# Insights in Human Epigenomic Dynamics through Comparative Primate Analysis

**Christopher G. Bell**[1-3]

[1]Epigenomic Medicine, Biological Sciences, University of Southampton;
[2]MRC Lifecourse Epidemiology Unit, Southampton General Hospital, University of Southampton; &
[3]Institute of Developmental Sciences, Southampton General Hospital, University of Southampton; Southampton, United Kingdom

## Abstract

Epigenomic analysis gives a molecular insight into cell-specific genomic activity. It provides a detailed functional plan to dissect an organism, tissue by tissue. Therefore comparative epigenomics may increase understanding of human-acquired traits, by revealing regulatory changes in systems such as the neurological, musculoskeletal, and immunological.

Enhancer loci evolve fast by hijacking elements from other tissues or rewiring and amplifying existing units for human-specific function. Promoters by contrast often require a CpG dense genetic infrastructure. Specific interplay occurs between the two, but also a shared modality of function, with coordination from global chromatin-modifying enzymes. Changes in specific transcription factor binding sites also facilitate the local epigenetic state. In the case of CTCF, these may further influence 3-dimensional structure and interaction.

How these mechanistic units are modulated between tissue and species enables more comprehensive understanding of human processes and pathology. With this, precise therapeutic targeting of these epigenetic modifications may become possible.

*Keywords*

Epigenomics, Epigenetics, Comparative Genomics, DNA methylation, Chromatin, Histone Modifications

## Introduction

Our understanding of the evolution of our species, *Homo sapiens*, will be transformed in the coming decades [1]. Genomic tools are now able to uncover molecular mechanisms not only comparatively across primate species [2], but also to peer into the past as demonstrated by the recent insights into Neanderthal, Denisovan, and still other unknown extinct hominins [3, 4]. These ancient DNA studies reveal that incorporated or 'introgressed' archaic DNA is found within the non-African modern human populations of today. This has both potential adaptive and disease susceptibility legacies [5-7], for example, evidence that the Denisovan genetic inheritance within the Tibetan population has aided their extreme environment adaptation [8].

Additionally, it is not only archaic genomes that are being explored, but also the computational reconstruction of the ancestral epigenetic state in hominins. The archaic DNA methylome can be indirectly estimated by quantifying the known degradation of methylated and unmethylated cytosines to thymine and uracil, respectively [9]. This has instigated a fascinating intersection of the fields of epigenetic, comparative and archaic genomics.

Epigenetic factors are principally the devices that each cell requires to regulate specialised functioning of its genome. They are the chemical marks and packaging of a genome that control the expression of the correct genes, at the appropriate time or under a particular condition. These include chemical modifications of DNA, post-translational modifications or variants of the histone proteins that DNA wraps around, and the complex interplay of certain non-coding RNA species [10]. This co-ordinated mechanism enables the cell to be propelled through development and into the specialised niche required for its synergistic role within the human body. The precise co-ordination of these mechanisms is therefore critical. Consequently, controlling the trajectory of global epigenetic modifiers is proposed to play a significant role in evolutionary change [11].

Epigenomic investigations may therefore lead to powerful functional insights, for example into the hidden complexity of the human brain [12]. A longstanding proposition is that epigenetic marks are a mechanism of molecular memory [13].

Decoding neuronal epigenomes and furthermore quantifying the plasticity of cellular states that enable complex higher functioning will help illuminate these biological processes. Therefore, primate comparison of local epigenetic factors as well as genetic changes in global modifying enzymes that mould the epigenome, will increase our understanding of human development and cognition [14, 15]. However, it is not only neurological disorders but all aspects of our evolutionarily-acquired vulnerabilities that may be informed by these comparative studies, including metabolic, immune and musculoskeletal [16]. This review will survey recent insights into the function of the human epigenome from comparative primate studies. The importance of sequence variation in facilitating the epigenetic machinery will be a major focus.

## Chromatin Modifications

Chromatin comprises the dynamic packaging of DNA that also includes regulatory cues [17]. Evidence from yeast models indicate the proteins involved, histones, can themselves maintain their epigenetic state through somatic mitosis independently of associated DNA sequences [18, 19]. Canonical histone modifications involve the modification of lysines (K) present in the tail of the histone 3 subunit (H3). These include signatures within latent promoter (H3K4me3) and enhancer regions (H3K4me1); an activity mark within both these regions (H3K27ac [20]); constitutive (H3K9me3) and facultative (H3K27me3) heterochromatin repressed regions [21]; as well as a combination of active and repressed marks indicating bivalent promoters [10]. Assaying a range of these chromatin marks has enabled combinatorial segmentation of the genome into demarcated tissue-specific functional regions [22]. These enhancer predictions [23], can be further functionally supported with associated expression of short enhancer-RNA (eRNA) [24].

The epigenetic machinery includes writers, erasers, readers and remodellers of the epigenome. It has been identified that there is a high mutational load in chromatin-modifying genes in developmental disorders [25]. These monogenic disorders strongly indicate the importance of chromatin dynamics and the

associated phenotypes commonly including significant behavioural and intellectual disabilities [26]. It is also of note that these genes have also been observed as somatic hotspots in cancers [27]. This indicates their critical function in a time-independent fashion [28] and further underlines the importance of the accurate orchestration of the cellular epigenome.

## DNA Modifications

Epigenetic modifications of DNA include the most common epigenetic mark, methylation of the 5' carbon of cytosine (5mC) [29]. In differentiated cells this occurs predominately in a CpG dinucleotide context, although non-CpG methylation is more prevalent in the brain. The default state of CpGs throughout the genome is methylated [30], the notable exceptions being dense clusters of CpGs termed 'CpG islands' (CGIs) that are predominately unmethylated. The low density and methylated CpGs are at an increased risk of loss over evolutionary time due to hypermutability and lack of recognition for repair [31]. Inversely, small increased GC clusters, including CpGs, can be generated as a by-product of recombination-associated Biased Gene Conversion (BGC) [31]. However, some lower density regions may possess an average low but not completely unmethylated DNA methylation state, previously identified as Low Methylation Regions (LMRs). These regions co-locate with enhancer evidence including H3K4me1 and the enhancer-related co-activator p300 [32].

The CpG dinucleotide is proposed to act as a genome-wide signalling molecule in its own right, recruiting factors dependent upon its methylation state [33]. Furthermore distinct genetic motifs influence methylation state, particularly with CpG dense regions. These Methylation Determining Regions (MDR) include particular transcription factor binding site (TFBS) motifs, such as SP1, CTCF and members of the RFX family [34]. Additionally, other short sequences including AT motifs are suggested to also direct and influence the epigenome [35].

The active removal of the DNA methylation mark occurs via an oxidative process catalysed by the TET enzymes [36]. This leads to the formation of

hydroxymethylated cytosine (5hmC) with further oxidative steps deriving 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), and then finally onto the unmodified cytosine base. The enzyme TDG may also be involved in these later changes. These modifications of DNA are recognised by particular readers [33]. Those that recognise 5mC include the methyl-CpG-binding protein (MECP2) and the methyl-CpG-binding-domain proteins (MBD1, MBD2 and MBD4), which enlist further molecules to create complexes that interact with chromatin remodellers [37].

The severe post-natal developmental decline that occurs in females with Rett syndrome is caused by mutation within the X-linked *MECP2* gene. It clearly demonstrates the importance of this epigenetically-selective 5mC binding protein in the maturation of the central nervous system [38]. Furthermore, evidence in mice indicates it is required even in adulthood for correct neurological function [39] and additionally overexpression of *MECP2* leads cynomolgus monkeys to display autistic behaviours [40]. Data has accrued that the MECP2 protein may also bind 5hmC in neuronal cells [41], as well as non-CpG methylated cytosines as the neuron matures [42, 43], indicating the potential epigenetic complexity in brain disorders [44].

The plasticity of DNA methylation enhances the gene transcriptional repertoire of neurons and is proposed to influence the synaptic wiring implicated in memory [45]. In post-mitotic neurons widespread active DNA methylation removal occurs, with TET enzymes playing a prominent role in synaptic activity and downstream consequences on gene expression [46]. Additionally, allele-specific parent-of-origin DNA methylation, termed imprinted loci, leads to allelic expression biases in brain tissue that can vary through different anatomical regions and over time. This thus further displays the intricacy of these epigenetic mechanisms [47].

## Contribution of Chromatin Modifications to Transcriptional Variation in Species

One of the first epigenomic primate comparative studies was performed by Cain *et al*. [48]. This study compared the promoter chromatin mark H3K4me3 genome-wide, across human, chimpanzee and rhesus macaque in lymphoblastoid cell lines (LCLs) by ChIP-seq (Chromatin Immunoprecipitation 2$^{nd}$-generation

sequencing) (and see Table 1 for Comparative Studies). Although H3K4me3 is found at promoters, it is not definitely indicative of their activity but merely their permissive potential. Cell line derived data is also subject to stochastic epigenetic variation due to its generation by immortalisation [49]. Despite these limitations this study estimated that ~7% of the comparative expression differences could be attributed to detected species-specific chromatin variation.

A more comprehensive analysis was performed later by the same research group in Zhou *et al*. [50], including now the combination of RNA polymerase II (Pol II) binding, and four histone modifications; H3K4me3, H3K4me1, H3K27ac and H3K27me3 [20]. This was in eight samples from each species. Whilst again suffering from the potential drawbacks of epigenomic analysis in LCLs, a combination of marks explained ~40% of the gene expression differences between the species. Although chromatin mark abundance measures are imprecise, this study found that it was not just Pol II co-localisation that was the sole determinate. The combinations preformed ~5% better than any individual mark alone and the well-known configuration of H3K4me3, H3K27ac, and Pol II was consistent with promoter transcription. However, pinpointing causality in these highly correlated regulatory interactions is difficult [51]. Comparative TFBS motif sequence changes within these regions were not implicitly explored but acknowledged as a potential driving influence. As expected, less difference was identified between human and chimpanzee than comparison of either of these species to the more distantly related rhesus macaque.

## Shared Features and Connections Between Promoters and Enhancers

A comparative analysis of promoter associated chromatin mark H3K4me3 was performed in prefrontal cortex (PFC) neurons, again between human, chimpanzee and rhesus macaque [52]. This is the cortical region associated with higher functioning, personality, and the evolution of the primate brain. H3K4me3 changes within this region have previously been associated with autism [53]. This study from Shulha *et al*. identified 410 human-specific gains out of a total of 34,639

human H3K4me3 peaks through comparative analysis and a further 61 loci where there was a human-specific loss. Of these 410 gains of H3K4me3, there was a twofold enrichment for peaks that reside within 0.5-1 Mb of each other with a proposed role in co-regulation through promoter-enhancer interactions. Whilst H3K4me1 is the predominate modification identified at enhancers, H3K4me3 is also found within defined strong enhancer regions and is seen within 75% of ChromHMM state 4 Strong Enhancers [54]. These data also support the architectural commonality that is becoming increasingly recognised between promoter and enhancers, with operative interchangeability and role switching between tissues or time-course [55] (see Figure 1). Both functional units share many features, including core promoter sequence elements, similar transcription factor binding and divergent RNA transcription [56]. Chromosome conformation capture data identified interaction between two peaks within *DPP10*, a gene previously associated with disorders such as autism [57] and bipolar disorder [58]. Of note, a neuronal antisense RNA, LOC389023, was also identified to originate from here. These same human-specific chromatin regions within *DPP10* had been previously identified to be enriched for human-specific CpG dinucleotides [59].

## Influence of CpG Clusters on Chromatin Structure

A six primate comparative sequence comparison identified human-specific CpGs (termed CpG 'beacons') [59] (See Figure 2). A dramatic increase in CpG 'beacons' was also shown to be strongly predictive of the human-specific H3K4me3 state from the above Shulha *et al*. dataset [60]. This is consistent with the influence of clustered unmethylated CpGs, as shown experimentally by Thomson *et al*., to lead to the formation of H3K4me3 by the binding of the lysine trimethyltransferase CFP1 via its CxxC domain [61]. Additionally, these CpG 'beacon' related human-specific H3K4me3 peaks were found to be enriched within the chromosome 2q fusion (2qfus) region, including the above mentioned locus of *DPP10* [60]. This is the historic site of the fusion of the two separate chromosomes comprising human chromosome 2 that are present in other primates. This 2qfus finding was also later

supported by Giannuzzi *et al.*, who also showed human-specific H3K4me3 enrichment in other human-specific cytogenic structures; the chromosome 1 and 18 pericentric inversion breakpoints [62]. In further analysis of PFC neurons, Dincer *et al*. identified that the broadest H3K4me3 domains (≥ 10 kb) related to synaptic function and interestingly were the most conserved across human and other primates [63]. Broad H3K4me3 peaks are also more predictive for cell-specific identity [64].

## Functional Rerouting and Up-regulation for Species-Specific Enhancer Formation

The evolution of mammalian enhancers was traced across 20 mammals by analysing liver tissue from these species for promoters and enhancers, via H3K4me3 and H3K27ac peaks [65]. Thus, the change of these genomic functional features could be examined over 180 million years of divergence. This identified far more rapid evolution in enhancers compared to promoters over this time course. Newly acquired enhancers were enriched for proximity to positively-selected protein-coding genes and were derived from ancestral DNA that had been functionally rerouted, not newly acquired DNA sequence.

Humans have evolved distinctive variation between our sets of limbs, with long legs and short arms, that distinguishes us even from other primates. This is due to independent adaptation including manual dexterity, bipedalism and endurance running [66]. The molecular mechanisms influencing this change were explored by ChIP-seq analysis of embryonic limb tissue through developmental stages in human, rhesus macaque and mouse, by comparison of the location of the activating histone modification H3K27ac [67]. Of the total promoters and enhancers, approximately 13% and 11% respectively, revealed a significant human-specific increase. The vast majority (~91%) of human limb gain-of-enhancer function was via the modification of enhancers that were also present in both rhesus macaque and mouse limbs. Whereas a small proportion (~6%) were recruited in human from other tissues in the other species. The small remainder (~3%) represented new enhancers found only in human. The highly conserved essential regulators of limb development, such as

*PITX1* and *TBX5*, showed no significant change in promoter or enhancer profiles across the species. Candidates for potential human-specific change were *ARHGAP6*, which has a markedly increased expression in human. In mouse, a mutation within a regulatory region in the orthologous locus is associated with polydactyly, as well as thickening and elongation of digit one in the hindlimb [68]. Sixteen regions that show human-specific genetic acceleration (HARs: Human Accelerated Region [69] or HACNS: Human Accelerated Conserved Noncoding Sequence [70]) also co-located with significant human-specific H3K27ac increase, including the HACNS1 locus [71].

## Non-Neuronal Regulatory Mechanism Hijacking in the Cortex

When comparing human to other mammals, development differences in the cerebral cortex begin within the first trimester with an increase in number, range, and migration trajectories of neurons [72]. A comparative study of cortical development performed by Reilly *et al*. [73], analysed promoter and enhancer profiles in human, rhesus macaque, and mouse by H3K27ac and H3K4me2 (a co-locating signal within both promoter and enhancers). This was implemented through stages 7, 8.5, and 12 post-conception weeks where human-specific changes in layers and features begin to develop. This was compared with developmentally equivalent time points in the rhesus macaque and mouse embryo. Approximately 85% of human promoters possessed collocating H3K4me2 and H3K27ac peaks but this was only 45% at enhancers, as these two marks separately recognised distinct enhancer loci [54]. Of this total set of 52,317 enhancers identified in a least one of the developmental stages, 16,473 were found to be potentially cortex-specific and enriched for proximity with genes associated with its development. 8,996 enhancers and 2,855 promoters were identified with human-specific gains and highlighted 301 genes with gain-enrichment hotspots. Of these gain regulatory loci, 48 showed overlap with previously genetically identified human-accelerated regions (HARs or HACNSs). Enrichments for gene co-expression networks included neuronal progenitor proliferation, homeobox genes, and the extracellular matrix (aiding progenitor cell renewal and neuronal migration). The enriched TFBS motifs in

common pathways implied that non-neuronal regulatory mechanisms had been harnessed, or older mammalian regulatory mechanisms in cortex developmental were hijacked, for human-specific adaptation. The co-ordinated changes needed to generate the human cortex therefore used and modified known developmental regulatory regions, in similar fashion to those identified above in the human limb.

## Critical Transcription Factor Modifications in Enhancer Loci

Human and chimpanzee facial structures have evolved distinct species-specific differences and in order to explore this, *in-vitro* derived human and chimpanzee cranial neural crest cells were compared [74]. Epigenomic analysis was performed by ChIP-seq for H3K4me1, H3K4me3, and H3K27ac. Also the enhancer related co-activator p300 was examined, as well as the cranial neural crest cell transcription factors (TFs) TFAP2A and NR2F1. Genome accessibility was assessed by a genome-wide assay for transposase-accessible chromatin (ATAC-seq). Employing these multiple strains of evidence, divergent facial enhancers were identified between the species. Although only small sequence variation was seen in divergent enhancers, with ~3–6 substitutions per 500 bp, it was suggested that change in critical key sequence-dependent TFs can considerably modify enhancer activity. Losses of the TFAP2A consensus motif as well as TFBSs for ALX homeobox factors were associated with a loss of H3K27ac. Changes in repressive TFs, where reduced motif strength is associated with increased H3K27ac peaks, included the known transcriptional repressors SNAI2 and the TBX family. The role of repeat classes within these species-specific enhancers was also noted, with endogenous retroviruses (ERV1, ERVL-MaLR, and ERVK) as well as L1 elements (LINE-1 retrotransposons) particularly enriched within these loci. Evidence of variation in the regulation of the primate-active L1s in pluripotent stem cells between humans and other apes is proposed to have differentially shaped these genomes [75].

## Sequence Redundancy In Enhancers

To explore *cis*-regulatory elements Vermunt *et al*. analysed H3K27ac ChIP-seq within eight anatomical regions of the brain in human, chimpanzee and rhesus macaque [76]. They identified that the location of the majority of peaks (~93%) had been conserved since the last common ancestor (LCA) of these three species ~25 million years ago (MYA). Thus supporting previous data that regulatory unit positioning is predominately conserved. As expected, likely promoter regions (co-locating Transcription Start Sites and H3K4me3) were consistent across the different areas of the brains, in contrast to potential enhancers that were more likely to be area-specific. Furthermore, co-locating enhancers were not highly conserved at the sequence level although they shared potential functionality. This lack of sequence constraint is consistent with data that enhancers do not require precise base-pair maintenance to retain function, and can be dynamic and rapid in their evolution [65] as long as critical TFBSs are maintained. As expected, on average only ~22.5% of the differences between human and rhesus macaque in the various tissues, were also seen between human and chimpanzee due to evolutionary distance. In fact, in this study only a small number of loci were defined to be novel in the human brain, 1,399 enhancers and 89 promoters. Each change localised predominately within one brain area, thus supporting the ability for enhancers to change more readily due to their less ubiquitous activity [77]. Although HAR sequences showed enriched overlap with the identified brain regulatory elements, they were not more prevalent at the human-specific enhancers. In another study by Boyd *et al*., a previous identified accelerated non-coding genetic locus [78] was implicated as a Human-specific enhancer, termed Human-Accelerated Regulatory Enhancer 5 (HARE5) of the *FZD8* gene [79]. This is a receptor in the Wnt pathway involved in neural progenitor proliferation and implicated in brain development and size.

## Sequence Influence on the DNA Methylome

The close interplay between genome sequence and CpG methylation may bring further insights into human evolution [80]. Classically, DNA methylation at CGI promoters is a repressive mechanism. However, the now recognised potential

functional role of CGI shores (2 kb either side of CGI) and the tidal nature of methylation changes that occur here [35], make CpG density changes within these regions of interest. Through genetic gain and loss of CpGs, permanent accretion and erosion of islands and change in their demarcated shores can occur. Also in low-density enhancer LMRs there is a potentially disproportionate importance of genetic CpG fluctuations due to the paucity of CpG sites.

Whilst CpG density is important in defining the DNA methylome, the influence of critical MDR TFs in instigating the local DNA methylation state is increasingly acknowledged. These mechanisms will affect the potential for CpGs to be lost due their hypermutability when methylated in the germline. The Influence of the TF Sp1 was documented by Macleod *et al.* in 1994 [81] and recent genome-wide data further supporting this and other significant TFs [34, 82]. Thus in studying the evolutionary divergence of the DNA methylome, investigating the underlying intertwined genetic changes in critical TFBSs and CpGs themselves is also required.

In comparative studies, array analysis, dependent upon probe affinity hybridisation, needs strong awareness of the confounding influence of sequence variation. Known variation can help excluding probes, however non-human primate polymorphic information is less complete. An inter-primate analysis was performed with the Illumina 450k DNA methylation array in peripheral blood [83]. Only less than a quarter of the probes (99,919) were assessed as complementary across five primate species, with 289,007 shared between human and chimpanzee. This identified ~9% of probes between human and chimpanzee were significantly different in methylation state, including 184 genes that were completely conserved at the protein level between the two species. Although the additional limitations of array targeted CpGs must also be recognised, in that they disproportionally lack coverage of the most dynamic CpGs [84].

A peripheral blood study across human, chimpanzee, and rhesus macaque by Wilson *et al*. [85] was performed genome-wide with MeDIP-seq (Methylated DNA Immunoprecipitation 2$^{nd}$-generation sequencing). This observed that species-specific Differentially Methylated Regions (s-DMRs) were more prevalent in CGI shores, in a similar fashion to previously recognised tissue-specific and cancer-specific changes [86]. Within orthologous CGIs, this study also showed increased human CTCF motif

strength in human hypomethylated s-DMRs when compared to both other primates, and the inverse in human hypermethylated s-DMRs. A strong human specific s-DMR in the promoter of the Leukotriene B4 receptor (*LTB4R)* gene was further examined. It showed consistent strong hypomethylation across the blood-cell sub-fractions that was also accompanied by human-specific H3K4me3 peaks and increased expression. The hypomethylated state of the human promoter CGI was likely caused by the RFX1 motif, an MDR [34], that was only present in humans not other primates (See Figure 3). This motif existed within both Neanderthal and Denisovan sequences, thus dating it to prior to ~0.6 MYA.

Martin *et al*. explored comparative neutrophils between three primates, human, chimpanzee, and orang-utan via the enzymatic method Methyl-seq[87]. This study estimated that ~10% of the CGI regions differed significantly between human and chimpanzee. A recent paper was performed with whole-genome bisulphite sequencing (BiS-seq) generating data for ~6 million well-covered CpGs, in peripheral blood across human, chimpanzee, gorilla, and orang-utan [88]. This identified that s-DMRs were also enriched for species-specific nucleotide changes and hypomethylated s-DMRs were specifically enriched twofold for endogenous retroviruses (ERVs). The potential facilitative role of transposable elements particularly ERVs was also seen in a multi-tissue examination of human epigenomes by Leung *et al*.[89] and is discussed recently by Lowdon *et al.* [90]. The PFC was compared between three humans and three chimpanzees by BiS-seq in Zeng *et al*. [91]. They reported in fact a small but discernable level of methylated non-CpG cytosine sites (1.3%–2.2%). Additionally, hundreds of gene promoters were found hypomethylated in human compared to chimpanzee and s-DMRs were enriched for neurological, psychological, and cancer disease loci.

## Germline Influences on the Evolution of DNA Methylation

The methylation state within the germ-line is very important in the evolution of the DNA methylome, as this is where the effect of 5mC hypermutability can lead to loss of CpG dinucleotides in subsequent generations [92]. Molaro *et al.*

investigated the sperm DNA methylome by BiS-seq between human and chimpanzee, and identified extended hypomethylation around CGI regions in comparison to embryonic stem cells, analogous to CGI shore regions [93]. A subsequent tidal reduction in size of these hypomethylated regions was observed with differentiation. A high proportion of hypomethylated human s-DMRs overlapped with the hominid-specific SVA repeats, that consist of an SINE, a VNTR and an *Alu* repeat [94]. This was particularly in the youngest members of this SVA repeat family. The lack of repressive DNA methylation and the mobilisation ability of these repeats enables germline insertions and indicates their significant potential in species-specific variation.

### Clustering of Species-Specific CpGs

With regard to evolutionary change in CpG dinucleotide itself, Tanay *et al*. identified that hyperconserved CpG domains between human and chimpanzee were enriched loci for Polycomb repressive complex 2 and were co-located near genes related to embryoic development [95]. The first estimation of the subset of human-specific CpGs, termed CpG 'beacons' mentioned above, was calculated by sequence comparison with five other primates, chimpanzee, gorilla, orang-utan, rhesus macaque, and marmoset [59], thus encapsulating a divergence time of ~35 MYA (see Figure 2). This estimated that ~1.19 million CpGs that are also non-polymorphic in 1,000 genomes data are human-specific from the total of ~28 million in the human genome. Twenty-one significant clusters of these CpG 'beacons' were determined by permutation and these loci were enriched for neurological disease due to co-location with monogenic neurological and developmental disorders. CpG 'beacons' individually were enriched within SVA repeats, the class that is enriched for sperm hypomethylation from Molaro *et al*.[93]. Furthermore, using primate comparative DNA methylome data, CpG density gain indicated by CpG 'beacons' correlated with reduced methylation in orthologous CGIs in these species over evolutionary time.

DNA methylome profiling across seven diverse vertebrates revealed a reasonably high level of conservation in hypomethylated loci from human to

zebrafish [96]. Although when taken to this range of species, these hypomethylated regions were in fact poorly captured by predicted CGI locations. However where broad non-methylated regions did exist, these were associated with polycomb-regulated developmental genes, consistent with the earlier findings of Tanay *et al*. [95].

## Evolutionary Role of Chromatin Modifiers and Readers

Chromatin modifying enzymes are potentially important in evolution due to their multilocus effects [11]. These include zinc finger proteins that have the ability to interact with cytosines in CGI specificity due to a CxxC domain that can 'read' CpG dense regions [97] (see Table 2). These comprise CFP1, the lysine methyltransferase that trimethylates H3K4 at unmethylated CGIs [61] and Lysine (K) demethylases KDM2A and KDM2B, the later of which protects Polycomb-bound genes from hypermethylation [98]. The Methyl Binding Domain 1 (MBD1) also contains this unmethylated cytosine sense unit as well as its Methyl-Binding Domain, although it is not required for its recruitment to methylated loci. It is postulated this CxxC domain may enable targeting when methylation levels are dramatically reduced, such as in pre-implantation embryos [99]. Other proteins in this group include the maintenance DNA methylator DNMT1, Lysine methytranferases MLL1 (KMT2A), and MLL2 (KMT2D). MLL1 is a significant modifier in both cortical and hippocampal development and has been proposed to play a role in neurobiology [100, 101] including working memory through prefrontal synaptic plasticity [102]. The TET enzymes (TET1, TET3) involved in the active demethylase process and subsequent formation of 5hmC and further derivatives, also use their CxxC domain to bind to unmethylated and hydroxymethylated CpGs [103]. It has also been recently identified that the CxxC domain protein IDAX binds unmethylated DNA, recruits its ancestral partner TET2 that now lacks this ability, and then down-regulates it consequently influencing the level of 5mC oxidation [104].

Other important chromatin modifiers in neurodevelopment include the lysine demethylase, *KDM5C*, with mouse evidence that mutations are involved in X-linked

intellectual disability [105]; the lysine methyltransferase of H3K9, *EHMT1* (*KMT1D*), involved in the developmental disorder Kleefstra syndrome [106]; and *CHD8*, a chromodomain helicase that binds to H3K4me3, which is autism-associated [107].

Accurate control in the sculpting and interpretation of the epigenome is required for normal development and these factors controlling this process are therefore able to have considerable effects in this process. One clear example of the influence of comparative species change in an epigenomic modifier is the major recombination regulator PRDM9 [108, 109]. In humans, it was first identified to recognise a 13bp motif that designates the location of hotspots of recombination via its zinc finger DNA-binding domain. PRDM9 is also a meiosis-specific histone methyltransferase catalysing the trimethylation of H3K4. This leads to nucleosome depletion and thus facilitates the subsequent double strain breaks required for recombination [110]. A comparison across 18 primates, revealed a high level of diversity in the zinc-finger domain due to rapid positive-selection driven evolution. This enables significant hotspot location variation both within and between species [111]. Differential hotspot use can lead to hybrid sterility and is therefore a proposed mechanism in the speciation of primates. Furthermore, PRDM9 degradation of its own binding sites occurs by meiotic drive, the process whereby mutations in hotspots are preferentially transmitted [112]. Biased Gene Conversion (BGC) is implicated in this self-destructive drive [113]. Human hotspots are identified to be young, having only been active 10% of the time since Denisovan divergence (~0.7 to 1.3 MYA) [114]. Comparative analysis of 29 mammalian genomes identified that ~15% of an expanded set of 563 HARs showed evidence of BGC GC-biased gene conversion and 9.4% overlapped chromatin enhancer evidence [115]. These recombination-associated genetic increases in GC [116] can also in cases help build local CpG density for potential promoter epigenetic function [60].

### Insights from the Interplay of the Epigenome and Long Non-Coding RNA

Epigenomic control is exerted by the co-ordinated modifications of the multiple intermeshed mechanisms, chromatin structure, DNA modifications, and

non-coding RNA that act to reinforce each other. As such, the inactivated X-chromosome in females is not reactivated on the administration of the demethylating drug azacytidine [117]. Regulatory factors including TFs and long non-coding RNA (lnc-RNA) are proposed to play a significant role in human evolution [118]. There is extensive RNA binding to chromatin-associated proteins and complexes [119]. These include lnc-RNAs that act as controllers of locus dynamics and molecular scaffolds for chromatin. p53-induced lnc-RNAs are shown to increase enhancer activity with recognised amplification of eRNA expression [120]. Furthermore, lnc-RNA are significantly involved in the regulation of development and fine-tuning of gene networks controlling organogenesis and as such are critical in the evolution of complex organisms [121]. For example, the WD repeat-containing protein 5 (WDR5) is required as a critical subunit by the MLL1 complex for H3K4me3 formation and thereby necessary for vertebrate development [122]. WDR5 itself needs the binding of specific lnc-RNAs essential to maintain this active chromatin state [123]. Recently, a primate-specific lnc-RNA HPAT5 was identified to be an integral regulator of pluripotency in preimplantation development [124]. Further untangling of the relationship of non-coding RNA and the epigenomic machinery will undoubtedly reveal detailed molecular mechanisms in species-specific function.

## Future Directions

The long-standing hypothesis of King and Wilson states that due to their strong coding sequence similarity, gene regulatory differences are the driver in the evolutionary divergence between humans and other closely related primates [125]. These regulatory genetic changes may facilitate epigenetic functional differences as well. Exploring comparative change in both genome and epigenome provides uniquely powerful instruments to understand their integrated function in biology, as well as their defects in pathophysiology. This will bring insights for the future field of epigenomic medicine, where the precise epigenetic modulation of genome activity has much potential for the treatment of human disorders, such as age-related

degenerative diseases [126]. Additionally, synthetic biology has also noted the possibilities of chromatin regulatory modification [127].

Further comparative analyses over more epigenomic mechanisms will undoubtedly deliver new insights, such as the role of histone variants. H3.3 is involved in the establishment of silenced chromatin states and maintenance of genome stability by silencing retrotransposons [128]. Also accumulating analyses indicate that 5hmC is not just a transitional state in the demethylation process, but an active operator in its own right [129]. Dynamic readers of 5hmC [130, 131] as well as 5caC [132] can influence TF binding. Current data indicate that 5hmC almost exclusively occurs within a CpG context [133]. Therefore within LMR enhancers the potential importance of the small number of CpGs where this hydroxymethylation occurs is of definite epigenetic interest.

Epigenetic modifications of RNA is an active field of research that would also be of relevance in comparative analyses. Recent Drosophila data on RNA hydroxymethylation implicated this mark in mRNA translation and blocking its catalyst, the TET enzyme, lead to impaired brain development [134]. Another recent finding requiring further comparative investigation, is the importance of the transcription factor *NRF1*. Its binding significantly influences nearby expression but it is strongly inhibited by DNA methylation [135]. Also, exploration of primate comparative CTCF differences may reveal insights particularly due to its orientation-specific influence on 3-Dimensional genome structure and looping [136].

Due to recent work on the epigenetic changes with age, ancient DNA methylation data from a 4000-yr-old Palaeolithic human was even able to estimate his age at death [137]. Human ageing DNA methylation changes are shown to have consistency with chimpanzees with small-scale array analysis [138], but any comparative differences that could be identified would be of great curiosity as to the cause for the divergence.

Future comparative analysis in homogeneous cells of the intertwined role of transcription factors, both drivers of the epigenomic state and those that are repelled or sensitive to it, will enable further species-specific discoveries. Mechanisms identified in the evolution of cancer cells can also be highly informative with respect to longer time frames[28]. The potential of factors such as CTCF to

modify long-range genomic interactions indicate that comparative analysis of chromatin neighbourhoods, by techniques such as ChIA-PET, may lead to new understanding [139].

Recent epigeneomic insights into neurological disease have come from analysis of the DNA methylation changes [140] as well immune enhancer signatures in Alzheimer's disease [141]. This later use of conserved enhancer signatures across species from human to mouse, was highly informative in identifying the pathogenic role of the immune system. This again demonstrates how these novel epigenomic tools, now and increasingly in the future, can be used to decipher the human condition.

## Figure Legends

Figure 1: Construction of active promoters and enhancers is dependent upon many factors. This includes genetic sequences appropriate to facilitate the correct epigenetic state, such as dense non-5mC CpGs in promoters, enabling H3K4me3 formation by CxxC-domain protein CFP1. Intermediate 5mC but also 5hmC in low density (but above baseline) CpGs present in enhancer loci is observed.

Figure 2:
Identification of Human-specific CpGs (CpG 'Beacons') indicated by arrows through Primate Comparative sequence analysis.

Figure 3: Significant differences in Methylation Determining Region (MDR) Transcription Factor Binding Sites (TFBSs) can lead to Human-specific Differentially Methylated Regions (DMRs). Here human-specific variation in the binding affinity for the *RFX1* motif in the promoter of *LTB4R* is shown from the TRAP algorithm[142]. This results in a human-specific hypomethylated DMR[85]. Comparative primate sequence for this *RFX1* motif in this location is shown below.

## Tables

| Analysis Target | Tissue | Species | Study |
|---|---|---|---|
| H3K4me3 | Lymphoblastoid Cell Lines | Hs, Pt, Mm | Cain *et al*. [48] |
| Pol II, H3K4me3, H3K4me1, H3K27ac, H3K27me3 | Lymphoblastoid Cell Lines | Hs, Pt, Mm | Zhou *et al*. [50] |
| H3K4me3 | Prefrontal Cortex Neurons | Hs, Pt, Mm | Shulha *et al*. [52] |
| H3K4me3, H3K27ac | Liver Hepatocytes | 20 Mammals including Hs, Mm, Csab, Cjac | Villar *et al*. [65] |
| H3K27ac | Embryonic Limb | Hs, Mm and Mouse | Cotney *et al*. [67] |
| H3K4me2, H3K27ac | Embryonic Cortex | Hs, Mm and Mouse | Reilly *et al*. [73] |
| H3K4me1, H3K4me3, H3K27ac, p300, TFAP2A, NR2F1 | Cranial neural crest cells | Hs, Pt | Prescott *et al*. [74] |
| H3K27ac | Brain Regions (8) | Hs, Pt, Mm | Vermunt *et al.* [76] |
| DNA methylation | Peripheral Blood | Hs, Pt, Pp, Gg, Pa/p | Hernando-Herraez *et al.* [83]. |
| DNA methylation | Peripheral Blood | Hs, Pt, Mm | Wilson *et al*. [85] |
| DNA methylation | Neutrophils | Hs, Pt, Pa | Martin *et al.*[87] |
| DNA methylation | Peripheral Blood | Hs, Pt, Gg, Pa | Hernando-Herraez *et al.* [88]. |
| DNA methylation | Prefrontal Cortex Neurons | Hs, Pt | Zeng *et al*. [91] |
| DNA methylation | Sperm | Hs, Pt | Molaro *et al.* [93] |

Table 1: Comparative Epigenomic Studies. Primate species analysed: Hs = *Homo sapiens*; Pt = *Pan troglodytes* (Chimpanzee); Pp = *Pan paniscus* (Bonobo); Gg = *Gorilla gorilla*; Pa/p = *Pongo abelii/pygmaeus* (Orang-utan); Mm = *Macaca Mulatta* (Rhesus macaque); Csab = *Chlorocebus aethiops sabaeus* (Vervet); Cjac = *Callithrix jacchus* (Marmoset).

| Gene | Also Known As | CxxC nomenclature |
|---|---|---|
| *CFP1* | CGBP, PHF18 | CXXC1 |
| *CXXC5* | | CXXC5 |
| *DNMT1* | MET1 | CXXC9 |
| *FBXL19* | | |
| *IDAX* | | CXXC4 |
| *KDM2A* | FBXL11, JHDM1A, NDY2 | CXXC8 |
| *KDM2B* | FBXL10, JHDM1B, NDY1 | CXXC2 |
| *MBD1* | PCM1 | CXXC3 |
| *MLL1* | ALL1, HTRX1, KMT2A | CXXC7 |
| *MLL2* | WBP7, KMT2B | |
| *TET1* | LCX | CXXC6 |
| *TET3* | | CXXC10 |

Table 2: Zinc Finger - CxxC domain containing genes (adapted from Long *et al*.) [97].

## Acknowledgements

CGB receives support from the EpiGen Global Research Consortium (www.epigengrc.com) and the MRC Lifecourse Epidemiology Unit.

## Abbreviations

5caC: 5-carboxylcytosine

5fC: 5-formylcytosine

5hmC: 5-hydroxymethylcytosine

5mC: 5-methylcytosine

ATAC-seq: Assay for Transposase-Accessible Chromatin 2$^{nd}$-generation sequencing

BiS-seq: Bisuphite 2$^{nd}$-generation sequencing

BGC: Biased Gene Conversion

CGI: CpG islands

ChIP-seq: Chromatin Immunoprecipitation 2$^{nd}$-generation sequencing

eRNA: enhancer-RNA

ERV: endogenous retrovirus

HAR: Human Accelerated Regions

HACN: Human Accelerated Conserved Noncoding

KDM: Lysine Demethylase

KMT: Lysine Methytransferases

LCA: Last Common Ancestor

LCL: Lymphoblastoid Cell Lines

LMR: Low Methylation Region

MBD: Methyl Binding Domain

MDR: Methylation Determining Region

MeDIP-seq: Methylated DNA Immunoprecipitation 2$^{nd}$-generation sequencing

PFC: Prefrontal Cortex

Pol II: RNA Polymerase II

TF: Transcription Factor

TFBS: Transcription Factor Binding Site

s-DMR: Species-specific Differentially Methylated Region

## References


[1] S. Paabo, The human condition-a molecular approach, Cell, 157 (2014) 216-226.
[2] J. Rogers, R.A. Gibbs, Comparative primate genomics: emerging patterns of genome content and dynamics, Nat Rev Genet, 15 (2014) 347-359.
[3] K. Prufer, F. Racimo, N. Patterson, F. Jay, S. Sankararaman, S. Sawyer, A. Heinze, G. Renaud, P.H. Sudmant, C. de Filippo, H. Li, S. Mallick, M. Dannemann, Q. Fu, M. Kircher, M. Kuhlwilm, M. Lachmann, M. Meyer, M. Ongyerth, M. Siebauer, C. Theunert, A. Tandon, P. Moorjani, J. Pickrell, J.C. Mullikin, S.H. Vohr, R.E. Green, I. Hellmann, P.L.F. Johnson, H. Blanche, H. Cann, J.O. Kitzman, J. Shendure, E.E. Eichler, E.S. Lein, T.E. Bakken, L.V. Golovanova, V.B. Doronichev, M.V. Shunkov, A.P. Derevianko, B. Viola, M. Slatkin, D. Reich, J. Kelso, S. Paabo, The complete genome sequence of a Neanderthal from the Altai Mountains, Nature, 505 (2014) 43-49.
[4] M. Meyer, M. Kircher, M.T. Gansauge, H. Li, F. Racimo, S. Mallick, J.G. Schraiber, F. Jay, K. Prufer, C. de Filippo, P.H. Sudmant, C. Alkan, Q. Fu, R. Do, N. Rohland, A. Tandon, M. Siebauer, R.E. Green, K. Bryc, A.W. Briggs, U. Stenzel, J. Dabney, J. Shendure, J. Kitzman, M.F. Hammer, M.V. Shunkov, A.P. Derevianko, N. Patterson, A.M. Andres, E.E. Eichler, M. Slatkin, D. Reich, J. Kelso, S. Paabo, A high-coverage genome sequence from an archaic Denisovan individual, Science, 338 (2012) 222-226.
[5] S. Sankararaman, S. Mallick, M. Dannemann, K. Prufer, J. Kelso, S. Paabo, N. Patterson, D. Reich, The genomic landscape of Neanderthal ancestry in present-day humans, Nature, 507 (2014) 354-357.
[6] F. Racimo, S. Sankararaman, R. Nielsen, E. Huerta-Sanchez, Evidence for archaic adaptive introgression in humans, Nat Rev Genet, 16 (2015) 359-371.
[7] S. Paabo, The diverse origins of the human gene pool, Nat Rev Genet, 16 (2015) 313-314.
[8] E. Huerta-Sanchez, X. Jin, Asan, Z. Bianba, B.M. Peter, N. Vinckenbosch, Y. Liang, X. Yi, M. He, M. Somel, P. Ni, B. Wang, X. Ou, Huasang, J. Luosang, Z.X.P. Cuo, K. Li, G. Gao, Y. Yin, W. Wang, X. Zhang, X. Xu, H. Yang, Y. Li, J. Wang, J. Wang, R. Nielsen, Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA, Nature, 512 (2014) 194-197.
[9] D. Gokhman, E. Lavi, K. Prüfer, M.F. Fraga, J.A. Riancho, J. Kelso, S. Pääbo, E. Meshorer, L. Carmel, Reconstructing the DNA Methylation Maps of the Neandertal and the Denisovan, Science, (2014).
[10] V.W. Zhou, A. Goren, B.E. Bernstein, Charting histone modifications and the functional organization of mammalian genomes, Nature reviews. Genetics, 12 (2011) 7-18.
[11] Maria J.E. Koster, B. Snel, H.T.M. Timmers, Genesis of Chromatin and Transcription Dynamics in the Origin of Species, Cell, 161 (2015) 724-736.
[12] J. Shin, G.-l. Ming, H. Song, Decoding neural transcriptomes and epigenomes via high-throughput sequencing, Nat Neurosci, 17 (2014) 1463-1475.
[13] F. Crick, Memory and molecular turnover, Nature, 312 (1984) 101.
[14] D. Gokhman, E. Meshorer, L. Carmel, Epigenetics: It's Getting Old. Past Meets Future in Paleoepigenetics, Trends Ecol Evol, (2016).
[15] Daniel H. Geschwind, P. Rakic, Cortical Evolution: Judge the Brain by Its Cover, Neuron, 80 633-647.

[16] M. O'Bleness, V.B. Searles, A. Varki, P. Gagneux, J.M. Sikela, Evolution of genetic and genomic features unique to the human lineage, Nat Rev Genet, 13 (2012) 853-866.
[17] R. Margueron, D. Reinberg, Chromatin structure and the inheritance of epigenetic information, Nat Rev Genet, 11 (2010) 285-296.
[18] P.N.C.B. Audergon, S. Catania, A. Kagansky, P. Tong, M. Shukla, A.L. Pidoux, R.C. Allshire, Restricted epigenetic inheritance of H3K9 methylation, Science, 348 (2015) 132-135.
[19] K. Ragunathan, G. Jih, D. Moazed, Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment, Science, 348 (2015) 1258699.
[20] M.P. Creyghton, A.W. Cheng, G.G. Welstead, T. Kooistra, B.W. Carey, E.J. Steine, J. Hanna, M.A. Lodato, G.M. Frampton, P.A. Sharp, L.A. Boyer, R.A. Young, R. Jaenisch, Histone H3K27ac separates active from poised enhancers and predicts developmental state, Proc Natl Acad Sci U S A, 107 (2010) 21931-21936.
[21] N. Saksouk, Teresa K. Barth, C. Ziegler-Birling, N. Olova, A. Nowak, E. Rey, J. Mateos-Langerak, S. Urbach, W. Reik, M.-E. Torres-Padilla, A. Imhof, J. Déjardin, Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation, Molecular Cell, 56 580-594.
[22] M.M. Hoffman, J. Ernst, S.P. Wilder, A. Kundaje, R.S. Harris, M. Libbrecht, B. Giardine, P.M. Ellenbogen, J.A. Bilmes, E. Birney, R.C. Hardison, I. Dunham, M. Kellis, W.S. Noble, Integrative annotation of chromatin elements from ENCODE data, Nucleic Acids Res, 41 (2013) 827-841.
[23] E. Calo, J. Wysocka, Modification of Enhancer Chromatin: What, How, and Why?, Molecular Cell, 49 (2013) 825-837.
[24] T.K. Kim, M. Hemberg, J.M. Gray, A.M. Costa, D.M. Bear, J. Wu, D.A. Harmin, M. Laptewicz, K. Barbara-Haley, S. Kuersten, E. Markenscoff-Papadimitriou, D. Kuhl, H. Bito, P.F. Worley, G. Kreiman, M.E. Greenberg, Widespread transcription at neuronal activity-regulated enhancers, Nature, 465 (2010) 182-187.
[25] U. Galderisi, G. Peluso, Defects in Chromatin Structure and Diseases, in: M.A. Hayat (Ed.) Tumor Dormancy, Quiescence, and Senescence, Volume 2, Springer Netherlands, 2014, pp. 143-153.
[26] H.T. Bjornsson, The Mendelian disorders of the epigenetic machinery, Genome Res, 25 (2015) 1473-1481.
[27] R.J. Ryan, B.E. Bernstein, Molecular biology. Genetic events that shape the cancer epigenome, Science, 336 (2012) 1513-1514.
[28] S. De, M.M. Babu, A time-invariant principle of genome evolution, Proc Natl Acad Sci U S A, 107 (2010) 13004-13009.
[29] D. Schubeler, Function and information content of DNA methylation, Nature, 517 (2015) 321-326.
[30] F. Schlesinger, A.D. Smith, T.R. Gingeras, G.J. Hannon, E. Hodges, De novo DNA demethylation and noncoding transcription define active intergenic regulatory elements, Genome Res, 23 (2013) 1601-1614.
[31] N.M. Cohen, E. Kenigsberg, A. Tanay, Primate CpG islands are maintained by heterogeneous evolutionary regimes involving minimal selection, Cell, 145 (2011) 773-786.

[32] M.B. Stadler, R. Murr, L. Burger, R. Ivanek, F. Lienert, A. Scholer, C. Wirbelauer, E.J. Oakeley, D. Gaidatzis, V.K. Tiwari, D. Schubeler, DNA-binding factors shape the mouse methylome at distal regulatory regions, Nature, 480 (2011) 490-495.
[33] A. Bird, The dinucleotide CG as a genomic signalling module, Journal of molecular biology, 409 (2011) 47-53.
[34] F. Lienert, C. Wirbelauer, I. Som, A. Dean, F. Mohn, D. Schubeler, Identification of genetic elements that autonomously determine DNA methylation states, Nature genetics, 43 (2011) 1091-1097.
[35] T. Quante, A. Bird, Do short, frequent DNA sequence motifs mould the epigenome?, Nat Rev Mol Cell Biol, (2016).
[36] P.W. Hill, R. Amouroux, P. Hajkova, DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story, Genomics, 104 (2014) 324-333.
[37] T. Bienvenu, J. Chelly, Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized, Nat Rev Genet, 7 (2006) 415-426.
[38] M.J. Lyst, R. Ekiert, D.H. Ebert, C. Merusi, J. Nowak, J. Selfridge, J. Guy, N.R. Kastan, N.D. Robinson, F. de Lima Alves, J. Rappsilber, M.E. Greenberg, A. Bird, Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor, Nat Neurosci, 16 (2013) 898-902.
[39] C.M. McGraw, R.C. Samaco, H.Y. Zoghbi, Adult neural function requires MeCP2, Science, 333 (2011) 186.
[40] Z. Liu, X. Li, J.T. Zhang, Y.J. Cai, T.L. Cheng, C. Cheng, Y. Wang, C.C. Zhang, Y.H. Nie, Z.F. Chen, W.J. Bian, L. Zhang, J. Xiao, B. Lu, Y.F. Zhang, X.D. Zhang, X. Sang, J.J. Wu, X. Xu, Z.Q. Xiong, F. Zhang, X. Yu, N. Gong, W.H. Zhou, Q. Sun, Z. Qiu, Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2, Nature, 530 (2016) 98-102.
[41] M. Mellen, P. Ayata, S. Dewell, S. Kriaucionis, N. Heintz, MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System, Cell, 151 (2012) 1417-1430.
[42] L. Chen, K. Chen, L.A. Lavery, S.A. Baker, C.A. Shaw, W. Li, H.Y. Zoghbi, MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome, Proc Natl Acad Sci U S A, (2015).
[43] H.W. Gabel, B. Kinde, H. Stroud, C.S. Gilbert, D.A. Harmin, N.R. Kastan, M. Hemberg, D.H. Ebert, M.E. Greenberg, Disruption of DNA-methylation-dependent long gene repression in Rett syndrome, Nature, advance online publication (2015).
[44] B. Kinde, H.W. Gabel, C.S. Gilbert, E.C. Griffith, M.E. Greenberg, Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2, Proceedings of the National Academy of Sciences, (2015).
[45] R. Halder, M. Hennion, R.O. Vidal, O. Shomroni, R.-U. Rahman, A. Rajput, T.P. Centeno, F. van Bebber, V. Capece, J.C.G. Vizcaino, A.-L. Schuetz, S. Burkhardt, E. Benito, M.N. Sala, S.B. Javan, C. Haass, B. Schmid, A. Fischer, S. Bonn, DNA methylation changes in plasticity genes accompany the formation and maintenance of memory, Nat Neurosci, 19 (2016) 102-110.
[46] H. Yu, Y. Su, J. Shin, C. Zhong, J.U. Guo, Y.-L. Weng, F. Gao, D.H. Geschwind, G. Coppola, G.-l. Ming, H. Song, Tet3 regulates synaptic transmission and homeostatic

plasticity via DNA oxidation and repair, Nat Neurosci, advance online publication (2015).
[47] J.D. Perez, N.D. Rubinstein, D.E. Fernandez, S.W. Santoro, L.A. Needleman, O. Ho-Shing, J.J. Choi, M. Zirlinger, S.-K. Chen, J.S. Liu, C. Dulac, Quantitative and functional interrogation of parent-of-origin allelic expression biases in the brain, 2015.
[48] C.E. Cain, R. Blekhman, J.C. Marioni, Y. Gilad, Gene expression differences among primates are associated with changes in a histone epigenetic modification, Genetics, 187 (2011) 1225-1234.
[49] K. Aberg, A.N. Khachane, G. Rudolf, S. Nerella, D.A. Fugman, J.A. Tischfield, E.J. van den Oord, Methylome-wide comparison of human genomic DNA extracted from whole blood and from EBV-transformed lymphocyte cell lines, Eur J Hum Genet, 20 (2012) 953-955.
[50] X. Zhou, C.E. Cain, M. Myrthil, N. Lewellen, K. Michelini, E.R. Davenport, M. Stephens, J.K. Pritchard, Y. Gilad, Epigenetic modifications are associated with inter-species gene expression variation in primates, Genome Biol, 15 (2014) 547.
[51] A.A. Pai, Y. Gilad, Comparative studies of gene regulatory mechanisms, Curr Opin Genet Dev, 29 (2014) 68-74.
[52] H.P. Shulha, J.L. Crisci, D. Reshetov, J.S. Tushir, I. Cheung, R. Bharadwaj, H.-J. Chou, I.B. Houston, C.J. Peter, A.C. Mitchell, W.-D. Yao, R.H. Myers, J.-f. Chen, T.M. Preuss, E.I. Rogaev, J.D. Jensen, Z. Weng, S. Akbarian, Human-Specific Histone Methylation Signatures at Transcription Start Sites in Prefrontal Neurons, PLoS Biol, 10 (2012) e1001427.
[53] H.P. Shulha, I. Cheung, C. Whittle, J. Wang, D. Virgil, C.L. Lin, Y. Guo, A. Lessard, S. Akbarian, Z. Weng, Epigenetic signatures of autism: trimethylated H3K4 landscapes in prefrontal neurons, Arch Gen Psychiatry, 69 (2012) 314-324.
[54] J. Ernst, P. Kheradpour, T.S. Mikkelsen, N. Shoresh, L.D. Ward, C.B. Epstein, X. Zhang, L. Wang, R. Issner, M. Coyne, M. Ku, T. Durham, M. Kellis, B.E. Bernstein, Mapping and analysis of chromatin state dynamics in nine human cell types, Nature, 473 (2011) 43-49.
[55] T.K. Kim, R. Shiekhattar, Architectural and Functional Commonalities between Enhancers and Promoters, Cell, 162 (2015) 948-959.
[56] L.J. Core, A.L. Martins, C.G. Danko, C.T. Waters, A. Siepel, J.T. Lis, Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers, Nat Genet, 46 (2014) 1311-1320.
[57] C.R. Marshall, A. Noor, J.B. Vincent, A.C. Lionel, L. Feuk, J. Skaug, M. Shago, R. Moessner, D. Pinto, Y. Ren, B. Thiruvahindrapduram, A. Fiebig, S. Schreiber, J. Friedman, C.E. Ketelaars, Y.J. Vos, C. Ficicioglu, S. Kirkpatrick, R. Nicolson, L. Sloman, A. Summers, C.A. Gibbons, A. Teebi, D. Chitayat, R. Weksberg, A. Thompson, C. Vardy, V. Crosbie, S. Luscombe, R. Baatjes, L. Zwaigenbaum, W. Roberts, B. Fernandez, P. Szatmari, S.W. Scherer, Structural variation of chromosomes in autism spectrum disorder, Am J Hum Genet, 82 (2008) 477-488.
[58] S. Djurovic, O. Gustafsson, M. Mattingsdal, L. Athanasiu, T. Bjella, M. Tesli, I. Agartz, S. Lorentzen, I. Melle, G. Morken, O.A. Andreassen, A genome-wide association study of bipolar disorder in Norwegian individuals, followed by replication in Icelandic sample, J Affect Disord, 126 (2010) 312-316.

[59] C.G. Bell, G.A. Wilson, L.M. Butcher, C. Roos, L. Walter, S. Beck, Human-specific CpG "beacons" identify loci associated with human-specific traits and disease, Epigenetics, 7 (2012) 1188-1199.
[60] C.G. Bell, G.A. Wilson, S. Beck, Human-specific CpG 'beacons' identify human-specific prefrontal cortex H3K4me3 chromatin peaks, Epigenomics, 6 (2014) 21-31.
[61] J.P. Thomson, P.J. Skene, J. Selfridge, T. Clouaire, J. Guy, S. Webb, A.R. Kerr, A. Deaton, R. Andrews, K.D. James, D.J. Turner, R. Illingworth, A. Bird, CpG islands influence chromatin structure via the CpG-binding protein Cfp1, Nature, 464 (2010) 1082-1086.
[62] G. Giannuzzi, E. Migliavacca, A. Reymond, Novel H3K4me3 marks are enriched at human- and chimpanzee-specific cytogenetic structures, Genome Res, (2014).
[63] A. Dincer, D.P. Gavin, K. Xu, B. Zhang, J.T. Dudley, E.E. Schadt, S. Akbarian, Deciphering H3K4me3 broad domains associated with gene-regulatory networks and conserved epigenomic landscapes in the human brain, Translational psychiatry, 5 (2015) e679.
[64] Bérénice A. Benayoun, Elizabeth A. Pollina, D. Ucar, S. Mahmoudi, K. Karra, Edith D. Wong, K. Devarajan, Aaron C. Daugherty, Anshul B. Kundaje, E. Mancini, Benjamin C. Hitz, R. Gupta, Thomas A. Rando, Julie C. Baker, Michael P. Snyder, J.M. Cherry, A. Brunet, H3K4me3 Breadth Is Linked to Cell Identity and Transcriptional Consistency, Cell, 158 (2014) 673-688.
[65] D. Villar, C. Berthelot, S. Aldridge, Tim F. Rayner, M. Lukk, M. Pignatelli, Thomas J. Park, R. Deaville, Jonathan T. Erichsen, Anna J. Jasinska, James M.A. Turner, Mads F. Bertelsen, Elizabeth P. Murchison, P. Flicek, Duncan T. Odom, Enhancer Evolution across 20 Mammalian Species, Cell, 160 (2015) 554-566.
[66] N.M. Young, G.P. Wagner, B. Hallgrimsson, Development and the evolvability of human limbs, Proc Natl Acad Sci U S A, 107 (2010) 3400-3405.
[67] J. Cotney, J. Leng, J. Yin, S.K. Reilly, L.E. DeMare, D. Emera, A.E. Ayoub, P. Rakic, J.P. Noonan, The evolution of lineage-specific regulatory activities in the human embryonic limb, Cell, 154 (2013) 185-196.
[68] T.A. Cormier, S.K. Prakash, D.B. Magner, H.Y. Zoghbi, I.B. Van den Veyver, Analysis of Mid1, Hccs, Arhgap6, and Msl3l1 in X-linked polydactyly (Xpl) and Patchy-fur (Paf) mutant mice, Mamm Genome, 12 (2001) 796-798.
[69] K.S. Pollard, S.R. Salama, N. Lambert, M.A. Lambot, S. Coppens, J.S. Pedersen, S. Katzman, B. King, C. Onodera, A. Siepel, A.D. Kern, C. Dehay, H. Igel, M. Ares, Jr., P. Vanderhaeghen, D. Haussler, An RNA gene expressed during cortical development evolved rapidly in humans, Nature, 443 (2006) 167-172.
[70] S. Prabhakar, J.P. Noonan, S. Paabo, E.M. Rubin, Accelerated evolution of conserved noncoding sequences in humans, Science, 314 (2006) 786.
[71] S. Prabhakar, A. Visel, J.A. Akiyama, M. Shoukry, K.D. Lewis, A. Holt, I. Plajzer-Frick, H. Morrison, D.R. Fitzpatrick, V. Afzal, L.A. Pennacchio, E.M. Rubin, J.P. Noonan, Human-specific gain of function in a developmental enhancer, Science, 321 (2008) 1346-1350.
[72] P. Rakic, Evolution of the neocortex: a perspective from developmental biology, Nat Rev Neurosci, 10 (2009) 724-735.
[73] S.K. Reilly, J. Yin, A.E. Ayoub, D. Emera, J. Leng, J. Cotney, R. Sarro, P. Rakic, J.P. Noonan, Evolutionary changes in promoter and enhancer activity during human corticogenesis, Science, 347 (2015) 1155-1159.

[74] S.L. Prescott, R. Srinivasan, M.C. Marchetto, I. Grishina, I. Narvaiza, L. Selleri, F.H. Gage, T. Swigut, J. Wysocka, Enhancer Divergence and cis-Regulatory Evolution in the Human and Chimp Neural Crest, Cell, 163 (2015) 68-83.
[75] M.C. Marchetto, I. Narvaiza, A.M. Denli, C. Benner, T.A. Lazzarini, J.L. Nathanson, A.C. Paquola, K.N. Desai, R.H. Herai, M.D. Weitzman, G.W. Yeo, A.R. Muotri, F.H. Gage, Differential L1 regulation in pluripotent stem cells of humans and apes, Nature, 503 (2013) 525-529.
[76] M.W. Vermunt, S.C. Tan, B. Castelijns, G. Geeven, P. Reinink, E. de Bruijn, I. Kondova, S. Persengiev, B. Netherlands Brain, R. Bontrop, E. Cuppen, W. de Laat, M.P. Creyghton, Epigenomic annotation of gene regulatory alterations during evolution of the primate brain, Nat Neurosci, advance online publication (2016).
[77] P.J. Wittkopp, G. Kalay, Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence, Nat Rev Genet, 13 (2012) 59-69.
[78] C.P. Bird, B.E. Stranger, M. Liu, D.J. Thomas, C.E. Ingle, C. Beazley, W. Miller, M.E. Hurles, E.T. Dermitzakis, Fast-evolving noncoding sequences in the human genome, Genome Biol, 8 (2007) R118.
[79] J.L. Boyd, S.L. Skove, J.P. Rouanet, L.J. Pilaz, T. Bepler, R. Gordan, G.A. Wray, D.L. Silver, Human-Chimpanzee Differences in a FZD8 Enhancer Alter Cell-Cycle Dynamics in the Developing Neocortex, Curr Biol, 25 (2015) 772-779.
[80] I. Hernando-Herraez, R. Garcia-Perez, A.J. Sharp, T. Marques-Bonet, DNA Methylation: Insights into Human Evolution, PLoS Genet, 11 (2015) e1005661.
[81] D. Macleod, J. Charlton, J. Mullins, A.P. Bird, Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island, Genes Dev, 8 (1994) 2282-2292.
[82] A. Feldmann, R. Ivanek, R. Murr, D. Gaidatzis, L. Burger, D. Schübeler, Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions, PLoS Genet, 9 (2013) e1003994.
[83] I. Hernando-Herraez, J. Prado-Martinez, P. Garg, M. Fernandez-Callejo, H. Heyn, C. Hvilsom, A. Navarro, M. Esteller, A.J. Sharp, T. Marques-Bonet, Dynamics of DNA methylation in recent human and great ape evolution, PLoS Genet, 9 (2013) e1003763.
[84] M.J. Ziller, H. Gu, F. Muller, J. Donaghey, L.T. Tsai, O. Kohlbacher, P.L. De Jager, E.D. Rosen, D.A. Bennett, B.E. Bernstein, A. Gnirke, A. Meissner, Charting a dynamic DNA methylation landscape of the human genome, Nature, 500 (2013) 477-481.
[85] G.A. Wilson, L.M. Butcher, H.R. Foster, A. Feber, C. Roos, L. Walter, G. Woszczek, S. Beck, C.G. Bell, Human-specific epigenetic variation in the immunological Leukotriene B4 Receptor (LTB4R/BLT1) implicated in common inflammatory diseases, Genome Med, 6 (2014) 19.
[86] R.A. Irizarry, C. Ladd-Acosta, B. Wen, Z. Wu, C. Montano, P. Onyango, H. Cui, K. Gabo, M. Rongione, M. Webster, H. Ji, J.B. Potash, S. Sabunciyan, A.P. Feinberg, The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores, Nat Genet, 41 (2009) 178-186.
[87] D.I. Martin, M. Singer, J. Dhahbi, G. Mao, L. Zhang, G.P. Schroth, L. Pachter, D. Boffelli, Phyloepigenomic comparison of great apes reveals a correlation between somatic and germline methylation states, Genome Res, 21 (2011) 2049-2057.
[88] I. Hernando-Herraez, H. Heyn, M. Fernandez-Callejo, E. Vidal, H. Fernandez-Bellon, J. Prado-Martinez, A.J. Sharp, M. Esteller, T. Marques-Bonet, The interplay

between DNA methylation and sequence divergence in recent human evolution, Nucleic Acids Res, (2015).
[89] D. Leung, I. Jung, N. Rajagopal, A. Schmitt, S. Selvaraj, A.Y. Lee, C.A. Yen, S. Lin, Y. Lin, Y. Qiu, W. Xie, F. Yue, M. Hariharan, P. Ray, S. Kuan, L. Edsall, H. Yang, N.C. Chi, M.Q. Zhang, J.R. Ecker, B. Ren, Integrative analysis of haplotype-resolved epigenomes across human tissues, Nature, 518 (2015) 350-354.
[90] R.F. Lowdon, H.S. Jang, T. Wang, Evolution of Epigenetic Regulation in Vertebrate Genomes, Trends in Genetics, 32 (2016) 269-283.
[91] J. Zeng, G. Konopka, B.G. Hunt, T.M. Preuss, D. Geschwind, S.V. Yi, Divergent Whole-Genome Methylation Maps of Human and Chimpanzee Brains Reveal Epigenetic Basis of Human Regulatory Evolution, Am J Hum Genet, (2012).
[92] A.Y. Panchin, V.J. Makeev, Y.A. Medvedeva, Preservation of methylated CpG dinucleotides in human CpG islands, Biol Direct, 11 (2016) 11.
[93] A. Molaro, E. Hodges, F. Fang, Q. Song, W.R. McCombie, G.J. Hannon, A.D. Smith, Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates, Cell, 146 (2011) 1029-1041.
[94] H. Wang, J. Xing, D. Grover, D.J. Hedges, K. Han, J.A. Walker, M.A. Batzer, SVA elements: a hominid-specific retroposon family, Journal of molecular biology, 354 (2005) 994-1007.
[95] A. Tanay, A.H. O'Donnell, M. Damelin, T.H. Bestor, Hyperconserved CpG domains underlie Polycomb-binding sites, Proc Natl Acad Sci U S A, 104 (2007) 5521-5526.
[96] H.K. Long, D. Sims, A. Heger, N.P. Blackledge, C. Kutter, M.L. Wright, F. Grutzner, D.T. Odom, R. Patient, C.P. Ponting, R.J. Klose, Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates, Elife, 2 (2013) e00348.
[97] H.K. Long, N.P. Blackledge, R.J. Klose, ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection, Biochem Soc Trans, 41 (2013) 727-740.
[98] M. Boulard, J.R. Edwards, T.H. Bestor, FBXL10 protects Polycomb-bound genes from hypermethylation, Nat Genet, 47 (2015) 479-485.
[99] T. Clouaire, J.I. de Las Heras, C. Merusi, I. Stancheva, Recruitment of MBD1 to target genes requires sequence-specific interaction of the MBD domain with methylated DNA, Nucleic Acids Res, 38 (2010) 4620-4634.
[100] D.A. Lim, Y.C. Huang, T. Swigut, A.L. Mirick, J.M. Garcia-Verdugo, J. Wysocka, P. Ernst, A. Alvarez-Buylla, Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells, Nature, 458 (2009) 529-533.
[101] S.Y. Kim, J.M. Levenson, S. Korsmeyer, J.D. Sweatt, A. Schumacher, Developmental regulation of Eed complex composition governs a switch in global histone modification in brain, J Biol Chem, 282 (2007) 9962-9972.
[102] M. Jakovcevski, H. Ruan, E.Y. Shen, A. Dincer, B. Javidfar, Q. Ma, C.J. Peter, I. Cheung, A.C. Mitchell, Y. Jiang, C.L. Lin, V. Pothula, A.F. Stewart, P. Ernst, W.D. Yao, S. Akbarian, Neuronal kmt2a/mll1 histone methyltransferase is essential for prefrontal synaptic plasticity and working memory, J Neurosci, 35 (2015) 5097-5108.
[103] M.R. Branco, G. Ficz, W. Reik, Uncovering the role of 5-hydroxymethylcytosine in the epigenome, Nat Rev Genet, 13 (2012) 7-13.
[104] M. Ko, J. An, H.S. Bandukwala, L. Chavez, T. Aijo, W.A. Pastor, M.F. Segal, H. Li, K.P. Koh, H. Lahdesmaki, P.G. Hogan, L. Aravind, A. Rao, Modulation of TET2

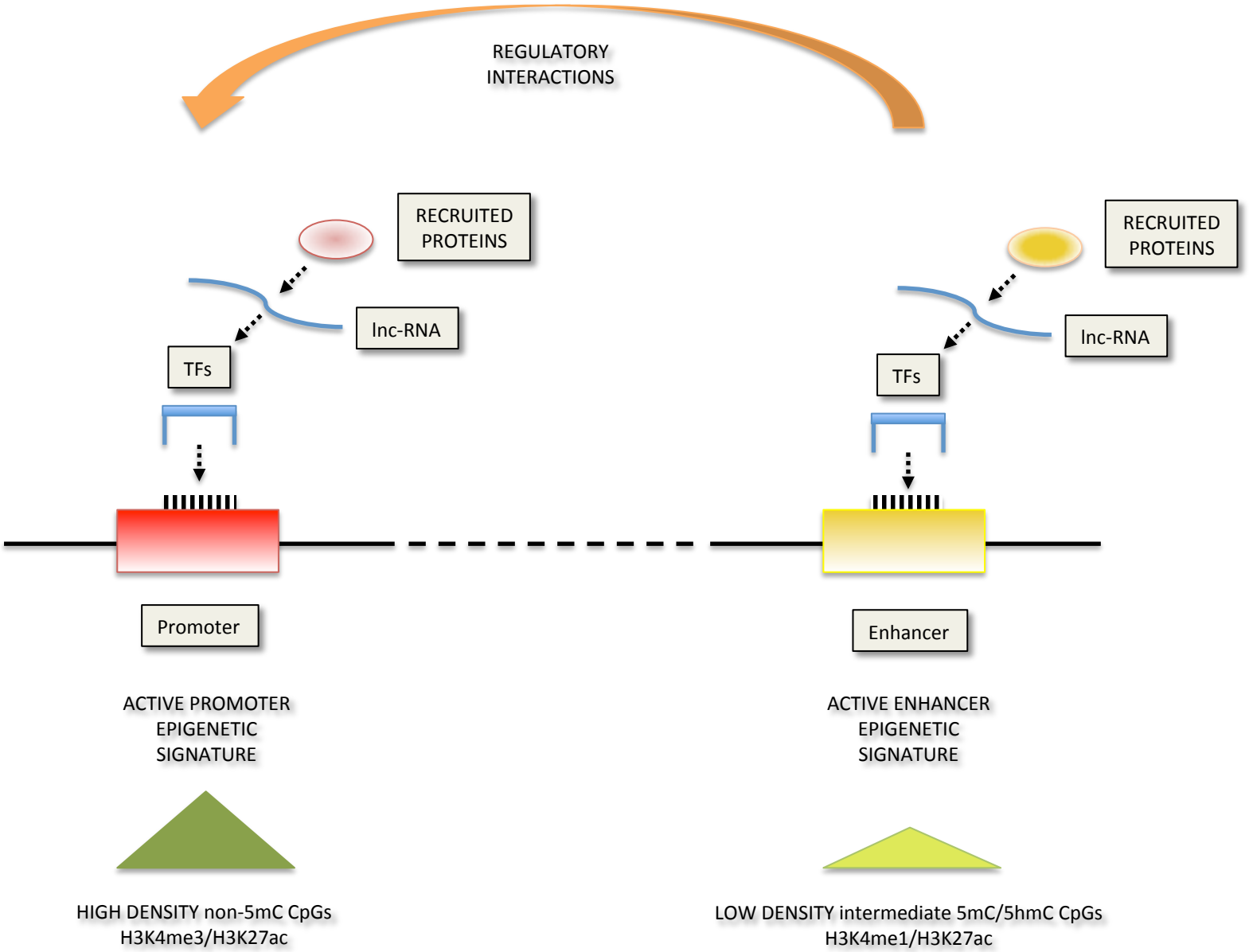
expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX, Nature, 497 (2013) 122-126.
[105] S. Iwase, E. Brookes, S. Agarwal, Aimee I. Badeaux, H. Ito, Christina N. Vallianatos, Giulio S. Tomassy, T. Kasza, G. Lin, A. Thompson, L. Gu, Kenneth Y. Kwan, C. Chen, Maureen A. Sartor, B. Egan, J. Xu, Y. Shi, A Mouse Model of X-linked Intellectual Disability Associated with Impaired Removal of Histone Methylation, Cell Rep, (2016).
[106] T. Kleefstra, H.G. Brunner, J. Amiel, A.R. Oudakker, W.M. Nillesen, A. Magee, D. Genevieve, V. Cormier-Daire, H. van Esch, J.P. Fryns, B.C. Hamel, E.A. Sistermans, B.B. de Vries, H. van Bokhoven, Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome, Am J Hum Genet, 79 (2006) 370-377.
[107] J. Cotney, R.A. Muhle, S.J. Sanders, L. Liu, A.J. Willsey, W. Niu, W. Liu, L. Klei, J. Lei, J. Yin, S.K. Reilly, A.T. Tebbenkamp, C. Bichsel, M. Pletikos, N. Sestan, K. Roeder, M.W. State, B. Devlin, J.P. Noonan, The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment, Nat Commun, 6 (2015).
[108] F. Baudat, J. Buard, C. Grey, A. Fledel-Alon, C. Ober, M. Przeworski, G. Coop, B. de Massy, PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice, Science, 327 (2010) 836-840.
[109] S. Myers, R. Bowden, A. Tumian, R.E. Bontrop, C. Freeman, T.S. MacFie, G. McVean, P. Donnelly, Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination, Science, 327 (2010) 876-879.
[110] C.L. Baker, S. Kajita, M. Walker, R.L. Saxl, N. Raghupathy, K. Choi, P.M. Petkov, K. Paigen, PRDM9 Drives Evolutionary Erosion of Hotspots in <italic>Mus musculus</italic> through Haplotype-Specific Initiation of Meiotic Recombination, PLoS Genet, 11 (2015) e1004916.
[111] J.J. Schwartz, D.J. Roach, J.H. Thomas, J. Shendure, Primate evolution of the recombination regulator PRDM9, Nat Commun, 5 (2014).
[112] B. Davies, E. Hatton, N. Altemose, J.G. Hussin, F. Pratto, G. Zhang, A.G. Hinch, D. Moralli, D. Biggs, R. Diaz, C. Preece, R. Li, E. Bitoun, K. Brick, C.M. Green, R.D. Camerini-Otero, S.R. Myers, P. Donnelly, Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice, Nature, 530 (2016) 171-176.
[113] B. Arbeithuber, A.J. Betancourt, T. Ebner, I. Tiemann-Boege, Crossovers are associated with mutation and biased gene conversion at recombination hotspots, Proceedings of the National Academy of Sciences, 112 (2015) 2109-2114.
[114] Y. Lesecque, S. Glemin, N. Lartillot, D. Mouchiroud, L. Duret, The red queen model of recombination hotspots evolution in the light of archaic and modern human genomes, PLoS Genet, 10 (2014) e1004790.
[115] K. Lindblad-Toh, M. Garber, O. Zuk, M.F. Lin, B.J. Parker, S. Washietl, P. Kheradpour, J. Ernst, G. Jordan, E. Mauceli, L.D. Ward, C.B. Lowe, A.K. Holloway, M. Clamp, S. Gnerre, J. Alfoldi, K. Beal, J. Chang, H. Clawson, J. Cuff, F. Di Palma, S. Fitzgerald, P. Flicek, M. Guttman, M.J. Hubisz, D.B. Jaffe, I. Jungreis, W.J. Kent, D. Kostka, M. Lara, A.L. Martins, T. Massingham, I. Moltke, B.J. Raney, M.D. Rasmussen, J. Robinson, A. Stark, A.J. Vilella, J. Wen, X. Xie, M.C. Zody, J. Baldwin, T. Bloom, C.W. Chin, D. Heiman, R. Nicol, C. Nusbaum, S. Young, J. Wilkinson, K.C. Worley, C.L. Kovar, D.M. Muzny, R.A. Gibbs, A. Cree, H.H. Dihn, G. Fowler, S. Jhangiani, V. Joshi, S.

Lee, L.R. Lewis, L.V. Nazareth, G. Okwuonu, J. Santibanez, W.C. Warren, E.R. Mardis, G.M. Weinstock, R.K. Wilson, K. Delehaunty, D. Dooling, C. Fronik, L. Fulton, B. Fulton, T. Graves, P. Minx, E. Sodergren, E. Birney, E.H. Margulies, J. Herrero, E.D. Green, D. Haussler, A. Siepel, N. Goldman, K.S. Pollard, J.S. Pedersen, E.S. Lander, M. Kellis, A high-resolution map of human evolutionary constraint using 29 mammals, Nature, 478 (2011) 476-482.
[116] A.L. Williams, G. Genovese, T. Dyer, N. Altemose, K. Truax, G. Jun, N. Patterson, S.R. Myers, J.E. Curran, R. Duggirala, J. Blangero, D. Reich, M. Przeworski, Non-crossover gene conversions show strong GC bias and unexpected clustering in humans, 2015.
[117] S. Karberg, Switching on Epigenetic Therapy, Cell, 139 (2009) 1029-1031.
[118] A. Perdomo-Sabogal, S. Kanton, M.B. Walter, K. Nowick, The role of gene regulatory factors in the evolutionary history of humans, Curr Opin Genet Dev, 29 (2014) 60-67.
[119] D. Hendrickson, D.R. Kelley, D. Tenen, B. Bernstein, J.L. Rinn, Widespread RNA binding by chromatin-associated proteins, Genome Biology, 17 (2016) 1-18.
[120] N. Leveille, C.A. Melo, K. Rooijers, A. Diaz-Lagares, S.A. Melo, G. Korkmaz, R. Lopes, F. Akbari Moqadam, A.R. Maia, P.J. Wijchers, G. Geeven, M.L. den Boer, R. Kalluri, W. de Laat, M. Esteller, R. Agami, Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA, Nat Commun, 6 (2015) 6520.
[121] P. Grote, B.G. Herrmann, Long noncoding RNAs in organogenesis: making the difference, Trends Genet, (2015).
[122] J. Wysocka, T. Swigut, T.A. Milne, Y. Dou, X. Zhang, A.L. Burlingame, R.G. Roeder, A.H. Brivanlou, C.D. Allis, WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development, Cell, 121 (2005) 859-872.
[123] Y.W. Yang, R.A. Flynn, Y. Chen, K. Qu, B. Wan, K.C. Wang, M. Lei, H.Y. Chang, Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency, 2014.
[124] J. Durruthy-Durruthy, V. Sebastiano, M. Wossidlo, D. Cepeda, J. Cui, E.J. Grow, J. Davila, M. Mall, W.H. Wong, J. Wysocka, K.F. Au, R.A. Reijo Pera, The primate-specific noncoding RNA HPAT5 regulates pluripotency during human preimplantation development and nuclear reprogramming, Nat Genet, 48 (2016) 44-52.
[125] M.C. King, A.C. Wilson, Evolution at two levels in humans and chimpanzees, Science, 188 (1975) 107-116.
[126] P. Hunter, The second coming of epigenetic drugs: A more strategic and broader research framework could boost the development of new drugs to modify epigenetic factors and gene expression, EMBO reports, (2015).
[127] A.J. Keung, J.K. Joung, A.S. Khalil, J.J. Collins, Chromatin regulation at the frontier of synthetic biology, Nat Rev Genet, 16 (2015) 159-171.
[128] S.J. Elsasser, K.M. Noh, N. Diaz, C.D. Allis, L.A. Banaszynski, Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells, Nature, (2015).
[129] D. Juan, J. Perner, E. Carrillo de Santa Pau, S. Marsili, D. Ochoa, H.-R. Chung, M. Vingron, D. Rico, A. Valencia, Epigenomic Co-localization and Co-evolution Reveal a

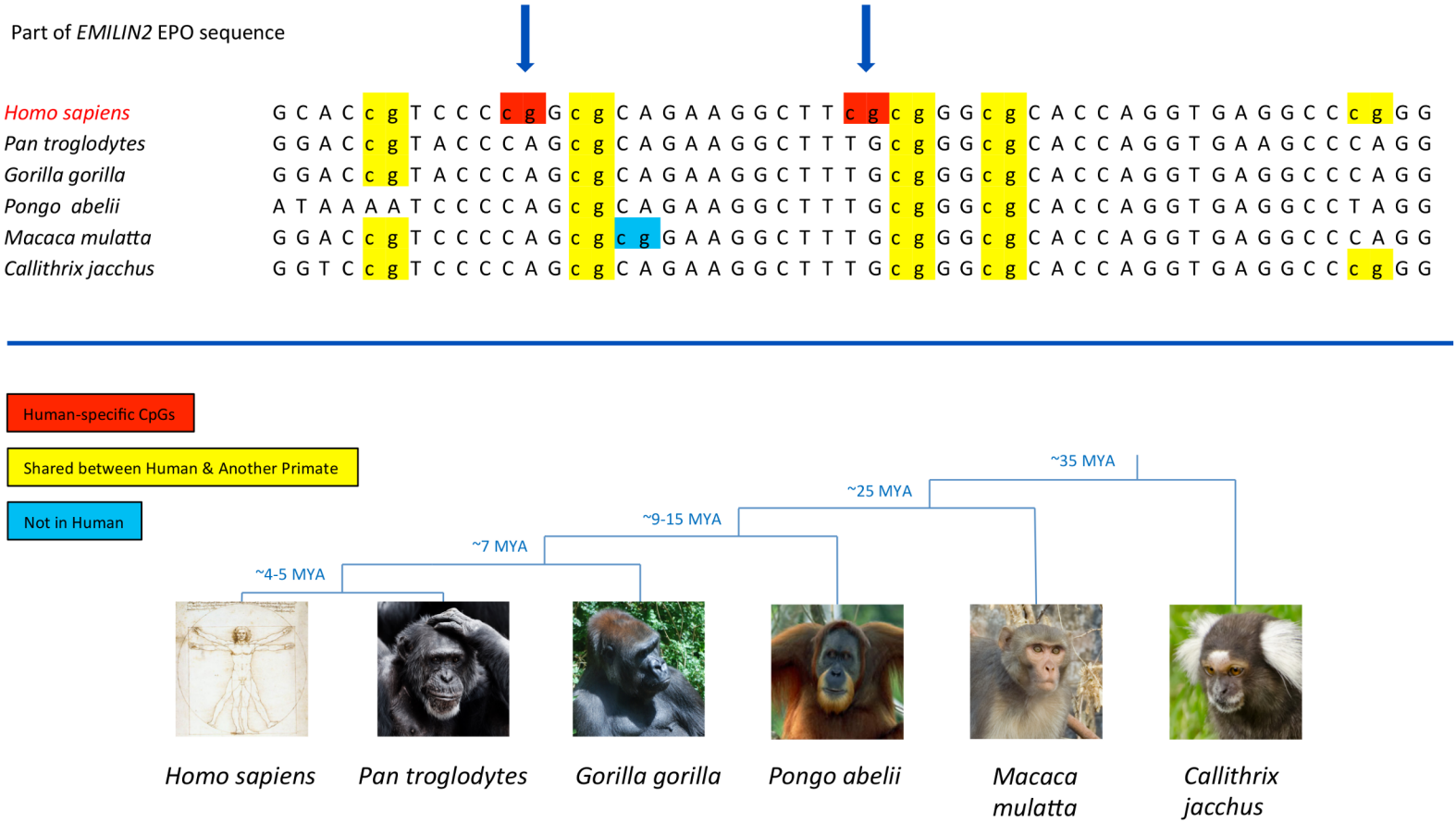
Key Role for 5hmC as a Communication Hub in the Chromatin Network of ESCs, Cell Rep, 14 (2016) 1246-1257.
[130] C.G. Spruijt, F. Gnerlich, A.H. Smits, T. Pfaffeneder, P.W. Jansen, C. Bauer, M. Munzel, M. Wagner, M. Muller, F. Khan, H.C. Eberl, A. Mensinga, A.B. Brinkman, K. Lephikov, U. Muller, J. Walter, R. Boelens, H. van Ingen, H. Leonhardt, T. Carell, M. Vermeulen, Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives, Cell, 152 (2013) 1146-1159.
[131] S.K. Sayeed, J. Zhao, B.K. Sathyanarayana, J.P. Golla, C. Vinson, C/EBPbeta (CEBPB) protein binding to the C/EBP|CRE DNA 8-mer TTGC|GTCA is inhibited by 5hmC and enhanced by 5mC, 5fC, and 5caC in the CG dinucleotide, Biochim Biophys Acta, (2015).
[132] J.P. Golla, J. Zhao, I.K. Mann, S.K. Sayeed, A. Mandal, R.B. Rose, C. Vinson, Carboxylation of cytosine (5caC) in the CG dinucleotide in the E-box motif (CGCAG|GTG) increases binding of the Tcf3|Ascl1 helix-loop-helix heterodimer 10-fold, Biochem Biophys Res Commun, 449 (2014) 248-255.
[133] M. Yu, G.C. Hon, K.E. Szulwach, C.X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J.H. Min, P. Jin, B. Ren, C. He, Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome, Cell, 149 (2012) 1368-1380.
[134] B. Delatte, F. Wang, L.V. Ngoc, E. Collignon, E. Bonvin, R. Deplus, E. Calonne, B. Hassabi, P. Putmans, S. Awe, C. Wetzel, J. Kreher, R. Soin, C. Creppe, P.A. Limbach, C. Gueydan, V. Kruys, A. Brehm, S. Minakhina, M. Defrance, R. Steward, F. Fuks, Transcriptome-wide distribution and function of RNA hydroxymethylcytosine, Science, 351 (2016) 282-285.
[135] S. Domcke, A.F. Bardet, P. Adrian Ginno, D. Hartl, L. Burger, D. Schubeler, Competition between DNA methylation and transcription factors determines binding of NRF1, Nature, 528 (2015) 575-579.
[136] E. de Wit, E.S. Vos, S.J. Holwerda, C. Valdes-Quezada, M.J. Verstegen, H. Teunissen, E. Splinter, P.J. Wijchers, P.H. Krijger, W. de Laat, CTCF Binding Polarity Determines Chromatin Looping, Mol Cell, 60 (2015) 676-684.
[137] J.S. Pedersen, E. Valen, A.M. Velazquez, B.J. Parker, M. Rasmussen, S. Lindgreen, B. Lilje, D.J. Tobin, T.K. Kelly, S. Vang, R. Andersson, P.A. Jones, C.A. Hoover, A. Tikhonov, E. Prokhortchouk, E.M. Rubin, A. Sandelin, M.T. Gilbert, A. Krogh, E. Willerslev, L. Orlando, Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome, Genome Res, 24 (2014) 454-466.
[138] S. Horvath, DNA methylation age of human tissues and cell types, Genome Biol, 14 (2013) R115.
[139] D. Hnisz, A.S. Weintraub, D.S. Day, A.-L. Valton, R.O. Bak, C.H. Li, J. Goldmann, B.R. Lajoie, Z.P. Fan, A.A. Sigova, J. Reddy, D. Borges-Rivera, T.I. Lee, R. Jaenisch, M.H. Porteus, J. Dekker, R.A. Young, Activation of proto-oncogenes by disruption of chromosome neighborhoods, Science, 351 (2016) 1454-1458.
[140] P.L. De Jager, G. Srivastava, K. Lunnon, J. Burgess, L.C. Schalkwyk, L. Yu, M.L. Eaton, B.T. Keenan, J. Ernst, C. McCabe, A. Tang, T. Raj, J. Replogle, W. Brodeur, S. Gabriel, H.S. Chai, C. Younkin, S.G. Younkin, F. Zou, M. Szyf, C.B. Epstein, J.A. Schneider, B.E. Bernstein, A. Meissner, N. Ertekin-Taner, L.B. Chibnik, M. Kellis, J. Mill, D.A. Bennett, Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci, Nat Neurosci, 17 (2014) 1156-1163.

[141] E. Gjoneska, A.R. Pfenning, H. Mathys, G. Quon, A. Kundaje, L.H. Tsai, M. Kellis, Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease, Nature, 518 (2015) 365-369.
[142] M. Thomas-Chollier, A. Hufton, M. Heinig, S. O'Keeffe, N.E. Masri, H.G. Roider, T. Manke, M. Vingron, Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs, Nature protocols, 6 (2011) 1860-1869.

Figure(s)

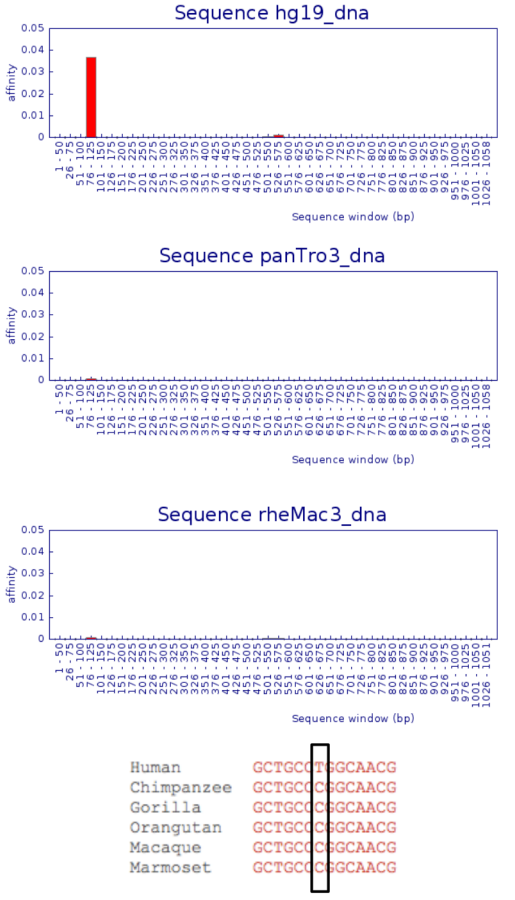

REGULATORY
INTERACTIONS
RECRUITED
PROTEINS
lnc-RNA
TFs
RECRUITED
PROTEINS
lnc-RNA
TFs
Promoter
Enhancer
ACTIVE PROMOTER
EPIGENETIC
SIGNATURE
ACTIVE ENHANCER
EPIGENETIC
SIGNATURE
HIGH DENSITY non-5mC CpGs
H3K4me3/H3K27ac
LOW DENSITY intermediate 5mC/5hmC CpGs
H3K4me1/H3K27ac

## Figure(s)


Part of EMILIN2 EPO sequence
Homo sapiens G C A C c g T C C C c g G c g C A G A A G G C T T c g c g G G c g C A C C A G G T G A G G C C c g G G
Pan troglodytes G G A C c g T A C C C A G c g C A G A A G G C T T T G c g G G c g C A C C A G G T G A A G C C C A G G
Gorilla gorilla G G A C c g T A C C C A G c g C A G A A G G C T T T G c g G G c g C A C C A G G T G A G G C C C A G G
Pongo abelii A T A A A A T C C C C A G c g C A G A A G G C T T T G c g G G c g C A C C A G G T G A G G C C T A G G
Macaca mulatta G G A C c g T C C C C A G c g c g G A A G G C T T T G c g G G c g C A C C A G G T G A G G C C C A G G
Callithrix jacchus G G T C c g T C C C C A G c g C A G A A G G C T T T G c g G G c g C A C C A G G T G A G G C C c g G G
Human-specific CpGs
Shared between Human & Another Primate
Not in Human
~35 MYA
~25 MYA
~9-15 MYA
~7 MYA
~4-5 MYA
Homo sapiens
Pan troglodytes
Gorilla gorilla
Pongo abelii
Macaca mulatta
Callithrix jacchus

**Figure(s)**

*RFX1* motif



***Research Highlights**

Highlights

- Epigenomic analysis reveals functional insights to cell-specific genomic activity
- Comparative primate epigenomics may give molecular understanding of human acquired traits
- Human-specific enhancers commonly redeploy existing regulatory elements
- Chromatin modifying enzymes may be important in evolution
- Transcription factors play a significant role in defining local epigenetic states