure 2A). This finding is consistent with previous research on primate orthologous gene pairs and human duplicated gene pairs,[12,22] suggesting that interspecies promoter methylation differences are correlated with interspecies changes in coding sequences.

We further investigated whether genes under distinct selective constraints showed variable degrees of promoter methylation. Indeed, there was a strong positive correlation between the average dN/dS ratio (non-synonymous substitution rate over synonymous substitution rate, a measure of the evolutionary pressure on the gene) of the orthologous gene in all species and the average promoter methylation level ($p < 2.2 \times 10^{-16}$, Figure 2B). The promoters of genes with low dN/dS ratios remained constitutively unmethylated across species, and these genes were usually evolutionarily conserved under strong purifying selection and may possess essential functions and active expression. In contrast, genes with higher dN/dS values among all species, which indicate a greater accumulation of non-synonymous mutations, are typically more likely to have high DNA methylation levels in their promoters. One possible driving force underlying the synchronized evolution of promoter methylation and coding sequence molecular evolution is the ability of epigenetic gene silencing to alleviate the impacts of genetic changes on drastic shifts in gene expression and immediate phenotypic effects while maintaining phenotypic flexibility. Another explanation is that sequence change is a result of the epigenetic silencing of genes, which relaxes the selection pressure on genes. These results together indicate that gene evolution and epigenomic evolution co-occur to regulate gene activities.

Lineage-specific methylation patterns of promoters and noncoding sequences are associated with species traits

Previous studies have shown that DNA methylation in one tissue could predict gene expression in other tissues.[23] Therefore, we utilized two samples from different developmental origins (endoderm and mesoderm) to account for most intraspecies differences in DNA methylation and represent the epigenomic

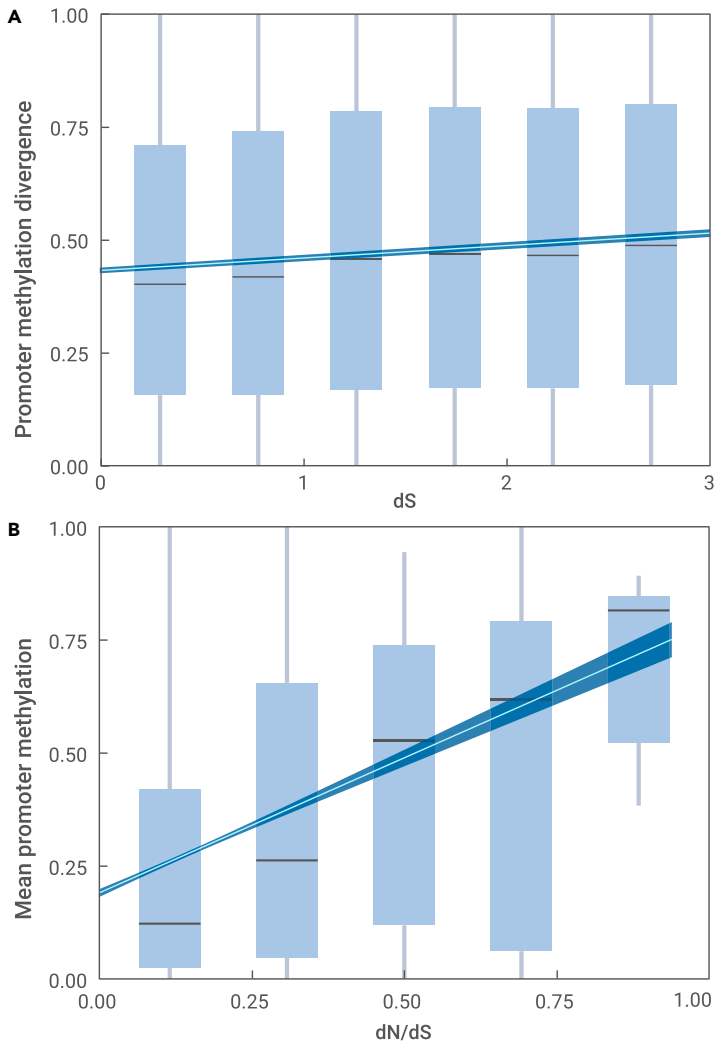

Figure 2. Epigenomic evolution of promoter methylation in synchrony with the molecular evolution of genes (A) The boxplot shows the distribution of promoter methylation divergence values for ~660,000 orthologous gene pairs with dS < 3, divided into six bins by dS values. The regression line with the 95% confidence interval demonstrates a significant correlation ($p < 2.2 \times 10^{-16}$) between promoter methylation divergence and dS of all gene pairs, indicating that gene pairs with larger dS values and higher evolutionary distances tended to differ in promoter methylation. (B) The boxplot shows the distribution of mean promoter methylation levels for more than 10,000 genes, divided into five bins by dN/dS values. The regression line with the 95% confidence interval demonstrates a significant correlation ($p < 2.2 \times 10^{-16}$) between mean promoter methylation and dN/dS of all these genes, indicating that genes with lower evolutionary rates tend to have lower mean promoter methylation levels.

profiles throughout the body in each species. Across all the orthologous genes examined, 12% of them had consistently hypomethylated promoters across species, whereas 5% had consistently hypermethylated promoters across species (Figure 3A). The genes with hypomethylated promoters showed significant gene ontology (GO) enrichment in important developmental processes, including the development of cells (such as cell morphogenesis, cell fate commitment, and negative regulation of cell differentiation), development of tissues and systems (tissue morphogenesis, head development, brain development, limb morphogenesis, and reproductive system development) as well as the development of embryos and entire organisms (embryonic morphogenesis, pattern specification process, and regionalization) (Figure 3B and Table S3). These genes with permissive epigenetic states are involved in developmental processes at various scales and are most likely activated to play essential roles in organism survival. In contrast, genes with heavily methylated promoters were primarily associated with immune system processes (Table S4). DNA hypermethylation targeting these genes likely maintains their repressed state in most somatic lineages, while they might be specifically turned on in the immune system.

A total of 2,701 genes exhibited species-specifically upregulated or downregulated methylation levels in promoters. In addition, another 145 gene promoters displayed altered methylation patterns that were shared by marsupials (Figure 3A). The lineage-specific altered methylation levels in promoters may reflect a different schema of gene expression, which might be associated with or result in phenotypic outcomes of evolutionary significance. We also identified differentially methylated nonpromoter and noncoding regions in each species and found that genes with altered promoter methylation patterns and the differentially methylated noncoding regions of species were enriched in many identical or closely related functions related to species-specific traits (Figures 3C and S7). For instance, genes in cows with altered promoter methylation as well as noncoding regions were significantly involved in digestive tract development and the response to alkaloid, which are relevant to the specialized anatomy of the ruminant digestive tract and diet adaptations. Dog-specific enriched biological processes include the development of sensory organs (ear and olfactory organs), sensory processing, and olfactory behavior. In pandas, differentially methylated genes played roles in pigmentation and embryo implantation, and the differentially methylated noncoding regions were related to *in utero* embryonic development. This may be correlated with obligate delayed implantation in the giant panda, as they produce underdeveloped cubs with the lowest neonate-maternal weight ratio among all eutherians.[24] Of special note were the genes showing alterations of promoter and noncoding region methylation unique to bats, which were significantly enriched in immune-related processes, such as innate immune response, modulation by host of symbiont transcription, regulation of viral entry into host cells, and regulation of viral genome replication. Limb development and embryonic digit morphogenesis were also enriched, and the differentially methylated genes include *SHH*, a key factor in limb development examined in vertebrates,[25–28] which has been reported to specifically influence bat digit development in previous comparative transcriptomic research.[29,30] Our finding

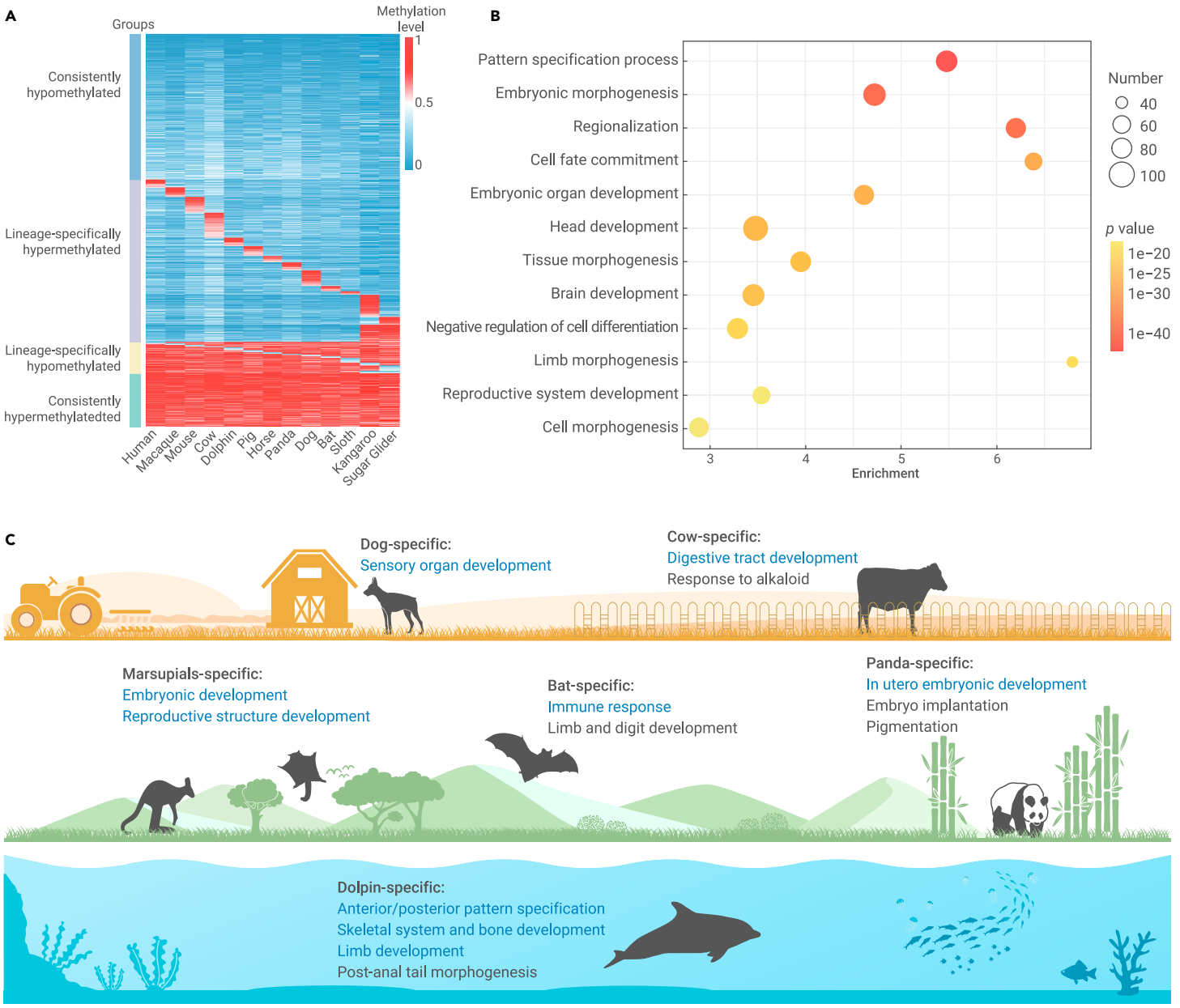

Figure 3. Promoter methylation patterns of orthologous genes are associated with important developmental processes and species traits (A) This heatmap shows the methylation levels of four groups of gene promoters: consistently hypomethylated, lineage-specifically hypermethylated, lineage-specifically hypomethylated, and consistently hypermethylated. (B) The functional analysis of consistently hypomethylated gene promoters reveals the enrichment of developmental processes, including cell development, tissue and system development, embryo development, and organism morphogenesis. (C) Promoters and noncoding sequences with lineage-specific altered methylation patterns exhibit enriched GO annotations (selected from all significant results) associated with species traits. The biological processes in black letters are enriched in the promoters, whereas the biological processes in blue letters are enriched in both the promoters and noncoding sequences.

of bat-specific altered methylation of the *SHH* promoter may contribute to the mechanism of the differential expression of this gene during development and assist in understanding the adaptive evolution of bat wings. In contrast to terrestrial mammals, dolphins exhibit a streamlined body shape and numerous skeletal adaptations facilitate swimming, including a telescoped skull, hindlimb loss, the flattening of the forelimbs into flippers, a large number of short vertebrae, and so forth.[31] Interestingly, dolphin-specific altered genes and noncoding regions were found to be enriched in developmental processes related to these features, including anterior/posterior pattern specification, skeletal system development, osteoblast differentiation, bone development, post-anal tail morphogenesis, and limb development. Among these, *SCUBE3* had a dolphin-specifically hypermethylated promoter, and this gene has been reported to cause malformations of vertebrae and shorter thigh bones in mutant mice.[32] The two marsupial species shared many specific modifications in gene promoters and noncoding regions, and these changes were involved in reproductive processes such as embryonic placenta development, post-embryonic development, reproductive structure development, mammary gland development, and so forth. These are related to the most distinct features between marsupials and eutherians, since marsupials typically exhibit a short gestation time but a long and sophisticated lactation period to nurture their highly altricial newborns.[33] These findings may help elucidate the regulatory mechanisms by which the mammary gland compensates for the functions of the short-lived, simple placenta of marsupials.[34] In addition, the reproductive system also evolved differently in marsupials and eutherians, as marsupial females have three vaginas and two separate uteri and males have a bifurcated penis. The epigenetic differences among these orthologous genes involved in the gene regulatory network of reproduction might partially explain the distinct discrepancies between marsupials and eutherians.

The correlations of DNA methylation and species traits were apparent in promoters and noncoding regions but not in the coding sequences of the genes, as very few genes exhibited altered methylation patterns in their coding sequences (less than 0.7% of orthologous genes in each species on average), and they were not significantly enriched in any of the above biological processes associated

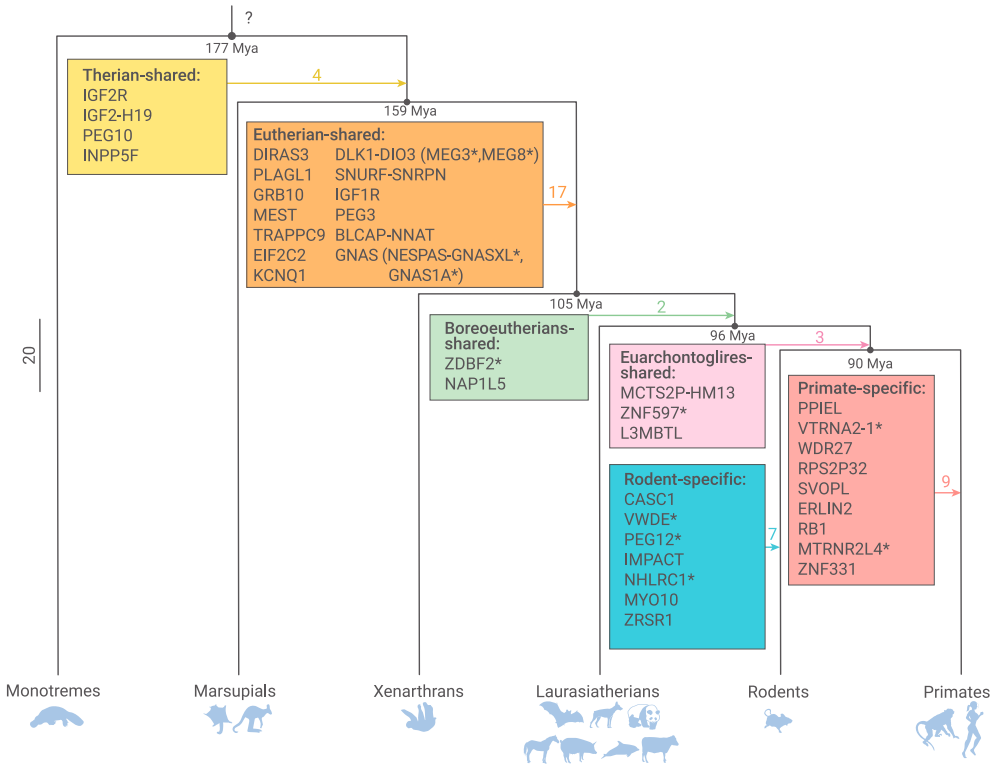

Figure 4. Evolutionary origins of documented human and mouse ICRs Arrows indicate that ICRs may arise during the evolution of the branch. Somatic ICRs are marked by an asterisk ("*"), and the rest are germline ICRs. Two somatic ICRs (MEG3 and MEG8) are located within the imprinting cluster of the germline ICR DLK1-DIO3, and two others (NESPAS-GNASXL, GNAS1A) are located within the cluster of the germline ICR GNAS. The divergence time of each node was extracted from TimeTree.[96]

with species traits. The lineage-specific differentially methylated noncoding regions may act as *cis*-regulatory elements to control the expression of neighboring genes. In fact, approximately 30% of these regions overlap with the conserved noncoding elements (CNEs) that we identified across all species, and they cover 2,576 CNEs with lineage-specific altered methylation patterns. Together, these results demonstrate that the DNA methylation of promoters and noncoding elements, but not coding sequences, can coordinately act on multiple levels of the regulatory cascades of mammals to facilitate many biological processes during development, hence potentially influencing or reflecting a variety of complex lineage-specific phenotypes.

Evolutionary history of known ICRs reveals stepwise accumulation of mammalian imprints

A fundamental process in establishing ICRs in mammals, which *cis*-regulate the allelic expression of neighboring genes in imprinting clusters, is the methylation of allelic DNA on one of the paternal chromosomes. Given the wealth of research on human and mouse imprinted genes, we collated information on 88 documented ICRs from humans and mice with expression data as a starting point for reconstructing the evolutionary history of imprinting in mammals (Table S5).[35–38] These ICRs included 53 known human ICRs, 56 known mouse ICRs, and 23 shared ICRs. We examined the overlap between the syntenic regions of the known ICRs and the *de novo* identified AMRs in each species. Most of the known ICRs were located in imprinting clusters with clear evolutionary origins that could be inferred from the methylation profiles of the species we studied (Figure S8). These known ICRs were grouped into the following categories: four therian-shared ICRs, 17 eutherian-shared ICRs, two Boreoeutherians-shared ICRs, three Euarchontoglires-shared ICRs, nine primate-specific ICRs, and seven rodent-specific ICRs (Figure 4). As epigenetic regulation may change as species evolve, leading to the loss or relaxation of DNA methylation marks at these loci, methylation patterns may not be strictly consistent in all species. In addition, our data may not be able to detect all AMRs in all species. Thus, we regarded the remaining documented ICRs with divergent or inconsistent methylation patterns relative to phylogeny to be non-conserved or data deficient, and we did not make inferences about their evolutionary history.

The evolutionary histories of the imprinted domains exhibited a stepwise accumulation process in mammals. Marsupials diverged with eutherians approximately 159 million years ago and do not appear to share many ICRs with eutherians. Three of the imprinting domains shared by marsupials and most eutherians are vital for embryonic and placental development (*PEG10*, *IGF2-H19*, *IGF2R*).[39–41] A large proportion of the imprinting domains arose during the early stage of eutherian evolution (i.e. before the divergence of Xenarthrans and Boreoeutherians), and they control the expression of a number of genes involved in cell proliferation, the cell cycle, cell differentiation, cell motility, and cell apoptosis (*DIRAS3*, *PLAGL1*, *GRB10*, *TRAPPC9*, *KCNQ1*, *DLK1*, *DIO3*, *MEG3*, *IGF1R*, *PEG3*, and *BLCAP*). These genes may shape some distinctive features shared by all eutherians, most strikingly the intricate maternal-fetal communication through a well-developed placenta during gestation, which may benefit from the parental-specific expression of development-related genes. In contrast, imprinted genes that emerged later in primates or rodents have more specialized functions and may be related to more specific features. For example, aberrant methylation of the primate-specific imprinted gene *PPIEL* is associated with intellectual disability and bipolar disorder,[42] whereas the mutation of another imprinted gene, *WDR27*, may cause developmental abnormalities of the brain.[43] In addition to the known human and mouse ICRs, other species have likely evolved their own imprints, as will be covered later.

Germline ICRs are more conserved than somatic ICRs

Using methylome data from human and mouse gametes and early embryos,[35,44] we classified ICRs as germline ICRs (also known as primary ICRs) and somatic ICRs (or secondary ICRs). Germline ICRs inherit allelic methylation from germ cells and are maintained after fertilization throughout development, whereas somatic ICRs are only established during early embryonic development.[35] We found that a larger proportion of the germline ICRs were shared at least among Boreoeutherians (emerged about 105 Mya), indicating that germline ICRs are significantly more phylogenetically conserved than somatic ICRs ($p < 0.05$, chi-square test). Indeed, many of the known somatic ICRs of humans and mice showed variable methylation patterns across species and even between different individuals or tissues of the same species, as demonstrated not only by our analysis but also by other studies.[45–47] For instance, we found in the longest contiguous allelic methylated regions in the human genome (once dysregulated, could lead to Prader-Willi syndrome),[48] the germline ICR at the *SNRPN-SNURF* locus was likely established in the ancestor of eutherians, whereas the methylation patterns of several somatic ICRs near the protein-coding genes *MKRN3*, *MAGEL2*, *NDN*, and long noncoding RNA (lncRNA) genes, varied considerably between species and even between the replicates of some species (Figure S9). Since only germline ICRs are established with differential methylation in maternal and paternal germ cells during gametogenesis, the less conserved somatic ICRs at surrounding loci may acquire allele-specific methylation in a germline ICR-dependent manner or as a result of germline ICR expansion. This distinction between germline ICRs and somatic ICRs in terms of conservatism is critical for determining their roles in embryonic development or in specific tissues.

Genomic imprinting mainly affects embryo development in mammals

To further understand the evolution of genomic imprinting in all mammalian species, we filtered the AMRs from prior analyses using more stringent criteria (see material and methods), only including AMRs supported by both replicates

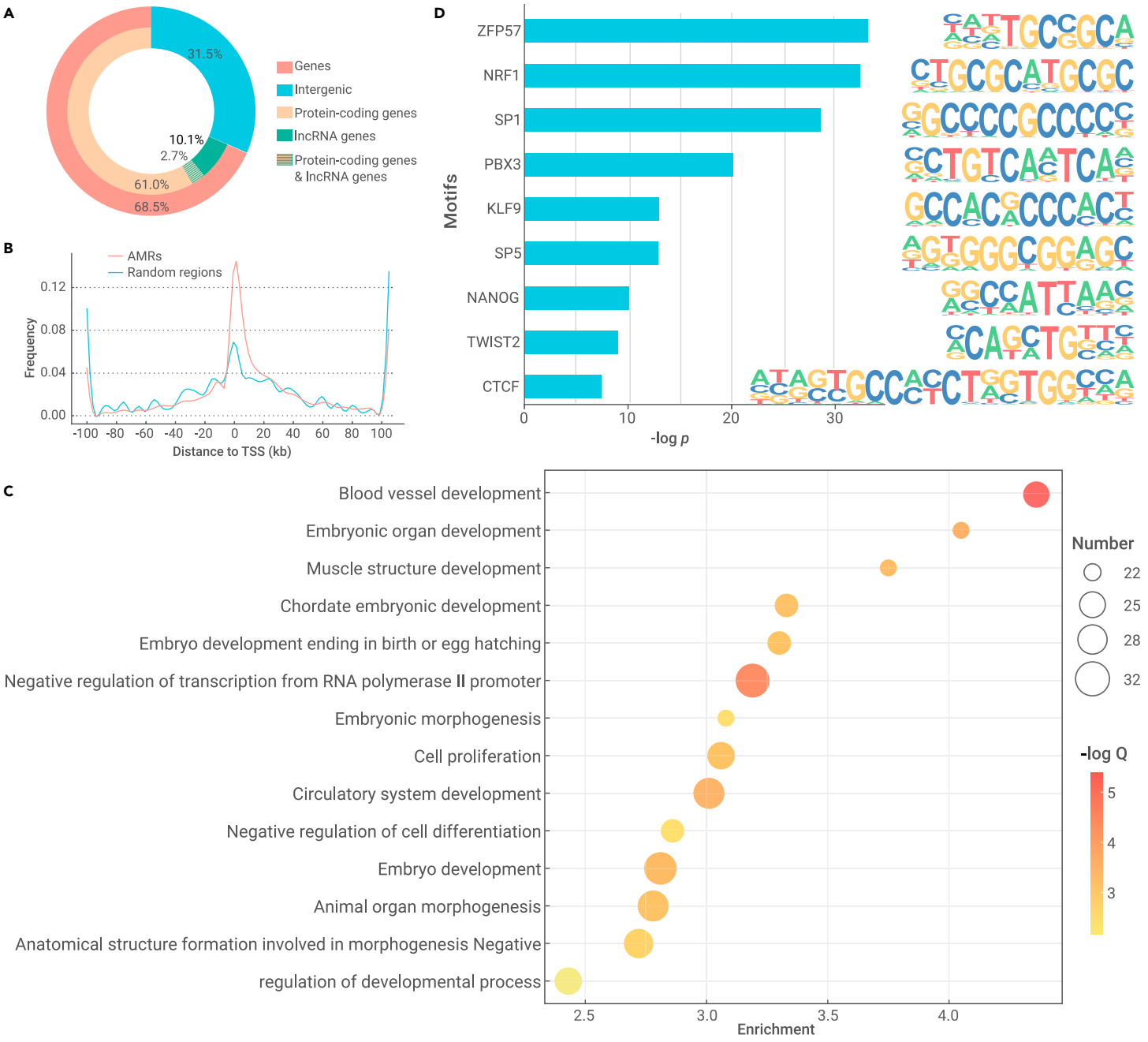

Figure 5. Genomic distribution, functional enrichment, and motif enrichment of AMRs (A) The pie chart indicates the proportions of AMRs that overlapped with genes (including protein-coding genes, lncRNA genes, and both) and intergenic regions. (B) The distances of the AMRs to the nearest transcription start sites of genes (black line) were significantly closer compared with those of random regions of similar length (gray line). (C) Gene functions enriched in the commonly shared AMRs were mostly associated with embryonic development. (D) Motifs enriched in the commonly shared AMRs were associated with development and gene expression regulation.

in each species. Thus, we identified the highly credible AMRs, which were putatively ICRs in each species, without referring to known imprints. Not surprisingly, the highly credible AMRs list also retained those conserved known ICRs described above, such as the ICRs at *GNAS*, *PEG10*, *PEG3*, *SNRPN-SNURF*, *MEG3*, and *BLCAP-NNAT*, demonstrating the effectiveness of our strategy in locating potential ICRs in other mammalian species.

Nearly 70% of the highly credible AMRs overlapped with genes, including 61.0% overlapping with protein-coding genes, 10.1% overlapping with lncRNA genes, and 2.7% overlapping with both (such as AMRs at *DIRAS3*, *GNAS*, *SNURF*, *MEST*, and *IGF2R*) (Figure 5A). Compared with random autosomal regions of similar length, the AMRs were significantly enriched around the TSS of genes ($p = 3.8 \times 10^{-13}$, chi-square test), indicating their important activities in regulating gene expression in *cis* (Figure 5B). Ideally, the imprinted genes with allele-specific expression could be identified based on heterozygous single nucleotide polymorphisms (SNPs) from transcriptome data. Here, due to the dynamics of gene expression and the lack of hybrids of diverged strains for most species, we only examined the relationship between AMRs and allele-specific expression genes (ASEGs) in several species with available data, including human, mouse, and pig. We found that the highly credible AMRs we identified overlapped with a large number of verified ASEGs in public datasets,[49] including 163 out of 226 human AMRs that overlapped with 181 ASEGs. For example, one human AMR located at the *ZNF597/NAA60* locus (chr16: 3442660–3445364) overlapped with the ASEGs *ZNF597* and *NAA60* in 125 independent human transcriptomes from 12 different cell lines and embryonic tissues; another AMR (chr22: 41681733–41683152) overlapped with *SNU13*, which is allelic expressed in 118 independent human transcriptomes from nine different sources. The allelic methylation patterns of these AMRs are shown in Figure S10. Similarly, 800 out of 1,356 mouse AMRs we identified overlapped with 723 ASEGs supported by 309 samples from at least 24 cell types. In addition, 405 out of 660 pig AMRs overlapped with 391 ASEGs from six available muscle transcriptomes. The full

lists of AMRs overlapping with ASEGs are provided in Tables S6–S8. The percentage of AMRs that overlap with the ASEGs found in other independent studies was similar to the percentage of AMRs that directly overlap with genes in all species (~70%), supporting that they potentially control the allelic expression of the imprinted genes, whereas the other AMRs located in intergenic regions may regulate gene expression as distal regulatory elements.

The functional enrichment of highly credible AMRs shared by at least three species revealed that these regions were significantly enriched in GO biological processes related to embryo development, embryo morphogenesis, cell differentiation, and regulation of transcription (Figure 5C). Additionally, these regions were significantly enriched in imprinting and embryonic development according to the mouse phenotype database (Figure S11). These findings suggest that the primary function of imprinting in mammals is to act as a key regulator of embryonic development, not just for the dozens of known imprints but also for the potential imprints in other species that have been subject to little research.

The highly credible AMRs were also enriched in specific transcription factor binding sites involved in the establishment and functioning of genomic imprinting (Figure 5D). These included the binding sites for ZFP57 and CTCF, two DNA binding factors that have been reported to regulate the allelic expression of imprinted genes. ZFP57 and its cofactor KAP1 could bind selectively to the methylated alleles at the ICRs to maintain asymmetric histone modifications and DNA methylation during early embryogenesis.[50,51] CTCF is reported to regulate genomic imprinting as a distal regulatory element through differential allelic binding at ICRs to organize the topologically associating domains (TADs) that insulate genes and their regulatory elements, which is best illustrated by the *IGF2/H19* locus.[52,53] The other enriched motifs, including NRF1, SP1, PBX3, SP5, NANOG, and TWIST2, are engaged in the expression regulation of cell growth or organ development, helping to elucidate the mechanisms whereby genomic imprinting influences embryo development.

Novel imprint origination is associated with more specific functions

We further identified species-specific AMRs to facilitate the understanding of the mechanisms of novel imprint origination. On average, 9.8% of the highly credible AMRs in each species were species-specific, and 59.9% of them overlapped with verified ASEGs in humans, mice, and pigs in public datasets.[49] As species-specific AMRs likely emerged more recently, their functions may differ from those AMRs shared by multiple taxa. We found that genes overlapping with species-specific AMRs were enriched in various biological processes but not in embryo development. The most frequently enriched GO terms were brain development, head development, regulation of neuron projection development, synapse organization, and behavior, all of which are related to the less-studied functions of imprinting in influencing gene expression in brains and behavior modes from the postnatal period to adulthood.[54,55] Thus, species-specific AMRs imply more specialized roles of imprinting during mammalian evolution that may influence brain function and behavior by affecting neurodevelopment.

DISCUSSION

To the best of our knowledge, this study presents the first comprehensive examination of epigenomic evolution spanning from marsupials across the major lineages of eutherian mammals using WGBS data. Several species were sequenced to examine DNA methylation for the first time, including two evolutionarily distinct marsupials. We found that the distribution of DNA methylation among genomic features was evolutionarily conserved in mammals and that the evolution of promoter methylation was significantly correlated with the molecular evolution of coding sequences. Through comparative analyses, our data also highlight both the conservation and divergence of promoter and noncoding element methylation as well as their potential roles in the evolution of traits. In addition, we performed analyses on the evolutionary origins of known imprints and the *de novo* identification of putative imprints in mammals, revealing their common functions in embryonic development through several transcription factor binding motifs.

Notably, we found that while the DNA sequences of orthologous genes and noncoding elements were mostly conserved, differing methylation levels across species at particular loci suggested that epigenomic state was decoupled from sequence conservation, leading to lineage-specific innovations in epigenetic regulation. These evolutionary innovations could result in alterations in gene expression regulation, which is a fundamental mechanism leading to the emergence of functional and structural novelty.[56] The placenta, for example, has been proposed as a good candidate for neo-organ origination research.[57] Here, we discovered that marsupials exhibit altered methylation patterns related to placentation and mammary glands, possibly helping to explain why marsupial and eutherian placentas display transcriptional differences; this could help us understand how marsupial mammary glands employ genes expressed in eutherian placentas to perform similar functions in supporting offspring development.[34,58] Thus, the evolution of placentation and embryogenesis that differentiates marsupials and eutherians may be partially due to or reflected by divergent strategies of epigenetic regulation.

Although our study demonstrates the potential for exploring DNA methylation in promoters and noncoding elements to reveal the regulatory underpinnings of phenotypic diversity and adaptive evolution in mammals, we were unable to determine when these changes initially occurred over the course of evolution. Questions also remain concerning the causes of epigenetic changes—whether they result from the expression dynamics of active genes or the effects of repressed genes, and whether and how other epigenetic marks, such as histone modifications, are involved in the process of epigenomic evolution. Studies focused on other types of epigenetic modifications are beginning to address these questions,[59–62] but there is still much to learn in this emerging and exciting field of research. With additional data integrating DNA methylation, histone marks, chromatin interaction, transcription, and genomics across taxa and tissue types, we can more carefully investigate how epigenetic marks precisely function in the evolution of complex traits.

By determining the origins of known imprints and performing *de novo* identification of potential imprints, our findings also help address hypotheses related to imprinting, such as the parental conflict hypothesis,[63] the coadaptation hypothesis,[64] and the host defense hypothesis,[65] all of which attempt to understand the evolutionary driving forces of imprinting evolution and the mechanisms of its origination. The parental conflict hypothesis proposes that the conflict between paternal and maternal interests in resource allocation to offspring drives the differential expression of paternal and maternal genes. The coadaptation hypothesis, however, suggests that imprinting evolved as a result of the coadaptation of the maternal hypothalamus and placenta, which governs the mother-infant relationship. These two hypotheses were first posed based on the functions of a limited number of imprinted genes (*IGF2*, *IGF2R*, *GRB10*, *PEG3*, etc.). Particularly, *IGF2* and *IGF2R* are expressed from different parental origins and act in synergy to control embryo size, which is assumed to be vital for regulating *in vivo* embryonic growth in mammals.[63] Our results concerning putative imprints on a larger scale confirm the critical role of imprinting in regulating embryonic development via a number of transcription factors involved in development. The hypothesis of genomic imprinting evolution in mammals also ties the phenomenon to viviparity due to its putative absence in monotremes (platypus and echidna).[66] Unfortunately, WGBS data are not available for monotremes, limiting our ability to conduct comparative epigenomic analysis of these ancient and enigmatic animals. In addition to the question of origination, it is also challenging to identify the evolutionary forces leading to imprint gains and losses during species evolution. However, the species-specific AMRs that we identified suggest that these AMRs may be coupled with more specific functions in brain development and behavior, rather than embryonic development. This could be explained by a model based on bet-hedging trade-offs between reproductive mean and variance that results in differing behavioral strategies in males and females under the influence of brain-expressed imprinted genes.[67] The variable reproductive modes in mammals may drive the evolution of lineage-specific imprinted genes that are expressed in the brain and contribute to specific cognitive and behavioral phenotypes.

CONCLUSION

By combining these findings, we conclude that epigenomics plays a crucial yet underappreciated role in evolution. Specifically, the epimodification of DNA methylation may not only evolve along with molecular changes in genes, but also affects gene expression to generate different phenotypes. Numerous prior studies have shown that the 5mC modification can directly facilitate nucleotide substitution, as it is more mutagenic than cytosine.[68–70] Several studies have consolidated the impacts of epigenomics on adaptation, speciation, species symbiosis, and so forth,[71–77] reiterating the notion that the research field of epigenomic evolution can shed new light on a range of evolutionary phenomena.

Our comparative analysis framework and methodology could be extended to other taxa, possibly including microorganisms, plants, other animals, and especially non-model organisms, to disentangle the mysteries of epigenomics throughout all manner of life forms. We suggest that evolutionary epigenomics should be more thoroughly and consistently incorporated into existing evolutionary theories to develop a more comprehensive and unified theoretical framework. In addition, future research that strengthens the interdisciplinary crossover and integration of epigenomics, genomics, and evolutionary biology will further our understanding of why and how life evolved into its current multitude of forms.

MATERIALS AND METHODS

WGBS of 13 mammals

Thirteen species were chosen for this study based on phylogenetic representativeness, the quality of their genome assemblies (Table S9), and sample availability. The procedures were conducted following the approval of the Animal Experiments Ethics Committee at the Institute of Zoology, Chinese Academy of Sciences. Genomic DNA was extracted from liver and skeletal muscle samples and pooled with lambda phage DNA, fragmented to 200–300 bp, bisulfite converted using an EZ DNA Methylation Gold Kit (Zymo) and PCR amplified. WGBS libraries were sequenced on a HiSeq sequencer (Illumina) to produce 2 × 150 bp paired-end reads. Reads were first trimmed with Trimmomatic v0.39[78] and then mapped to the corresponding genomes using Bismark v0.23[79] with the default parameters and cleaned to remove duplicates. CpG methylation calls were extracted using in-house scripts. Public WGBS sequencing data were analyzed using the same pipeline. The mapping efficiency was ~70% for all samples, and the average sequencing depth was 13.2x (ranging from 10.5x to 15.6x) after deduplication, which is sufficient for methylation analysis according to Ziller et al.,[80] and the mean CpG methylation level was 69.0%.

Multi-species whole-genome alignment and identification of CNEs and orthologous genes

The human genome was used as the reference and aligned with other genomes using LASTZ v1.02[81] and the whole-genome alignment of 13 species was generated using Multiz-TBA v11.2.[82] A total of 714,784 CNEs longer than 50 bp were identified using the PHAST package.[83] The homology between the coding genes of 13 species was determined using OrthoFinder v2.5.4[84] with sequence similarity searches performed via DIAMOND.[85] We identified 7,365 single-copy orthologous genes shared by all 13 species. Given the relatively high evolutionary distance between marsupials and eutherians, we included another 2,241 single-copy orthologous genes from all eutherians and one marsupial species as well as 613 single-copy orthologous genes from all eutherians. Therefore, we conducted comparative analyses with a total of 10,219 orthologous genes shared by at least 11 eutherian species.

Hierarchical clustering analysis based on CpG and syntenic region methylation profiles

The epigenetic distances between samples were calculated using the "dist" function with the Euclidean distance measure in R, and hierarchical clustering analysis was performed with the "hclust" function with Ward's minimum variance method.[86] The methylation levels of shared CpG sites and syntenic regions longer than 100 bp with at least three CpGs were used to calculate sample distances. Since the methylation levels of replicates were highly correlated, we merged the data of each species to represent their methylation profiles in the following analysis.

Gene evolution analyses

We used MACSE v2.04[87] to align the coding sequences of orthologous genes and used gBlocks v0.91[88] to obtain conserved blocks. The dS of orthologous gene pairs was calculated using KaKs_Calculator v2.0[89] with the parameter "-m YN," and the dN/dS ratios of each gene in all species were estimated using PAML under the "free ratio" model.[90] Methylation divergence was calculated as (M1−M2)/(M1+M2), where M1 and M2 are the methylation levels of two gene promoters. Only promoters with at least three CpG sites were used to calculate methylation levels.

Consistent and divergent DNA methylation of promoters and noncoding sequences

Promoter methylation levels were determined using 2,000 bp upstream to 50 bp downstream of TSS of genes. Based on the bimodal distribution of promoter methylation, we defined consistently hypermethylated promoters as those showing methylation levels ≥50% in all species with an average methylation level above 70% and consistently hypomethylated promoters as those showing methylation levels <50% in all species with an average methylation level below 30%. Lineage-specific altered methylation indicates that the species or taxon exhibits unique hypermethylation (≥50%) or hypomethylation (<50%), where the methylation difference between this species/taxon and the average methylation level of other species is greater than 30%. Functional enrichment analyses of genes with consistent and specific promoter methylation were conducted using Metascape.[91] To identify differentially methylated noncoding regions, we scanned the multi-species syntenic regions using a sliding window approach with a window length of 200 bp and a step size of 50 bp in a manner independent of genomic annotations. Blocks with lengths less than 200 bp were also retained. Then, we excluded regions that overlapped with promoters and coding sequences and calculated the methylation levels of these noncoding regions containing at least three CpG sites. Species-specific differential methylation was defined using the same standard as described above, and differentially methylated regions located within 200 bp were merged. Functional enrichment analyses of these differentially methylated noncoding regions were conducted using GREAT v4.0.4 with the "basal plus extension" model and default parameters.[92]

Identification and functional analyses of AMRs

We identified the AMRs in each species with AMRFinder from MethPipe v4.1.1[93] with the parameters "-g 1000 -b." The minimum coverage per CpG to test windows was set to half of the genome depth after deduplication to adjust for differences in sequencing depths. These AMRs in each species were used to check if they overlapped with the syntenic regions of known human and mouse ICRs to determine the evolutionary origins of the known ICRs. The syntenic regions of known ICRs were first identified based on whole-genome alignment and double checked against gene orthology and gene structure.

To obtain more accurate results for potential imprints, we further filtered AMRs with the following criteria: (1) the region was covered by at least eight all-methylated reads (>80% of CpG sites were methylated per read) or all-unmethylated reads (>80% of CpG sites were unmethylated per read); (2) the ratio of all-methylated to all-unmethylated reads covered in the region is between 0.5 and 2; (3) the average methylation level of the region is between 30% and 70%; (4) AMRs containing SNPs at CpG sites, identified with BS-SNPer v1.1,[94] were excluded to mitigate the influence of individual genetic polymorphisms on DNA methylation; and (5) AMRs located at sex chromosomes were excluded. Finally, AMRs located within 1,000 bp of each other were merged in each sample, and only the overlapping AMRs among the replicates from the same species (including AMRs located within 1,000 bp) were retained to represent highly credible AMRs in each species. The bottlenose dolphin was excluded from these analyses due to the lack of replicates. Altogether, this filtering process resulted in an average of roughly 1,300 highly credible AMRs in each species, with an average length of 2,000 bp and an average methylation level of 54%, in accordance with the known ICRs. Species-specific AMRs were defined as the highly credible AMRs in the target species shared by at least eight eutherians, where the syntenic regions in other species did not cover any allelic methylated fraction of a window size of 10 CpG sites.

GO enrichment analyses of AMRs were conducted with Metascape,[91] and the motif enrichment analysis was performed using HOMER with the default parameters.[95]

DATA AND CODE AVAILABILITY

The WGBS data generated by this research are deposited in NCBI (PRJNA896705) and other data needed to evaluate the conclusions are present in this paper and/or the supplementary materials. Original scripts used in the analysis are available upon request.

REFERENCES

1. Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. ***9***, 465–476.
2. Zemach, A., McDaniel, I.E., Silva, P., et al. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science ***328***, 916–919.
3. Greenberg, M.V.C., and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. Nat. Rev. Mol. Cell Biol. ***20***, 590–607.
4. Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. ***25***, 1010–1022.
5. Yin, Y., Morgunova, E., Jolma, A., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science ***356***, eaaj2239.
6. Hon, G.C., Rajagopal, N., Shen, Y., et al. (2013). Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nat. Genet. ***45***, 1198–1206.
7. Du, W., Shi, G., Shan, C.-M., et al. (2022). Mechanisms of chromatin-based epigenetic inheritance. Sci. China Life Sci. ***65***, 2162–2190.
8. Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat. Rev. Genet. ***13***, 153–162.
9. Guo, H., Zhu, P., Yan, L., et al. (2014). The DNA methylation landscape of human early embryos. Nature ***511***, 606–610.

10. Pai, A.A., Bell, J.T., Marioni, J.C., et al. (2011). A genome-wide study of DNA methylation patterns and gene expression levels in multiple human and chimpanzee tissues. PLoS Genet. ***7***, e1001316.
11. Zeng, J., Konopka, G., Hunt, B.G., 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.
12. Hernando-Herraez, I., Prado-Martinez, J., Garg, P., et al. (2013). Dynamics of DNA methylation in recent human and great ape evolution. PLoS Genet. ***9***, e1003763.
13. Hernando-Herraez, I., Heyn, H., Fernandez-Callejo, M., et al. (2015). The interplay between DNA methylation and sequence divergence in recent human evolution. Nucleic Acids Res. ***43***, 8204–8214.
14. Mendizabal, I., Shi, L., Keller, T.E., et al. (2016). Comparative methylome analyses identify epigenetic regulatory loci of human brain evolution. Mol. Biol. Evol. ***33***, 2947–2959.
15. Blake, L.E., Roux, J., Hernando-Herraez, I., et al. (2020). A comparison of gene expression and DNA methylation patterns across tissues and species. Genome Res. ***30***, 250–262.
16. Sahm, A., Koch, P., Horvath, S., et al. (2021). An analysis of methylome evolution in primates. Mol. Biol. Evol. ***38***, 4700–4714.
17. Barlow, D.P., Stöger, R., Herrmann, B.G., et al. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the *Tme* locus. Nature ***349***, 84–87.
18. Barlow, D.P., and Bartolomei, M.S. (2014). Genomic imprinting in mammals. Cold Spring Harbor Perspect. Biol. ***6***, a018382.
19. Lowdon, R.F., Jang, H.S., and Wang, T. (2016). Evolution of epigenetic regulation in vertebrate genomes. Trends Genet. ***32***, 269–283.
20. Lokk, K., Modhukur, V., Rajashekar, B., et al. (2014). DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. Genome Biol. ***15***, r54.
21. Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature ***518***, 317–330.
22. Keller, T.E., and Yi, S.V. (2014). DNA methylation and evolution of duplicate genes. Proc. Natl. Acad. Sci. USA ***111***, 5932–5937.
23. Gunasekara, C.J., Scott, C.A., Laritsky, E., et al. (2019). A genomic atlas of systemic interindividual epigenetic variation in humans. Genome Biol. ***20***, 105.
24. Zhang, H., Li, D., Wang, C., et al. (2009). Delayed implantation in giant pandas: the first comprehensive empirical evidence. Reproduction ***138***, 979–986.
25. Litingtung, Y., Dahn, R.D., Li, Y., et al. (2002). *Shh* and *Gli3* are dispensable for limb skeleton formation but regulate digit number and identity. Nature ***418***, 979–983.
26. Shapiro, M.D., Hanken, J., and Rosenthal, N. (2003). Developmental basis of evolutionary digit loss in the Australian lizard *Hemiergis*. J. Exp. Zool. B Mol. Dev. Evol. ***297***, 48–56.
27. Stopper, G.F., and Wagner, G.P. (2007). Inhibition of Sonic hedgehog signaling leads to posterior digit loss in *Ambystoma mexicanum*: parallels to natural digit reduction in urodeles. Dev. Dynam. ***236***, 321–331.
28. Vargas, A.O., and Wagner, G.P. (2009). Frame-shifts of digit identity in bird evolution and Cyclopamine-treated wings. Evol. Dev. ***11***, 163–169.
29. Hockman, D., Cretekos, C.J., Mason, M.K., et al. (2008). A second wave of *Sonic hedgehog* expression during the development of the bat limb. Proc. Natl. Acad. Sci. USA ***105***, 16982–16987.
30. Maier, J.A., Rivas-Astroza, M., Deng, J., et al. (2017). Transcriptomic insights into the genetic basis of mammalian limb diversity. BMC Evol. Biol. ***17***, 86.
31. Cozzi, B., Huggenberger, S., and Oelschläger, H.A. (2016). Anatomy of Dolphins: Insights into Body Structure and Function (Academic Press).
32. Fuchs, H., Sabrautzki, S., Przemeck, G.K.H., et al. (2016). The first *Scube3* mutant mouse line with pleiotropic phenotypic alterations. G3 (Bethesda) ***6***, 4035–4046.
33. Renfree, M.B. (1983). Marsupial reproduction: the choice between placentation and lactation. Oxford reviews of reproductive biology, ***vol 5*** (Clarendon Press), pp. 1–29.
34. Guernsey, M.W., Chuong, E.B., Cornelis, G., et al. (2017). Molecular conservation of marsupial and eutherian placentation and lactation. Elife ***6***, e27450.
35. Wang, L., Zhang, J., Duan, J., et al. (2014). Programming and inheritance of parental DNA methylomes in mammals. Cell ***157***, 979–991.
36. Court, F., Tayama, C., Romanelli, V., et al. (2014). Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. Genome Res. ***24***, 554–569.
37. Joshi, R.S., Garg, P., Zaitlen, N., et al. (2016). DNA methylation profiling of uniparental disomy subjects provides a map of parental epigenetic bias in the human genome. Am. J. Hum. Genet. ***99***, 555–566.
38. Zink, F., Magnusdottir, D.N., Magnusson, O.T., et al. (2018). Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. Nat. Genet. ***50***, 1542–1552.
39. Ono, R., Nakamura, K., Inoue, K., et al. (2006). Deletion of *Peg10*, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. Nat. Genet. ***38***, 101–106.
40. Constância, M., Hemberger, M., Hughes, J., et al. (2002). Placental-specific IGF-II is a major modulator of placental and fetal growth. Nature ***417***, 945–948.
41. Lau, M.M., Stewart, C.E., Liu, Z., et al. (1994). Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. Genes Dev. ***8***, 2953–2963.
42. Kuratomi, G., Iwamoto, K., Bundo, M., et al. (2008). Aberrant DNA methylation associated with bipolar disorder identified from discordant monozygotic twins. Mol. Psychiatry ***13***, 429–441.
43. Conti, V., Carabalona, A., Pallesi-Pocachard, E., et al. (2013). Periventricular heterotopia in 6q terminal deletion syndrome: role of the *C6orf70* gene. Brain ***136***, 3378–3394.
44. Li, C., Fan, Y., Li, G., et al. (2018). DNA methylation reprogramming of functional elements during mammalian embryonic development. Cell Discov. ***4***, 41.
45. Sakatani, T., Wei, M., Katoh, M., et al. (2001). Epigenetic heterogeneity at imprinted loci in normal populations. Biochem. Biophys. Res. Commun. ***283***, 1124–1130.
46. Li, C., Bin, Y., Curchoe, C., et al. (2008). Genetic imprinting of *H19* and *IGF2* in domestic pigs (*Sus scrofa*). Anim. Biotechnol. ***19***, 22–27.
47. Ferrón, S.R., Radford, E.J., Domingo-Muelas, A., et al. (2015). Differential genomic imprinting regulates paracrine and autocrine roles of IGF2 in mouse adult neurogenesis. Nat. Commun. ***6***, 8265.
48. Mendiola, A.J.P., and LaSalle, J.M. (2021). Epigenetics in Prader-Willi Syndrome. Front. Genet. ***12***, 624581.
49. Zhou, Q., Guan, P., Zhu, Z., et al. (2022). ASMdb: a comprehensive database for allele-specific DNA methylation in diverse organisms. Nucleic Acids Res. ***50***, D60–D71.
50. Li, X., Ito, M., Zhou, F., et al. (2008). A maternal-zygotic effect gene, *Zfp57*, maintains both maternal and paternal imprints. Dev. Cell ***15***, 547–557.
51. Quenneville, S., Verde, G., Corsinotti, A., et al. (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol. Cell ***44***, 361–372.
52. Kurukuti, S., Tiwari, V.K., Tavoosidana, G., et al. (2006). CTCF binding at the *H19* imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*. Proc. Natl. Acad. Sci. USA ***103***, 10684–10689.
53. Llères, D., Moindrot, B., Pathak, R., et al. (2019). CTCF modulates allele-specific sub-TAD organization and imprinted gene activity at the mouse *Dlk1-Dio3* and *Igf2-H19* domains. Genome Biol. ***20***, 272.
54. Davies, W., Isles, A.R., and Wilkinson, L.S. (2005). Imprinted gene expression in the brain. Neurosci. Biobehav. Rev. ***29***, 421–430.
55. Wilkinson, L.S., Davies, W., and Isles, A.R. (2007). Genomic imprinting effects on brain development and function. Nat. Rev. Neurosci. ***8***, 832–843.
56. McKenna, K.Z., Wagner, G.P., and Cooper, K.L. (2021). A developmental perspective of homology and evolutionary novelty. In Current Topics in Developmental Biology, S.F. Gilbert, ed. (Academic Press), pp. 1–38.
57. Griffith, O.W., Chavan, A.R., Protopapas, S., et al. (2017). Embryo implantation evolved from an ancestral inflammatory attachment reaction. Proc. Natl. Acad. Sci. USA ***114***, E6566–E6575.
58. Kin, K., Maziarz, J., Chavan, A.R., et al. (2016). The transcriptomic evolution of mammalian pregnancy: gene expression innovations in endometrial stromal fibroblasts. Genome Biol. Evol. ***8***, 2459–2473.
59. Xiao, S., Xie, D., Cao, X., et al. (2012). Comparative epigenomic annotation of regulatory DNA. Cell ***149***, 1381–1392.
60. Villar, D., Berthelot, C., Aldridge, S., et al. (2015). Enhancer evolution across 20 mammalian species. Cell ***160***, 554–566.
61. Vermunt, M.W., Tan, S.C., Castelijns, B., et al. (2016). Epigenomic annotation of gene regulatory alterations during evolution of the primate brain. Nat. Neurosci. ***19***, 494–503.
62. Liu, S., Yu, Y., Zhang, S., et al. (2020). Epigenomics and genotype-phenotype association analyses reveal conserved genetic architecture of complex traits in cattle and human. BMC Biol. ***18***, 80.
63. Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet. ***7***, 45–49.
64. Curley, J.P., Barton, S., Surani, A., et al. (2004). Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. Proc. Biol. Sci. ***271***, 1303–1309.
65. Barlow, D.P. (1993). Methylation and imprinting: from host defense to gene regulation? Science ***260***, 309–310.
66. Renfree, M.B., Suzuki, S., and Kaneko-Ishino, T. (2013). The origin and evolution of genomic imprinting and viviparity in mammals. Proc. Biol. Sci. ***368***, 20120151.
67. Wilkins, J.F., and Bhattacharya, T. (2019). Intragenomic conflict over bet-hedging. Proc. Biol. Sci. ***374***, 20180142.
68. Coulondre, C., Miller, J.H., Farabaugh, P.J., et al. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. Nature ***274***, 775–780.
69. Jiang, C., and Zhao, Z. (2006). Directionality of point mutation and 5-methylcytosine deamination rates in the chimpanzee genome. BMC Genom. ***7***, 316.
70. Chuang, T.-J., and Chen, F.-C. (2014). DNA methylation is associated with an increased level of conservation at nondegenerate nucleotides in mammals. Mol. Biol. Evol. ***31***, 387–396.
71. Hu, J., Askary, A.M., Thurman, T.J., et al. (2019). The epigenetic signature of colonizing new environments in *Anolis* lizards. Mol. Biol. Evol. ***36***, 2165–2170.
72. Wilkinson, G.S., Adams, D.M., Haghani, A., et al. (2021). DNA methylation predicts age and provides insight into exceptional longevity of bats. Nat. Commun. ***12***, 1615.
73. Pinho, G.M., Martin, J.G.A., Farrell, C., et al. (2022). Hibernation slows epigenetic ageing in yellow-bellied marmots. Nat. Ecol. Evol. ***6***, 418–426.
74. Hu, Y., Wang, X., Xu, Y., et al. (2023). Molecular mechanisms of adaptive evolution in wild animals and plants. Sci. China Life Sci. ***66***, 453–495.
75. Skinner, M.K., Gurerrero-Bosagna, C., Haque, M.M., et al. (2014). Epigenetics and the evolution of Darwin's finches. Genome Biol. Evol. ***6***, 1972–1989.
76. Smith, T.A., Martin, M.D., Nguyen, M., et al. (2016). Epigenetic divergence as a potential first step in darter speciation. Mol. Ecol. ***25***, 1883–1894.
77. Li, Y., Liew, Y.J., Cui, G., et al. (2018). DNA methylation regulates transcriptional homeostasis of algal endosymbiosis in the coral model Aiptasia. Sci. Adv. ***4***, eaat2142.
78. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics ***30***, 2114–2120.
79. Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics ***27***, 1571–1572.

80. Ziller, M.J., Hansen, K.D., Meissner, A., et al. (2015). Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. Nat. Methods ***12***, 230–232.
81. Harris, R.S. (2007). Improved Pairwise Alignmnet of Genomic DNA. PhD Thesis (The Pennsylvania State University).
82. Blanchette, M., Kent, W.J., Riemer, C., et al. (2004). Aligning multiple genomic sequences with the threaded blockset aligner. Genome Res. ***14***, 708–715.
83. Siepel, A., Bejerano, G., Pedersen, J.S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. ***15***, 1034–1050.
84. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. ***20***, 238.
85. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods ***12***, 59–60.
86. R Core Team (2021). R: A Language and Environment for Statistical Computing.
87. Ranwez, V., Douzery, E.J.P., Cambon, C., et al. (2018). MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Mol. Biol. Evol. ***35***, 2582–2584.
88. Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. ***56***, 564–577.
89. Wang, D., Zhang, Y., Zhang, Z., et al. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics Proteomics Bioinformatics ***8***, 77–80.
90. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. ***24***, 1586–1591.
91. Zhou, Y., Zhou, B., Pache, L., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. ***10***, 1523.
92. McLean, C.Y., Bristor, D., Hiller, M., et al. (2010). GREAT improves functional interpretation of *cis*-regulatory regions. Nat. Biotechnol. ***28***, 495–501.
93. Fang, F., Hodges, E., Molaro, A., et al. (2012). Genomic landscape of human allele-specific DNA methylation. Proc. Natl. Acad. Sci. USA ***109***, 7332–7337.
94. Gao, S., Zou, D., Mao, L., et al. (2015). BS-SNPer: SNP calling in bisulfite-seq data. Bioinformatics ***31***, 4006–4008.
95. Heinz, S., Benner, C., Spann, N., 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.
96. Kumar, S., Suleski, M., Craig, J.M., et al. (2022). TimeTree 5: an expanded resource for species divergence times. Mol. Biol. Evol. ***39***, msac174.

ACKNOWLEDGMENTS

This study is funded by National Natural Science Foundation of China (31821001), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000), and the PI Project of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2020GD0804).

AUTHOR CONTRIBUTIONS

F.W. conceptualized the project. Y.H., S.Y., X.D. designed the methodology. Y.H. performed data analysis and visualization. Y.H. wrote the manuscript with contributions from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

It can be found online at https://doi.org/10.1016/j.xinn.2023.100434.

LEAD CONTACT WEBSITE

Fuwen Wei: http://sourcedb.ioz.cas.cn/yw/scs/casm/200907/t20090716_2088462.html.