

REVIEW

# DNA Methylation: Insights into Human Evolution

#### Irene Hernando-Herraez<sup>1</sup><sup>©</sup>, Raquel Garcia-Perez<sup>1</sup><sup>©</sup>, Andrew J. Sharp<sup>2</sup>\*, Tomas Marques-Bonet<sup>1,3,4</sup>\*

1 Institute of Evolutionary Biology (UPF-CSIC), PRBB, Barcelona, Spain, 2 Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai School, New York, New York, United States of America, 3 Catalan Institution of Research and Advanced Studies (ICREA), Barcelona, Spain, 4 Centro Nacional de Analisis Genomico (CNAG), PCB, Barcelona, Spain

• These authors contributed equally to this work.

\* andrew.sharp@mssm.edu (AJS); tomas.marques@upf.edu (TMB)

## Abstract

A fundamental initiative for evolutionary biologists is to understand the molecular basis underlying phenotypic diversity. A long-standing hypothesis states that species-specific traits may be explained by differences in gene regulation rather than differences at the protein level. Over the past few years, evolutionary studies have shifted from mere sequence comparisons to integrative analyses in which gene regulation is key to understanding species evolution. DNA methylation is an important epigenetic modification involved in the regulation of numerous biological processes. Nevertheless, the evolution of the human methylome and the processes driving such changes are poorly understood. Here, we review the close interplay between Cytosine-phosphate-Guanine (CpG) methylation and the underlying genome sequence, as well as its evolutionary impact. We also summarize the latest advances in the field, revisiting the main literature on human and nonhuman primates. We hope to encourage the scientific community to address the many challenges posed by the field of comparative epigenomics.

## Introduction

Methylation of nucleotide bases is the only covalent modification of DNA commonly found across many different taxa. So far, three types have been described: N6-methyladenine (6mA), N4-methylcytosine (4mC), and 5-methylcytosine (5mC). While 6mA and 4mC are restricted to prokaryotes and certain eukaryotes [1-3], 5mC is the predominant epigenetic modification in eukaryotic DNA [4]. In mammals, 5mC mainly occurs in the context of Cytosine-phosphate-Guanine (CpG) dinucleotides [5].

The distribution of CpG methylation is uneven across mammalian genomes. While the majority of CpGs (~60%–80%) are methylated [5], regions of densely clustered CpGs, known as CpG islands (CGIs), are often devoid of methylation [6]. Many CGIs are found in the vicinity of gene promoters, with approximately two-thirds of genes having a CGI at their promoter. Methylation of promoter CGIs provokes long-term transcriptional repression of the associated genes [5,7]. Classical examples include X chromosome inactivation [8] and genomic imprinting [9]. In contrast, the functional impact of DNA methylation outside gene promoters is not



# 

**Citation:** Hernando-Herraez I, Garcia-Perez R, Sharp AJ, Marques-Bonet T (2015) DNA Methylation: Insights into Human Evolution. PLoS Genet 11(12): e1005661. doi:10.1371/journal.pgen.1005661

Editor: Rebecca J Oakey, King's College London, UNITED KINGDOM

Published: December 10, 2015

**Copyright:** © 2015 Hernando-Herraez et al. This is an open access article distributed under the terms of the <u>Creative Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: TMB is supported by ICREA (www.icrea. cat), EMBO YIP (www.embo.org) 2013, MICINN BFU2014-55090-P (www.mecd.gob.es), BFU2015-7116-ERC and BFU2015-6215-ERC. AJS is supported by NIH (www.nih.gov) grants DA033660, HG006696, HD073731, and MH097018, and research grant 6-FY13-92 from the March of Dimes Foundation. IHH is supported by AGAUR (Generalitat de Catalunya, Spain; www.gencat.cat/agaur/) and RGP by a fellowship from MICINN (www.mecd.gob. es). We also acknowledge the Barcelona Zoo (Ajuntament de Barcelona) for an award to IHH. The funders had no role in the preparation of the article.

**Competing Interests:** The authors have declared that no competing interests exist.

so well understood. Gene body methylation has been reported to be involved in alternative splicing [10], and methylation of transposable elements is involved in the suppression of retrotransposition [11]. Finally, certain unmethylated domains in CpG-poor regions have been shown to coincide with distal regulatory elements, such as active enhancers, and usually colocate with other epigenetic marks, such as histone modifications [12–15]. Together, these findings suggest that DNA methylation patterns are complex and highly dependent on the genomic context [16].

Non-CpG methylation, the methylation of cytosines followed by a base other than guanine, has also been reported. Indeed, non-CpG methylation is abundant in plants [17,18] and recently has also been identified in some mammalian cell types [12,18–20]. Furthermore, additional chemical modifications arisen from the oxidative conversion of 5mC have also been described [21–24]. The significance of these epigenetic marks in the mammalian genome is currently poorly understood, so our focus is on methylation that occurs in CpG dinucleotides

Despite being a key regulator of genomic function, the role of DNA methylation in species evolution is only beginning to be explored. More than four decades ago, it was first proposed that regulatory changes could lead to species-specific adaptations as well as phenotypic variability [25]; however, technical limitations at that time prevented this hypothesis from being tested. Within the past decade, the development of high-throughput genomic technologies has allowed the study of species evolution from a molecular perspective. Taking advantage of these technologies, recent studies have provided the first insights into the evolution of the epigenome [26]. Some of the key questions in the field are starting to be addressed: Is regulatory evolution coupled with sequence evolution? To what extent do regulatory changes affect transcription? What are the evolutionary landmarks of DNA methylation? In this review, we survey the incipient field of evolutionary epigenomics, focusing on CpG methylation. We discuss the current state of the field and conclude by suggesting future research avenues.

#### The Interplay between the Genome and the Methylome

There is a close interplay between DNA methylation and the underlying nucleotide sequence. The mutagenic nature of 5mC is one of the primary players in this crosstalk, as methylcytosines are prone to spontaneous deamination to thymine, resulting in a CpG-to-Thymine-phosphate-Guanine (TpG) mutation rate that is ~10-fold higher than for other dinucleotides [27]. This is reflected in the nucleotide divergence between primates: While the average human-chimpan-zee divergence is ~1% across the genome, at CpG sites it increases to ~15% [28]. Furthermore, heavily methylated, CpG-rich, subtelomeric regions exhibit a high rate of deamination, which is balanced by the rapid gain of guanines and cytosines, commonly attributed to biased gene conversion [29]. Nonetheless, in a proportion (~15%) of these genomic regions, the loss of CpG dinucleotides is not compensated [29]. Indeed, high deamination rates over evolutionary time are thought to be the cause of the overall depletion of CpGs that is characteristic of mammalian genomes.

In contrast to the mutagenic effect of 5mC, many genomic regions with high densities of CpGs are constitutively unmethylated. Recent studies suggest that this CpG richness can be solely accounted for by low CpG deamination rates, rather than being a byproduct of purifying selection [29]. About 80% of these regions are located close to (<10 Kb) annotated transcription start sites (TSSs), and meet the classic definition of CpG islands [30]. Therefore, the methylation state plays an important role in the dynamics of CpG dinucleotides.

On the other hand, several lines of evidence indicate that the genome harbors the information necessary to establish local DNA methylation patterns. How this genomic information is encoded and read is still not fully understood, but recent work is beginning to shed light on the



**Fig 1. The interplay between the genome and the methylome.** A) Methylated cytosines tend to deaminate over evolutionary time and, thus, the methylation state of cytosines in different species influences the evolution of the underlying genome sequence. B) Species-specific nucleotide changes that disrupt transcription factor (TF) binding sites can alter the methylation state of nearby CpG dinucleotides and, as a consequence, establish species-specific differentially methylated regions (DMRs). C) The insertion of transposable elements in a particular lineage, along with the accumulation of nucleotide changes, can lead to the emergence of novel CpG dinucleotides, creating species-specific regulatory regions.

doi:10.1371/journal.pgen.1005661.g001

regulatory mechanisms. Fragment-insertion experiments have demonstrated that both a high Guanine-phosphate-Cytosine (GC) content and a high density of CpG dinucleotides are necessary and sufficient requirements to induce an unmethylated state [31,32]. From a mechanistic standpoint, it is unclear how CpG and GC content richness drive the unmethylated state. Proteins that specifically recognize unmethylated CpGs via CXXC domains could prevent DNA methylation [33–35]. Other hypotheses posit the existence of proteins that recruit DNA methyltransferases through the recognition of Adenosine-phosphate-Thymine (AT)-rich regions [31].

Therefore, evolutionary changes in CpG dinucleotides and GC content could contribute to alter the methylation landscape of a species (Fig 1A). With this idea in mind, regions with high CpG density that are present exclusively in the human lineage have been identified and termed "CpG beacons" [36]. Although their methylation state has not been studied, these regions are good candidates to show differential methylation between species. Interestingly, human-specific CpG beacons were found to be enriched for genes related to cognition and behavior, including the well-characterized *HAR1A* gene, which plays a crucial role in cortical development [37].

Different mechanisms govern methylation patterning at CpG-poor regions. Some compelling studies have revealed that methylation patterns at such regions are likely driven by the binding of transcription factors (TFs), as the presence of specific motifs suffices to create unmethylated domains [13,32,38,39]. Importantly, this mechanism likely plays a crucial role during cell differentiation [40]. Further evidence results from the analysis of methylation patterns after recent duplication events in the human genome. Though most duplicated segments share conserved methylation states, duplication pairs with discordant methylation are associated with nucleotide changes at particular binding motifs [41].

Evolutionary studies in human populations have identified genetic variants that associate with methylation levels at nearby CpGs, termed methylation quantitative trait loci (mQTL) [42–45]. Consistent with the role of TFs in determining methylation patterns, genetic variants that disrupt TF binding sites are more frequently associated with changes in methylation (Fig 1B). Furthermore, TF binding sites within regions that show differential methylation in the human lineage compared to great apes show an increase of human-specific mutations [46]. Altogether, these findings suggest that evolutionary changes in TF binding motifs contribute to shaping the methylome between species.

Finally, it has also been suggested that retrotransposition events coupled with gradual nucleotide changes could lead to the accumulation of novel CpG sites [47]. These regions would provide the necessary grounds for the emergence of new regulatory regions (Fig 1C). Indeed, several reports suggest that imprinting at certain genes could have arisen as a byproduct of DNA methylation silencing of retrotransposons. One such case is the *RB1* imprinted gene, which results from the differential methylation of a processed pseudogene [48].

The emerging picture is that the joint actions of all the above-discussed mechanisms throughout the evolutionary scale have modeled the architecture of the methylome ( $\underline{Fig 1}$ ).

### **Comparative Epigenomics**

Comparative epigenomics, the interspecies comparison of DNA and histone modifications, is a promising research field that can enlighten our understanding of mammalian epigenome evolution [49]. Furthermore, together with TF footprints and chromatin accessibility maps, it can be used as a tool to map regulatory elements [50,51]. Evolutionary studies have shown that 90% of unmethylated regions that are associated with gene promoters are shared across distant vertebrates, from zebrafish to humans, indicating that the unmethylated state at gene promoters is a deeply conserved epigenetic feature [52]. Histone modifications are also conserved at promoters, in sharp contrast to enhancers, whose evolution is much more rapid and less constrained [50].

Importantly, co-localization of different epigenetic marks, including DNA methylation, several histone modifications, and H2A.Z, has been used to study the evolution of the epigenome, the transcriptome, and the genome in humans, mice, and pigs [49]. This study revealed several important points: (i) the co-occurrence of different epigenetic marks is conserved across species, (ii) histone modifications are predictive of gene expression levels in all three species and can explain more than 50% of the variance in gene expression [49,53,54], and (iii) interspecies epigenetic variation often occurs in neutrally evolving regions, leading to the suggestion that the epigenome might act as a buffer by masking genetic changes from immediate phenotypic changes [49]. Such deep comparative epigenetic analyses are extremely valuable and provide the groundwork to characterize general features of the species epigenomes [49,50,55]. Yet, only when we identify the differences that separate us from our closest nonhuman relatives do we get closer to understanding the uniqueness of our species.

#### **Human and Nonhuman Primates**

Though methylomes are overall quite well-conserved between primate species [56,57], several studies have now identified hundreds of species-specific differentially methylated regions (sDMRs) (Table 1). Early studies focused exclusively on methylation levels at gene promoters and CGIs [58,59], providing some interesting examples, including sDMRs in the brain and associated with known disease genes [60-62]. On the premise that changes in gene expression

Reference	Species	Methodology	Tissue	Highlights
Wang, J. (2012)	Human, Macaque	MeDIP-chip and SEQUENOM MassARRAY	Prefrontal cortex	>100 differentially methylated regions; Validated DMRs associated with genes with neural functions and with schizophrenia and Alzheimer's disease
Pai, A. (2011)	Human, Chimpanzee	Illumina 27K array	Liver, heart, and kidney	14.5% of promoter CpG sites are differentially methylated between tissues; 8.6% of promoter CpG sites are differentially methylated between species; Interspecies differences in promoter methylation underlie 12%–18% of gene expression differences
Molaro, A. (2011)	Human, Chimpanzee	Whole-genome bisulfite sequence	Sperm	70% of genes are hypomethylated in both chimpanzee and human sperm; 6% and 35% of orthologous SVAs had a methylation level below 50% in chimpanzee and human sperm, respectively
Martin, D.I.K. (2011)	Human, Chimpanzee, Orangutan	MethylSeq	Neutrophils	10% of CpG islands-like regions present different methylation states between chimpanzees and humans; Regions with differential methylation might have diverged in gene regulatory function
Fukuda, K. (2013)	Human, Chimpanzee	MeDIP-chip (chromosomes 21 and 22)	Peripheral blood leukocytes	16 sDMRs between chimpanzees and humans in chromosomes 21 and 22; Genetic changes underlying these differences in methylation include gain/loss of CTCF-binding sites and changes in CpG density
Hernando- Herraez, I. (2013)	Human, Chimpanzee, Bonobo, Gorilla, Orangutan	Illumina 450K array	Peripheral blood	~9% of the assayed CpG sites showed significant methylation differences between chimpanzees and humans; 184 genes perfectly conserved at protein level show significant epigenetic differences between chimpanzees and humans
Hernando- Herraez, I. (2015)	Human, Chimpanzee, Gorilla, Orangutan	Whole-genome bisulfite sequence	Peripheral blood	72% of the hypomethylated regions (HMRs) were shared among all four species; 42.6% of HMRs were on human CpG islands; 52.6% of HMRs were on human CpG shores
Gokhman, D. (2014)	Neandertal, Denisovan	Deamination rate as a proxy for DNA methylation	Femur, costae, and tibia bones	>2,000 DMRs between archaic and present-day humans; Substantial changes in methylation in the <i>HOXD</i> cluster
Fraser, H. B (2012)	Human	Illumina 27K array	Lymphoblastoid cell lines	21.4% of CpG sites differed in methylation between populations; 5.4% of these CpG sites were strongly associated with local SNPs
Heyn, H. 2013	Human	Illumina 450K array	Lymphoblastoid cell lines	439 population-specific differentially methylated CpG sites (pop-CpG); Significantly decreased gene expression associated to promoter hypermethylation in 12.9% (13 out of 101) of pop-CpG; Significantly increased gene expression associated to gene body methylation in 23.9% (27 out of 113) of pop-CpG

#### Table 1. Comparative studies of DNA methylation patterns in primates.

doi:10.1371/journal.pgen.1005661.t001

patterns are major contributors to evolutionary innovation [63–66], several studies have also attempted to link expression differences with underlying epigenetic changes at regulatory regions. Early studies using methylation arrays with limited coverage suggested that differential promoter methylation underlies only ~12%–18% of observed gene expression differences between human and chimpanzee [67]. However, other studies focusing on histone modifications indicate that a larger proportion (~40%) of differentially expressed genes can be accounted for by interspecies epigenetic differences [49,53,54]. More recent genome-wide DNA methylation studies have shown that most sDMRs actually locate distal TSS, and that these sites are also enriched in particular histone modifications [46]. This finding highlights the importance of epigenetic modifications at distal regulatory elements, as well as the close interplay between different epigenetic marks. Of note, it has been observed that regions of methylation divergence often coincide not only with regulatory elements of genes functionally related to the tissue being studied but also with genes with functions specific to other tissues and developmental time points [56,62,67]. Though it is unclear how to explain the latter observation, it could be due to vestigial regulatory elements that were functional during development, and their methylation state has remained unaltered in adult tissues [68]. This is an engaging interpretation, since many human-specific morphological features emerge during early development. Indeed, developmental stages are crucial for the acquisition of human-specific traits, and species-specific differences in histone modifications have also been reported in genes involved in developmental processes. Recent studies have shown that many genes that have gained H3K27ac marks since the human-macaque split have important roles in limb development or the acquisition of human-specific traits [69]. Similarly, thousands of promoters and enhancers associated with genes crucial in cortical development were reported to display human-specific gains both in H3K27ac and H3K4me2 [70].

In sum, from primate epigenetic comparative studies, we have started to gain insight into the molecular mechanisms through which evolutionary novelties may have arisen in the human lineage. These studies suggest that epigenetic modifications in regulatory elements could have altered expression patterns of key genes, leading to evolutionary adaptations.

#### **Humans and Extinct Hominids**

Genome sequencing of extinct hominids has provided a better understanding of the population history and genome evolution of our species [71,72]. However, additional studies have been limited due to the tiny amounts and degraded nature of DNA that can be extracted from ancient bones. Recently, a novel method was developed to infer DNA methylation patterns from the genome sequence of Neandertals and Denisovans, opening the possibility of studying for the very first time the methylome of extinct species [73]. This method is based on the different spontaneous deamination rates of methylated and unmethylated cytosines, and takes advantage of the characteristic CpG-to-TpG substitution pattern of methylated cytosines that accumulate over thousands of years of chemical degradation. Using this approach, over a thousand DMRs among humans, Neandertals, and Denisovans were identified. Of particular note, one human-specific DMR was located in the HOXD cluster, a key regulator of limb development [74]. Whereas the HOXD9 promoter and HOXD10 gene body are hypomethylated in humans, both archaic species were hypermethylated at this locus, while the gene body of HOXD9 was also hypermethylated in the Denisovan genome. As a result, the authors postulated that this differential methylation in the HOXD cluster could account for some of the anatomical differences between archaic and present-day humans [73]. However, these results are based on the analysis of the genome of just one individual of each species, and thus it is possible that the observed differences could simply represent epigenetic polymorphisms. Furthermore, to determine if these DMRs truly represent human-specific changes, the ancestral methylation state is required. To date, this has not been feasible due to the lack of primate DNA methylation data from bone tissues.

Few studies have addressed DNA methylation patterns across human populations [44,75]. Fraser et al. [75] reported differences in DNA methylation near TSS between European and African populations, although these changes were relatively small in magnitude and did not show any apparent correlation with gene expression levels. Importantly, while over half of these differences could be accounted for by common cis-linked genetic variants, they also acknowledge that more complex genetic interactions and environmental factors may contribute to DNA methylation variation. Heyn et al. [44] reported that methylation patterns characteristic of three distinct human populations (Caucasian, African-American, and Han Chinese) were able to recapitulate the demographic history of each group. Population-specific DMRs, besides being associated with several histone modifications and TF binding sites, were found to occur in genes related to susceptibilities to different diseases and xenobiotic response factors. These findings suggest that certain methylation changes might be due to local selective pressures, such as geographic differences in pathogens or other environmental factors.

There is now accumulating evidence indicating the influence of different environmental factors on methylation patterns [76]. Although it is unclear the extent to which this phenomenon occurs, it is reasonable to assume that different environments could lead to epigenetic divergence. Given that it is not possible to control the environment in most studies involving nonmodel organisms, the results need to be carefully interpreted: some of the observed differences between species or populations might be due to environmental factors and not the result of heritable regulatory changes. Further research is required to understand the crosstalk between the genome, the epigenome, and the environment.

#### Inheritance of Epigenetic Variation

DNA methylation patterns can be influenced by stochastic events or environmental factors creating a source of epigenetic variability [77,78]. Whether this variation could be passed on to future generations has been the subject of intense research [79,80]. A major barrier for its transmission is the robust reprogramming that resets almost all epigenetic marks both in the germline and the zygote [81]. In plants, however, it is a relatively common phenomenon, mainly attributed to the limited reprogramming that occurs in the germline [82–84]. On the contrary, the inheritance of epigenetic changes in mammals is extremely rare. Nonetheless, if some loci were to escape the DNA demethylation process, they would be good candidates for experiencing epigenetic inheritance. Recent genome-wide studies assessing mouse and human germ cell reprogramming have reported a few genes that avoided this erasure [85–87]. Remarkably, some of these genes were related to metabolic and neurological disorders [87]. Other studies have also described cases in which transgenerational epigenetic inheritance might have occurred [88–91]. These epigenetic alterations are unstable over time, and the phenotypic effects disappear in a few generations. Moreover, changes in the genetic sequence are often disregarded as sources of observed heritability.

The notion of transgenerational epigenetic inheritance is very appealing, particularly when suggested that it could respond to environmental challenges [92]. In such situations, it has been proposed that the persistence of epigenetic changes across generations could be considered adaptive and, therefore, it might impact fitness and influence species evolution [93]. Despite this provocative idea, to date there is no solid evidence to substantiate this claim, particularly in mammals, and experimental attempts have failed to prove an adaptive role of epigenetic variation [79].

#### **Future Directions**

Our understanding of DNA methylation and evolution has substantially grown over the past years. Nonetheless, the field of evolutionary epigenomics is still in its infancy, and several steps are required to complete an accurate and detailed picture of the human epigenome.

Despite remarkable achievements accomplished through the parallel survey of genetic, epigenetic, and transcriptional information [40,94,95], a comprehensive evolutionary perspective is required to elucidate the complexity of regulatory mechanisms and to assess the significance of epigenetic changes. Such integrative studies will also be key to interpreting noncoding variation. Besides, many comparative studies disregard the spatial and temporal dynamics of DNA methylation and have exclusively focused on adult organs. Notably, studying embryonic developmental stages is crucial to understanding species evolution, since many of the phenotypically relevant changes occur at these stages. The usage of bulk tissue samples, composed by multiple and variable proportions of cell types, is also problematic. However, the recently developed single-cell bisulfite sequencing will likely overcome this limitation [96].

Future studies also need to determine which genomic regions are more susceptible to experiencing epigenetic changes induced by environmental factors, since such regions could represent confounding factors when studying interspecies epigenetic differences. Studies of methylation patterns in populations with similar genetic backgrounds but different environments would be a good starting point. On the other hand, non-CpG methylation as well as 5mC derivatives have only been recently reported, and they have not been studied in an evolutionary context.

Finally, the major challenge for the next years is to move beyond mere comparative descriptions and offer insight into phenotypic consequences. To that end, experimental assays are required. In this regard, humanized mice have proved useful resources [97,98]. However, they might not be suitable for interpreting certain phenotypes; perhaps induced pluripotent stem cells (iPSCs) could fill this gap [99]. iPSCs can be differentiated to several cell types, providing a system in which to investigate the phenotypic effects of interspecies differences. We foresee an exciting decade for the field.

#### References

- Fu Y, Luo GZ, Chen K, Deng X, Yu M, et al. N6-Methyldeoxyadenosine Marks Active Transcription Start Sites in Chlamydomonas. Cell. 2015; 161: 879–892. doi: <u>10.1016/j.cell.2015.04.010</u> PMID: <u>25936837</u>
- Zhang G, Huang H, Liu D, Cheng Y, Liu X, et al. N6-Methyladenine DNA Modification in Drosophila. Cell. 2015; 161: 893–906. doi: 10.1016/j.cell.2015.04.018 PMID: 25936838
- Greer EL, Blanco MA, Gu L, Sendinc E, Liu J, et al. DNA Methylation on N6-Adenine in C. elegans. Cell. 2015; 161: 868–878. doi: <u>10.1016/j.cell.2015.04.005</u> PMID: <u>25936839</u>
- Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science. 2010; 328: 916–919. doi: <u>10.1126/science.1186366</u> PMID: <u>20395474</u>
- Jones PA. The Role of DNA Methylation in Mammalian Epigenetics. Science. 2001; 293: 1068–1070. PMID: <u>11498573</u>
- Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011; 25: 1010–1022. doi: 10.1101/gad.2037511 PMID: 21576262
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012; 13: 484–492. doi: <u>10.1038/nrg3230</u> PMID: <u>22641018</u>
- Sharp AJ, Stathaki E, Migliavacca E, Brahmachary M, Montgomery SB, et al. DNA methylation profiles of human active and inactive X chromosomes. Genome Res. 2011; 21: 1592–1600. doi: <u>10.1101/gr.</u> <u>112680.110</u> PMID: <u>21862626</u>
- Barlow DP. Methylation and imprinting: from host defense to gene regulation? Science. 1993; 260: 309–310. PMID: <u>8469984</u>
- Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res. 2013; 23: 1256–1269. doi: <u>10.1038/cr.</u> <u>2013.110</u> PMID: <u>23938295</u>
- Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet. 1998; 20: 116–117. PMID: <u>9771701</u>
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009; 462: 315–322. doi: <u>10.1038/nature08514</u> PMID: <u>19829295</u>
- Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature. 2011; 480: 490–495. doi: <u>10.1038/nature10716</u> PMID: <u>22170606</u>
- Bock C, Beerman I, Lien WH, Smith ZD, Gu H, et al. DNA Methylation Dynamics during In Vivo Differentiation of Blood and Skin Stem Cells. Mol Cell. Elsevier; 2012; 47: 633–647.
- 15. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? Mol Cell. Elsevier Inc.; 2013; 49: 825–837.

- Schübeler D. Function and information content of DNA methylation. Nature. 2015; 517: 321–326. doi: 10.1038/nature14192 PMID: 25592537
- Becker C, Hagmann J, Müller J, Koenig D, Stegle O, et al. Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. Nature. 2011; 480: 245–249. doi: 10.1038/nature10555 PMID: 22057020
- Chen CL, Rappailles A, Duquenne L, Huvet M, Guilbaud G, et al. Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. Genome Res. 2010; 20: 447–457. doi: <u>10.</u> <u>1101/gr.098947.109</u> PMID: <u>20103589</u>
- Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, et al. Global epigenomic reconfiguration during mammalian brain development. Science. 2013; 341: 1237905. doi: <u>10.1126/science.1237905</u> PMID: <u>23828890</u>
- Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature. 2015; 523: 212–216. doi: <u>10.1038/</u> nature14465 PMID: 26030523
- Wen L, Li X, Yan L, Tan Y, Li R, et al. Whole-genome analysis of 5-hydroxymethylcytosine and 5methylcytosine at base resolution in the human brain. Genome Biol. 2014; 15: R49. doi: <u>10.1186/gb-2014-15-3-r49</u> PMID: <u>24594098</u>
- Wang T, Pan Q, Lin L, Szulwach KE, Song CX, et al. Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. Hum Mol Genet. 2012; 21: 5500–5510. doi: 10.1093/hmg/dds394 PMID: 23042784
- Plongthongkum N, Diep DH, Zhang K. Advances in the profiling of DNA modifications: cytosine methylation and beyond. Nat Rev Genet. 2014; 15: 647–661. doi: <u>10.1038/nrg3772</u> PMID: <u>25159599</u>
- 24. Bachman M, Uribe-Lewis S, Yang X, Burgess HE, Iurlaro M, et al. 5-Formylcytosine can be a stable DNA modification in mammals. Nat Chem Biol. 2015; 11: 555–557. doi: <u>10.1038/nchembio.1848</u> PMID: <u>26098680</u>
- King MC, Wilson AC. Evolution at two levels in humans and chimpanzees. Science. 1975; 188: 107– 116. PMID: 1090005
- Mendizabal I, Keller TE, Zeng J, Yi SV. Epigenetics and evolution. Integr Comp Biol. 2014; 54: 31–42. doi: 10.1093/icb/icu040 PMID: 24838745
- Cooper DN, Mort M, Stenson PD, Ball EV, Chuzhanova NA. Methylation-mediated deamination of 5methylcytosine appears to give rise to mutations causing human inherited disease in CpNpG trinucleotides, as well as in CpG dinucleotides. Hum Genomics. 2010; 4: 406–410. PMID: 20846930
- 28. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature. 2005; 437: 69–87. PMID: <u>16136131</u>
- Cohen NM, Kenigsberg E, Tanay A. Primate CpG islands are maintained by heterogeneous evolutionary regimes involving minimal selection. Cell. 2011; 145: 773–786. doi: <u>10.1016/j.cell.2011.04.024</u> PMID: <u>21620139</u>
- Bird AP. DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res. 1980; 8: 1499– 1504. PMID: <u>6253938</u>
- Wachter E, Quante T, Merusi C, Arczewska A, Stewart F, et al. Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. Elife. 2014; 3: e03397. doi: <u>10.7554/eLife.03397</u> PMID: <u>25259796</u>
- Krebs AR, Dessus-Babus S, Burger L, Schübeler D. High-throughput engineering of a mammalian genome reveals building principles of methylation states at CG rich regions. Elife. 2014; 3: e04094. doi: <u>10.7554/eLife.04094</u> PMID: <u>25259795</u>
- Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. Nat Rev Genet. 2009; 10: 295–304. doi: <u>10.1038/nrg2540</u> PMID: <u>19308066</u>
- Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007; 448: 714–717. PMID: <u>17687327</u>
- Long HK, Blackledge NP, Klose RJ. ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. Biochem Soc Trans. 2013; 41: 727–740. doi: <u>10.1042/BST20130028</u> PMID: <u>23697932</u>
- Bell CG, Wilson GA, Beck S. Human-specific CpG "beacons" identify human-specific prefrontal cortex H3K4me3 chromatin peaks. Epigenomics. 2014; 6: 21–31. doi: <u>10.2217/epi.13.74</u> PMID: <u>24579944</u>
- Pollard KS, Salama SR, Lambert N, Lambot MA, Coppens S, et al. An RNA gene expressed during cortical development evolved rapidly in humans. Nature. 2006; 443: 167–172. PMID: <u>16915236</u>
- Hodges E, Molaro A, Dos Santos CO, Thekkat P, Song Q, et al. Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. Mol Cell. 2011; 44: 17–28. doi: 10.1016/j.molcel.2011.08.026 PMID: 21924933

- Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013; 500: 477–481. doi: <u>10.1038/nature12433</u> PMID: <u>23925113</u>
- Tsankov AM, Gu H, Akopian V, Ziller MJ, Donaghey J, et al. Transcription factor binding dynamics during human ES cell differentiation. Nature. 2015; 518: 344–349. doi: <u>10.1038/nature14233</u> PMID: <u>25693565</u>
- **41.** Prendergast JGD, Chambers EV, Semple CAM. Sequence-level mechanisms of human epigenome evolution. Genome Biol Evol. 2014; 6: 1758–1771. doi: <u>10.1093/gbe/evu142</u> PMID: <u>24966180</u>
- 42. Shoemaker R, Deng J, Wang W, Zhang K. Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. Genome Res. 2010; 20: 883–889. doi: <u>10.1101/gr.104695.109</u> PMID: <u>20418490</u>
- **43.** Zhang D, Cheng L, Badner JA, Chen C, Chen Q, et al. Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain. Am J Hum Genet. 2010; 86: 411–419. doi: <u>10.1016/j.ajhg.</u> 2010.02.005 PMID: 20215007
- Heyn H, Moran S, Hernando-Herraez I, Sayols S, Gomez A, et al. DNA methylation contributes to natural human variation. Genome Res. 2013; 23: 1363–1372. doi: <u>10.1101/gr.154187.112</u> PMID: <u>23908385</u>
- **45.** Banovich NE, Lan X, McVicker G, van de Geijn B, Degner JF, et al. Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. PLoS Genet. 2014; 10: e1004663. doi: <u>10.1371/journal.pgen.1004663</u> PMID: <u>25233095</u>
- **46.** Hernando-Herraez I, Heyn H, Fernandez-Callejo M, Vidal E, Fernandez-Bellon H, et al. The interplay between DNA methylation and sequence divergence in recent human evolution. Nucleic Acids Res. 2015.
- Suzuki S, Shaw G, Kaneko-Ishino T, Ishino F, Renfree MB. The evolution of mammalian genomic imprinting was accompanied by the acquisition of novel CpG islands. Genome Biol Evol. 2011; 3: 1276–1283. doi: <u>10.1093/gbe/evr104</u> PMID: <u>22016334</u>
- Kanber D, Berulava T, Ammerpohl O, Mitter D, Richter J, et al. The human retinoblastoma gene is imprinted. PLoS Genet. 2009; 5: e1000790. doi: <u>10.1371/journal.pgen.1000790</u> PMID: <u>20041224</u>
- Xiao S, Xie D, Cao X, Yu P, Xing X, et al. Comparative epigenomic annotation of regulatory DNA. Cell. 2012; 149: 1381–1392. doi: <u>10.1016/j.cell.2012.04.029</u> PMID: <u>22682255</u>
- 50. Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, et al. Enhancer Evolution across 20 Mammalian Species. Cell. 2015; 160: 554–566. doi: <u>10.1016/j.cell.2015.01.006</u> PMID: <u>25635462</u>
- Stergachis AB, Neph S, Sandstrom R, Haugen E, Reynolds AP, et al. Conservation of trans-acting circuitry during mammalian regulatory evolution. Nature. 2014; 515: 365–370. doi: <u>10.1038/nature13972</u> PMID: <u>25409825</u>
- Long HK, Sims D, Heger A, Blackledge NP, Kutter C, et al. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. Elife. 2013; 2: e00348. doi: <u>10.7554/eLife.00348</u> PMID: <u>23467541</u>
- Cain CE, Blekhman R, Marioni JC, Gilad Y. Gene expression differences among primates are associated with changes in a histone epigenetic modification. Genetics. 2011; 187: 1225–1234. doi: <u>10.1534/genetics.110.126177</u> PMID: <u>21321133</u>
- Zhou X, Cain CE, Myrthil M, Lewellen N, Michelini K, et al. Epigenetic modifications are associated with inter-species gene expression variation in primates. Genome Biol. 2014; 15: 547. PMID: <u>25468404</u>
- Ward MC, Wilson MD, Barbosa-Morais NL, Schmidt D, Stark R, et al. Latent regulatory potential of human-specific repetitive elements. Mol Cell. 2013; 49: 262–272. doi: <u>10.1016/j.molcel.2012.11.013</u> PMID: <u>23246434</u>
- 56. Molaro A, Hodges E, Fang F, Song Q, McCombie WR, et al. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. Cell. 2011; 146: 1029–1041. doi: <u>10.1016/j.cell.</u> <u>2011.08.016</u> PMID: <u>21925323</u>
- Martin DIK, Singer M, Dhahbi J, Mao G, Zhang L, et al. Phyloepigenomic comparison of great apes reveals a correlation between somatic and germline methylation states. Genome Res. 2011; 21: 2049– 2057. doi: <u>10.1101/gr.122721.111</u> PMID: <u>21908772</u>
- 58. Enard W, Fassbender A, Model F, Adorján P, Pääbo S, et al. Differences in DNA methylation patterns between humans and chimpanzees. Curr Biol. 2004; 14: R148–9. PMID: <u>15027464</u>
- Farcas R, Schneider E, Frauenknecht K, Kondova I, Bontrop R, et al. Differences in DNA methylation patterns and expression of the CCRK gene in human and nonhuman primate cortices. Mol Biol Evol. 2009; 26: 1379–1389. doi: <u>10.1093/molbev/msp046</u> PMID: <u>19282513</u>
- **60.** Fukuda K, Ichiyanagi K, Yamada Y, Go Y, Udono T, et al. Regional DNA methylation differences between humans and chimpanzees are associated with genetic changes, transcriptional divergence and disease genes. J Hum Genet. 2013; 58: 446–454. doi: 10.1038/jhg.2013.55 PMID: 23739127

- Zeng J, Konopka G, Hunt BG, Preuss TM, Geschwind D, et al. Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. Am J Hum Genet. 2012; 91: 455–465. doi: <u>10.1016/j.ajhg.2012.07.024</u> PMID: <u>22922032</u>
- Hernando-Herraez I, Prado-Martinez J, Garg P, Fernandez-Callejo M, Heyn H, et al. Dynamics of DNA methylation in recent human and great ape evolution. PLoS Genet. 2013; 9: e1003763. doi: <u>10.1371/</u> journal.pgen.1003763 PMID: 24039605
- Cáceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, et al. Elevated gene expression levels distinguish human from non-human primate brains. Proc Natl Acad Sci U S A. 2003; 100: 13030–13035. PMID: <u>14557539</u>
- Uddin M, Wildman DE, Liu G, Xu W, Johnson RM, et al. Sister grouping of chimpanzees and humans as revealed by genome-wide phylogenetic analysis of brain gene expression profiles. Proc Natl Acad Sci USA. 2004; 101: 2957–2962. PMID: <u>14976249</u>
- Blekhman R, Oshlack A, Chabot AE, Smyth GK, Gilad Y. Gene regulation in primates evolves under tissue-specific selection pressures. PLoS Genet. 2008; 4: e1000271. doi: <u>10.1371/journal.pgen.1000271</u> PMID: <u>19023414</u>
- 66. Khaitovich P, Hellmann I, Enard W, Nowick K, Leinweber M, et al. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science. 2005; 309: 1850–1854. PMID: 16141373
- Pai AA, Bell JT, Marioni JC, Pritchard JK, Gilad Y. A genome-wide study of DNA methylation patterns and gene expression levels in multiple human and chimpanzee tissues. PLoS Genet. 2011; 7: e1001316. doi: <u>10.1371/journal.pgen.1001316</u> PMID: <u>21383968</u>
- Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nat Genet. 2013; 45: 1198–1206. doi: <u>10.1038/ng.2746</u> PMID: <u>23995138</u>
- Cotney J, Leng J, Yin J, Reilly SK, DeMare LE, et al. The evolution of lineage-specific regulatory activities in the human embryonic limb. Cell. 2013; 154: 185–196. doi: <u>10.1016/j.cell.2013.05.056</u> PMID: <u>23827682</u>
- Reilly SK, Yin J, Ayoub AE, Emera D, Leng J, et al. Evolutionary changes in promoter and enhancer activity during human corticogenesis. Science. 2015; 347: 1155–1159. doi: <u>10.1126/science.1260943</u> PMID: <u>25745175</u>
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, et al. A draft sequence of the Neandertal genome. Science. 2010; 328: 710–722. doi: 10.1126/science.1188021 PMID: 20448178
- 72. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, et al. A high-coverage genome sequence from an archaic Denisovan individual. Science. 2012; 338: 222–226. doi: <u>10.1126/science.1224344</u> PMID: <u>22936568</u>
- 73. Gokhman D, Lavi E, Prüfer K, Fraga MF, Riancho JA, et al. Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. Science. 2014; 344: 523–527. doi: <u>10.1126/science.1250368</u> PMID: 24786081
- 74. Goodman FR. Limb malformations and the human HOX genes. Am J Med Genet. 2002; 112: 256–265. PMID: <u>12357469</u>
- 75. Fraser HB, Lam LL, Neumann SM, Kobor MS. Population-specificity of human DNA methylation. Genome Biol. 2012; 13: R8. doi: 10.1186/gb-2012-13-2-r8 PMID: 22322129
- Dominguez-Salas P, Moore SE, Baker MS, Bergen AW, Cox SE, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. Nat Commun. Nature Publishing Group; 2014; 5: 3746.
- 77. Shah S, McRae AF, Marioni RE, Harris SE, Gibson J, et al. Genetic and environmental exposures constrain epigenetic drift over the human life course. Genome Res. 2014; 24: 1725–1733. doi: <u>10.1101/gr.</u> <u>176933.114</u> PMID: <u>25249537</u>
- Kaminsky ZA, Tang T, Wang SC, Ptak C, Oh GHT, et al. DNA methylation profiles in monozygotic and dizygotic twins. Nat Genet. 2009; 41: 240–245. doi: <u>10.1038/ng.286</u> PMID: <u>19151718</u>
- Heard E, Martienssen RA. Transgenerational epigenetic inheritance: myths and mechanisms. Cell. 2014; 157: 95–109. doi: <u>10.1016/j.cell.2014.02.045</u> PMID: <u>24679529</u>
- Lim JP, Brunet A. Bridging the transgenerational gap with epigenetic memory. Trends Genet. 2013; 29: 176–186. doi: 10.1016/j.tig.2012.12.008 PMID: 23410786
- 81. Feng S, Jacobsen SE, Reik W. Epigenetic reprogramming in plant and animal development. Science. 2010; 330: 622–627. doi: <u>10.1126/science.1190614</u> PMID: <u>21030646</u>
- Schmitz RJ, Ecker JR. Epigenetic and epigenomic variation in Arabidopsis thaliana. Trends Plant Sci. 2012; 17: 149–154. doi: <u>10.1016/j.tplants.2012.01.001</u> PMID: <u>22342533</u>

- Weigel D, Colot V. Epialleles in plant evolution. Genome Biol. 2012; 13: 249. doi: <u>10.1186/gb-2012-13-10-249</u> PMID: <u>23058244</u>
- Hauser MT, Aufsatz W, Jonak C, Luschnig C. Transgenerational epigenetic inheritance in plants. Biochim Biophys Acta. 2011; 1809: 459–468. doi: <u>10.1016/j.bbagrm.2011.03.007</u> PMID: <u>21515434</u>
- Hackett JA, Zylicz JJ, Surani MA. Parallel mechanisms of epigenetic reprogramming in the germline. Trends Genet. 2012; 28: 164–174. doi: <u>10.1016/j.tig.2012.01.005</u> PMID: <u>22386917</u>
- 86. Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, et al. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell. 2012; 48: 849–862. doi: <u>10.1016/j.molcel.2012.11.001</u> PMID: <u>23219530</u>
- Tang WWC, Dietmann S, Irie N, Leitch HG, Floros VI, et al. A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development. Cell. 2015; 161: 1453–1467. doi: <u>10.1016/j.cell.</u> 2015.04.053 PMID: 26046444
- Morgan HD, Sutherland HG, Martin DI, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. Nat Genet. 1999; 23: 314–318. PMID: <u>10545949</u>
- Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, et al. Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. Proc Natl Acad Sci USA. 2003; 100: 2538–2543. PMID: <u>12601169</u>
- Carone BR, Fauquier L, Habib N, Shea JM, Hart CE, et al. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. Cell. 2010; 143: 1084–1096. doi: 10.1016/j.cell.2010.12.008 PMID: 21183072
- Dias BG, Ressler KJ. Parental olfactory experience influences behavior and neural structure in subsequent generations. Nat Neurosci. 2014; 17: 89–96. doi: <u>10.1038/nn.3594</u> PMID: <u>24292232</u>
- Jablonka E, Raz G. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Q Rev Biol. 2009; 84: 131–176. PMID: <u>19606595</u>
- Richards EJ. Inherited epigenetic variation—revisiting soft inheritance. Nat Rev Genet. 2006; 7: 395– 401. PMID: <u>16534512</u>
- Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489: 57–74. doi: <u>10.1038/nature11247</u> PMID: <u>22955616</u>
- Consortium RE, Kundaje A, Meuleman W, Ernst J, Bilenky M, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518: 317–330. doi: <u>10.1038/nature14248</u> PMID: <u>25693563</u>
- 96. Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods. 2014; 11: 817–820. doi: <u>10.1038/</u> <u>nmeth.3035</u> PMID: <u>25042786</u>
- Enard W, Gehre S, Hammerschmidt K, Hölter SM, Blass T, et al. A Humanized Version of Foxp2 Affects Cortico-Basal Ganglia Circuits in Mice. Cell. 2009; 137: 961–971. doi: <u>10.1016/j.cell.2009.03.041</u> PMID: 19490899
- Prabhakar S, Visel A, Akiyama JA, Shoukry M, Lewis KD, et al. Human-specific gain of function in a developmental enhancer. 2008; 7811: 1346–1351.
- Gallego-Romero I, Pavlovic BJ, Hernando-Herraez I, Zhou X, Ward MC, et al. A panel of induced pluripotent stem cells from chimpanzees: a resource for comparative functional genomics. Elife. 2015; 4: 1– 29.