Aging Cell (2012) **11**, pp1055–1064 Doi: 10.1111/acel.12007

Stress-associated H3K4 methylation accumulates during postnatal development and aging of rhesus macaque brain

Yixing Han,[1,2,3†] Dali Han,[1,2,3†] Zheng Yan,[1] Jerome D. Boyd-Kirkup,[1] Christopher D. Green,[1] Philipp Khaitovich[1,4] and Jing-Dong J. Han[1]

[1]Chinese Academy of Sciences Key laboratory for Computational Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai, 200031, China
[2]Center for Molecular Systems Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Datun Road, Beijing, 100101, China
[3]Graduate University of Chinese Academy of Sciences, Yuquan Road, Beijing, 100049, China
[4]Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig, Germany

Summary

Epigenetic modifications are critical determinants of cellular and developmental states. Epigenetic changes, such as decreased H3K27me3 histone methylation on insulin/IGF1 genes, have been previously shown to modulate lifespan through gene expression regulation. However, global epigenetic changes during aging and their biological functions, if any, remain elusive. Here, we examined the histone modification H3K4 dimethylation (H3K4me2) in the prefrontal cortex of individual rhesus macaques at different ages by chromatin immunoprecipitation, followed by deep sequencing (ChIP-seq) at the whole genome level. Through integrative analysis of the ChIP-seq profiles with gene expression data, we found that H3K4me2 increased at promoters and enhancers globally during postnatal development and aging, and those that correspond to gene expression changes in *cis* are enriched for stress responses, such as the DNA damage response. This suggests that metabolic and environmental stresses experienced by an organism are associated with the progressive opening of chromatin. In support of this, we also observed increased expression of two H3K4 methyltransferases, SETD7 and DPY30, in aged macaque brain.

Key words: aging; ChIP-seq; histone methylation; prefrontal cortex; rhesus macaque; stress response.

Introduction

Aging is characterized by a progressive decline of numerous physiological functions and homeostasis, leading to an increased risk of 'aging-related' diseases such as Alzheimer's disease and, ultimately, death (Campisi, 2005; Kirkwood, 2005). Recent studies have revealed evolutionarily conserved aging regulatory pathways through genetic screens and characterized naturally occurring mutations and aging-related diseases in various model organisms. DNA microarray and RNA-seq studies in model organisms, and various human tissues, have demonstrated that gene expression varies remarkably during aging at the whole transcriptome level (Lu *et al.*, 2004; Fraser *et al.*, 2005). Integrative analyses, together with protein–protein interaction networks, indicate that these changes are modular and reflect concerted changes involving transcriptional regulators (Xue *et al.*, 2007). At the functional level, these changes reflect increased DNA repair activity (Lu *et al.*, 2004) and dysregulation of cellular state switches, for example, the switch between proliferation and differentiation phases (Xue *et al.*, 2007).

Epigenetic modifications play important roles in locking in transcriptional status and cellular states during the development of an organism (Mikkelsen *et al.*, 2007; Hirabayashi & Gotoh, 2010). They are an important interface for receiving and memorizing cues from the environment, allowing heritable and conditionally programmed gene expression from the genome (Berger, 2007; Li *et al.*, 2007; Yu *et al.*, 2008). In this way, epigenetic modifications could potentially mediate environmental influences on the rate of aging. Recent findings that *Caenorhabditis elegans* lifespan is regulated by histone methyltransferase, and demethylase enzymes capable of acting on histone H3 at lysine4 (H3K4me2/3, associated with transcriptional activation) and lysine27 (H3K27me2/3, associated with transcriptional repression), respectively, offer direct evidence for epigenetic regulation of aging and lifespan through specific alterations to chromatin architecture (Greer *et al.*, 2010; Jin *et al.*, 2011). In addition, deregulation of H4 lysine12 acetylation (H4K12ac) and lysine16 acetylation (H4K16ac), which are associated with gene activation, in the aging mouse brain has been linked to aging phenotypes including memory decline and DNA repair deficiency (Peleg *et al.*, 2010; Krishnan *et al.*, 2011). Chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) analysis of H3K4 trimethylation (H3K4me3) has also identified changes in the neurons of normal and autistic human brains, focused primarily at the early developmental stage (Cheung *et al.*, 2010). Nevertheless, whole genome-wide evidence for epigenetic regulation of chromatin plasticity during aging has, by and large, not been investigated.

Rhesus macaque (*Macaca mulatta*) is one of the most important primate model organisms because of its evolutionary closeness to human and has been widely used to study human diseases and aging (Colman *et al.*, 2009). With a genome 93% identical to *Homo sapiens* (Gibbs *et al.*, 2007), its maximal lifespan is much shorter than that of humans at only 25–30 years, making it a practical system to study aging in primates (Colman *et al.*, 2009).

Similar to the human brain, the rhesus macaque frontal cortex is associated with intellectual abilities (Margulies *et al.*, 2009) and is of great importance to the aging process. Here, we employed a ChIP-seq approach to examine the profiles of the active histone modification mark H3K4me2 in the prefrontal cortex (PFC) of rhesus macaques at different ages. H3K4me2 and H3K4me3 are both associated with active gene expressions, but unlike H3K4me3 that locates exclusively to transcription start sites (TSSs), H3K4me2 marks both TSSs and enhancers (He *et al.*,

Correspondence
Jing-Dong J. Han, Chinese Academy of Sciences Key laboratory for Computational Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China. Tel.: 8621 54920458; fax: 8621 54920451; e-mail: jdhan@picb.ac.cn

[†]These authors contributed equally to this work.

Accepted for publication *09 September 2012*

Aging Cell

2010). We found that the H3K4me2 modification levels at the promoter region and non-TSS H3K4me2 modification sites (or enhancer-like sites) increased during postnatal development and aging. Functional analysis, in conjunction with gene expression data, indicated that H3K4me2 changes, in particular those that conferred corresponding gene transcription changes, reflected progressively more active and transcriptionally accessible (or open) chromatin structures associated with metabolic and environmental stresses that induced DNA damage and inflammation. As a result, such changes may form an epigenetic memory of stress and damage experienced by the organism.

Results

H3K4me2 is positively correlated with gene expression within each sample

To probe the global changes in histone modification during postnatal development and aging, we conducted H3K4me2 ChIP-seq using rhesus macaque PFC samples. H3K4me2 marks not only transcription start sites (TSSs), but also enhancer regions, and as an activating mark, its level is correlated with downstream gene expression (He *et al.*, 2010). To facilitate the identification of functional H3K4me2 changes, we selected four rhesus macaque subjects that best represent the age-related gene expression changes among 12 macaque PFC samples of different postnatal development or aging stages, based on gene expression patterns previously profiled (Somel *et al.*, 2010, Figs 1A,B and S1A–C). These four samples are derived from 0.4-, 9-, 22-, and 26-year-old monkeys to represent four very different ontogenetic stages: infant, young, old, and very old, designated as 'A', 'B', 'C', and 'D' samples (Table S1 and Fig. 1C). We obtained 11 641 976, 13 140 748, 13 975 031, and 14 705 802 total reads for these four samples, respectively (Table S1). On average, 74% of the reads could be uniquely mapped to the rhesus macaque genome (rheMac2). The quality of the reads and the mapping rates were confirmed by the ShortRead package (Morgan *et al.*, 2009). The reads were also saturated for peak identification (i.e. significantly more reads than actually required for the highest peak confidence) based on the peaks identified by the SICER package (Zang *et al.*, 2009, Fig. S2). As expected, within each sample, H3K4me2 was significantly positively correlated with gene expression levels, consistent with it being an activating mark for gene expression (Fig. 1D–G).

More open and active chromatin at functional regions during postnatal development and aging

To examine the overall profile changes in H3K4me2 during rhesus macaque PFC aging, we first plotted the fraction of ChIP-seq short reads (shifted based on the average fragment length of ChIP) across different genomic regions after conventional RPM (reads per million) normalization (Heintzman *et al.*, 2009). Here, we adopted the routinely used operational definitions for different genomic regions: a promoter region encloses −2 to +2 kb surrounding the TSS; a 5′ or 3′ proximal region is −10 to −2 kb upstream of TSS or 0 to +10 kb downstream of transcription termination site (TTS); a 5′ or 3′ distal region is −100 to −10 kb upstream of TSS or +10 to +100 kb downstream of TTS; and those 100 kb away from TSS or TTS and not located in the above regions are defined as intergenic regions. As expected, H3K4me2 is concentrated at the gene promoters and intragenic regions, and given that promoter regions are much smaller than intergenic regions, this reflects a much higher read density, and therefore greater concentration of H3K4me2, at promoters (Fig. 2A). The asymmetric distribution with the inflection (representing the nucleosome depletion) at the TSS region is exactly as expected. Intriguingly, the fraction of reads at the promoter region also gradually increases with age (Fig. 2A). When focused on the TSS+/−2 kb region, we found that the normalized reads that mapped to either the positive or negative strands relative to the orientation of gene transcription showed age-related increases in both intensity and breath (Fig. 2B). The age-related increase in H3K4me2 density is most pronounced at two asymmetric peaks flanking the TSS (Figs 2C and S3A).

Since H3K4me2 marks both promoter and enhancer regions, we also examined the H3K4me2 levels on enhancer-like regions. Such regions were defined as the H3K4me2 islands remaining after discarding all H3K4me2 islands falling within TSS+/−10 kb. Similar to the age-related changes in promoters, H3K4me2 also increased with age at these enhancer-like regions (Fig. 2D).

To further verify our observation, we examined the published H3K4me3 profiles in five normal adult humans' PFC neurons (Shulha *et al.*, 2012). Consistent with the H3K4me2 profiles in macaque PFC, the H3K4me3 levels in the two older samples (age 38 and 55 years) were much higher ($P < 10e{-}20$ by paired Student's *t*-test across the whole transcriptome) than that in the three younger (age 20, 24, and 28 years) samples (Fig. S3B). However, it should be noted that unlike our macaque PFC samples, these samples were not selected to represent typical aging profiles, and therefore, it is possible that large individual variations could mask age-related changes.

We further examined the H3K4me2 profiles on individual genes using the UCSC genome browser (Fig. S4). Six known nontissue-specific aging-related genes (PPARGC1B, SIRT1, NFKB2, PRDX2, IRS1/2, and WNT11, see more information in Table S2) that had shown increased H3K4me2 levels in old macaque samples were selected for validation by ChIP-qPCR using rhesus macaque gluteal muscle tissue. ChIP-qPCR confirmed that H3K4me2 levels were significantly increased for PPARGC1B, SIRT1, and WNT11 ($P = 0.0007$, 0.04, and 0.01, respectively) and marginally increased for NFKB2, PRDX2, and IRS1/2 ($P = 0.25$, 0.1, and 0.14, respectively, Fig. S4) at these genes' promoters in aged macaque muscle (age 20, 21, and 22 years) when compared to younger individuals (age 8, 9, and 9 years; Fig. S4, qPCR primers are listed in Table S2).

H3K4me2 is associated with stress response genes during development and aging

What are the functional implications of such global opening of the chromatin? As H3K4me2/3 marks are associated with gene activation and a more open chromatin structure (Martin & Zhang, 2005; Barski *et al.*, 2007; Lieberman-Aiden *et al.*, 2009; Verzi *et al.*, 2010), we searched for genes that showed a consistent age-related increase/decrease in both H3K4me2 and gene expression, which we termed 'H3K4me2 *cis*-regulated genes' (H3K4me2 CRGs). We illustrated the definition of this group of genes and other similarly defined groups of genes using various Venn diagrams in Fig. 3A.

We first identified genes showing age-related expression changes using a previously published regression-based method (Somel *et al.*, 2009). Gene expression values for these genes were then clustered across the 12 rhesus macaque PFC samples using a *k*-means algorithm with $k = 4$ (Fig. S1A). Similar analyses were also performed on the normalized H3K4me2 counts at the promoters (TSS+/−2 kb). Consistent with the global profiles (Fig. 2A–C), many more promoters showed increased H3K4me2 with age, than those that show decreased H3K4me2 (Figs 3A, S1A and S5A, Table S3). This is very different from

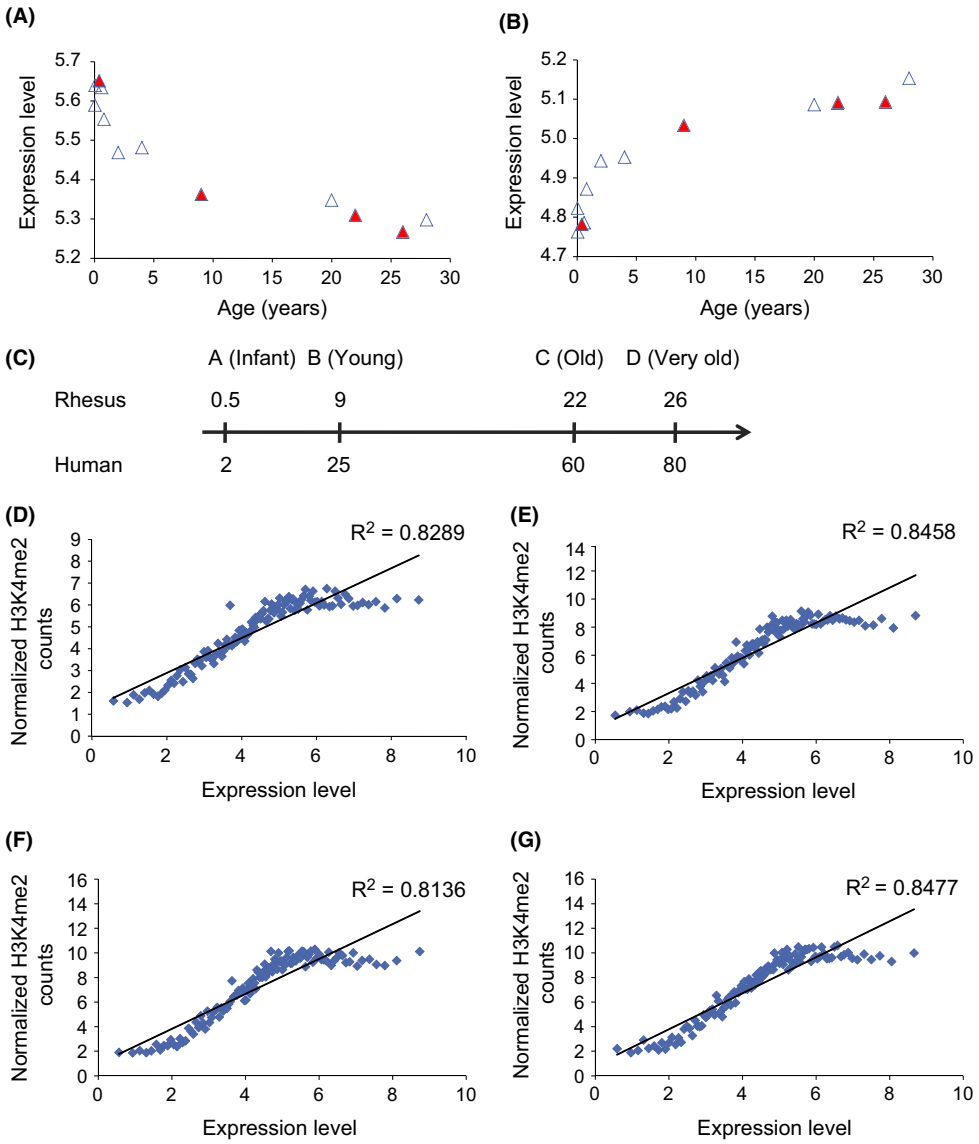

Fig. 1 Sample selection and genome-wide correlation of H3K4me2 to gene expression. (A) and (B) The four selected macaque samples fit the average age-related gene expression profile of the 12 macaque samples visualized by the age-related increased (A) or decreased (B) gene expression cluster (Fig. S1A). Each triangle represents a rhesus macaque individual's average gene expression level in the cluster. Red triangles mark the selected samples A, B, C, and D of ages 0.4, 9, 22, and 26 years, respectively. *x*-axis indicates the age of the macaques (years). (C) Schematic representation of the ages of the four selected rhesus macaques relative to the developmental and aging stages in rhesus macaque and human. (D–G) Linear regression of the normalized H3K4me2 ChIP-seq reads, within TSS+/−2 kb, with gene expression across all genes probed on the microarray in sample A, B, C, and D, respectively. Each diamond point represents 100 genes grouped by normalized gene expression values measured by microarray in each sample. Panels D–G are data from sample A–D, respectively. The average normalized H3K4me2 ChIP-seq signal within the TSS+/−2 kb of 100 genes in each group (indicated by the *y*-axis) is plotted against their average expression level (indicated by the *x*-axis). Pearson correlation coefficients (PCCs) were calculated between the group average expression levels and H3K4me2 levels across the 100 groups.

the age-related expression changes, which contained approximately equal numbers of genes that show expression increases or decreases with age (Figs 3A left panels and S1A). The genes that show an increase in expression overlap significantly (P = 9.3e10−4 by Fisher's exact test) with the genes that show an increase in H3K4me2 at the promoter, while those that show a decrease in expression do not significantly overlap with those that decrease in promoter H3K4me2 (P = 0.638). This indicates that not all expression changes are the direct result of *cis* changes in promoter H3K4me2 on the same gene, and this is particularly true for those genes that show expression decreases with age. We therefore focused on genes that show both age-related increase/decrease in gene expression and a corresponding increase/decrease in H3K4me2 at promoter CRGs (genes are listed in Table S4). The age-related promoter H3K4me2 increased CRGs (pICRGs) share significant overlap with genes that also show an age-related increase in H3K4me3 at promoters in human brain neurons (P = 0.012 by GSEA, Experimental procedures). Additionally, genes that show a significant age-related increase in H3K4me2 at macaque brain promoters also overlap significantly with human genes that show increased H3K4me3 at TSSs (P = 0.001). However, neither the age-related promoter decreased CRGs (pDCRGs) nor the age-related H3K4me2-down-regulated macaque genes show significant overlap with the H3K4me3-down-regulated human genes. These results suggest that promoter H3K4me2 up-regulation is more biologically relevant to the aging process than down-regulation.

Consistently among the pICRGs, in addition to chromatin modification functions, aging-related functions, such as oxidative stress and DNA damage, and repair-related GO terms and KEGG pathways are also significantly enriched (Fig. 3B left panel, Table S5). In contrast, the pDCRGs are not enriched for any biological functions. We further performed a literature co-citation analysis of the pICRGs with the terms 'aging', 'DNA repair', 'inflammation', 'metabolism', and 'stress' (Table S6). These terms were also significantly more co-cited with pICRGs than genes that merely show age-related expression increases or the average of all genes probed on the Affymetrix array (Fig. 3C, left panel). Interestingly, 'inflammation' and 'metabolism' terms were also more co-cited with genes showing age-related H3K4me2 increase and pDCRGs, respectively, compared to the background co-citation rate of all the genes on the microarray.

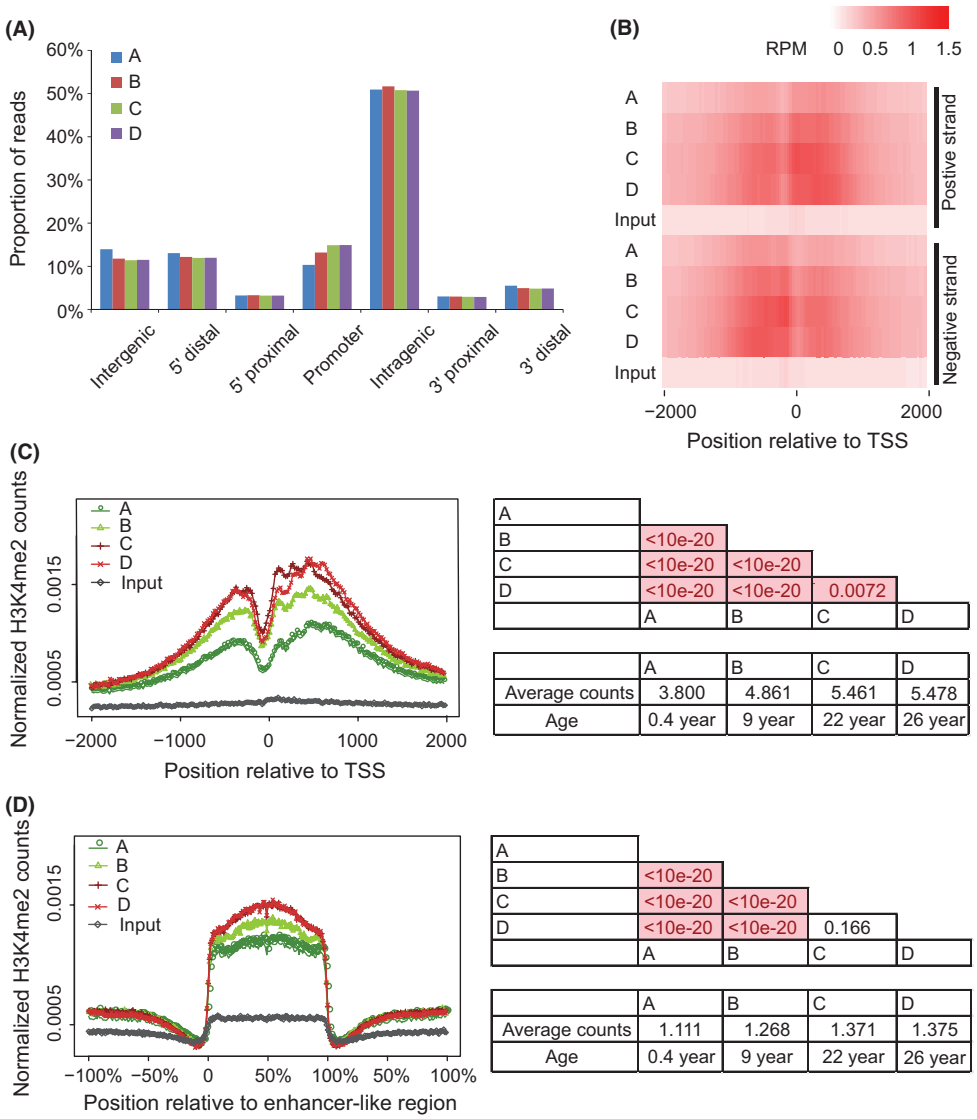

Fig. 2 H3K4me2 at promoter and enhancer-like regions increases with age. (A) Distribution (in percentage) of H3K4me2 ChIP-seq reads at different genomic regions. H3K4me2 modification is enriched at promoter and intragenic regions and increases at the promoter region with age. (B) H3K4me2 ChIP-seq read density on the positive (upper panel) or negative strand (lower panel) surrounding transcription start sites (TSSs) in A, B, C, D, and input samples. (C) H3K4me2 ChIP-seq read density surrounding TSSs in A, B, C, D, and input samples based on the combined reads on the positive and negative strands, shifted by 63 bp corresponding to half of the average fragment length. (D) Profiles of H3K4me2 density on all enhancer-like regions in A, B, C, D, and input samples.

These co-citation rates are robust across different fractions of data (sorted by significance *P*-value) and are generally (except for co-citation with the 'aging' term) higher among genes that display a trend for stronger age-related increase in both gene expression and H3K4me2 (Fig. S6A–D). The pICRGs show overlapping co-citation with many of these stress-related terms and are mostly enclosed by co-citation with the general 'stress' term (Figs 3D and S7). In addition, there is a significant enrichment of functional interactions (Experimental procedures) among these genes (Fig. 3D, empirical $P = 0.015$ based on 1000 permutations). Taken together, these data indicate that promoter-associated increases of H3K4me2 with age are most associated with stress responses.

We performed similar analyses on H3K4me2-marked enhancer-like regions and identified H3K4me2 CRGs based on H3K4me2 changes to the enhancer-like regions closest to the gene (eICRGs). Similar to the results from the promoter analysis, there were many more enhancer-like regions with an age-dependent increase in H3K4me2, than those with decreased H3K4me2 (Figs 3A right panels and S5B, Table S7). Although there is only a small overlap between pICRGs and eICRGs (Fig. 3A, Table S8), the eICRGs were also enriched for chromatin regulation functions (Fig. 3B right panel, Table S9). This suggests that chromatin modifiers themselves are a major target of age-related gene expression up-regulation by H3K4me2 at both promoters and enhancers, implicating extensive cross-talk, feedback, or feed forward controls among different chromatin modification events. As for pICRGs, the eICRGs were also more frequently co-cited with stress-related terms compared to genes with age-related decreases, or the average of all genes probed on the microarray (Figs 3C right panel for cocitation ≥ 1 and S7 for cocitation ≥ 2 or 3). eICRGs that are co-cited with 'aging', 'DNA repair', 'inflammation', and 'metabolism' also interact extensively among themselves (empirical $P = 0.01$ based on 1000 permutations), and both the genes and interactions are predominately covered by their co-citation with 'stress' (Fig. 3E).

Interestingly, of the specific stresses induced by UV, unfolded protein, heat shock and oxidation (co-citation terms listed on Table S6), oxidative stress-related genes are the most over-represented in pICRGs and eICRGs, followed by UV stress-related genes (Figs 4). This, together with the enriched DNA repair functions (Fig. 3D,E), points to the possibility that oxidative stress and UV-induced DNA damage might trigger H3K4 methylation.

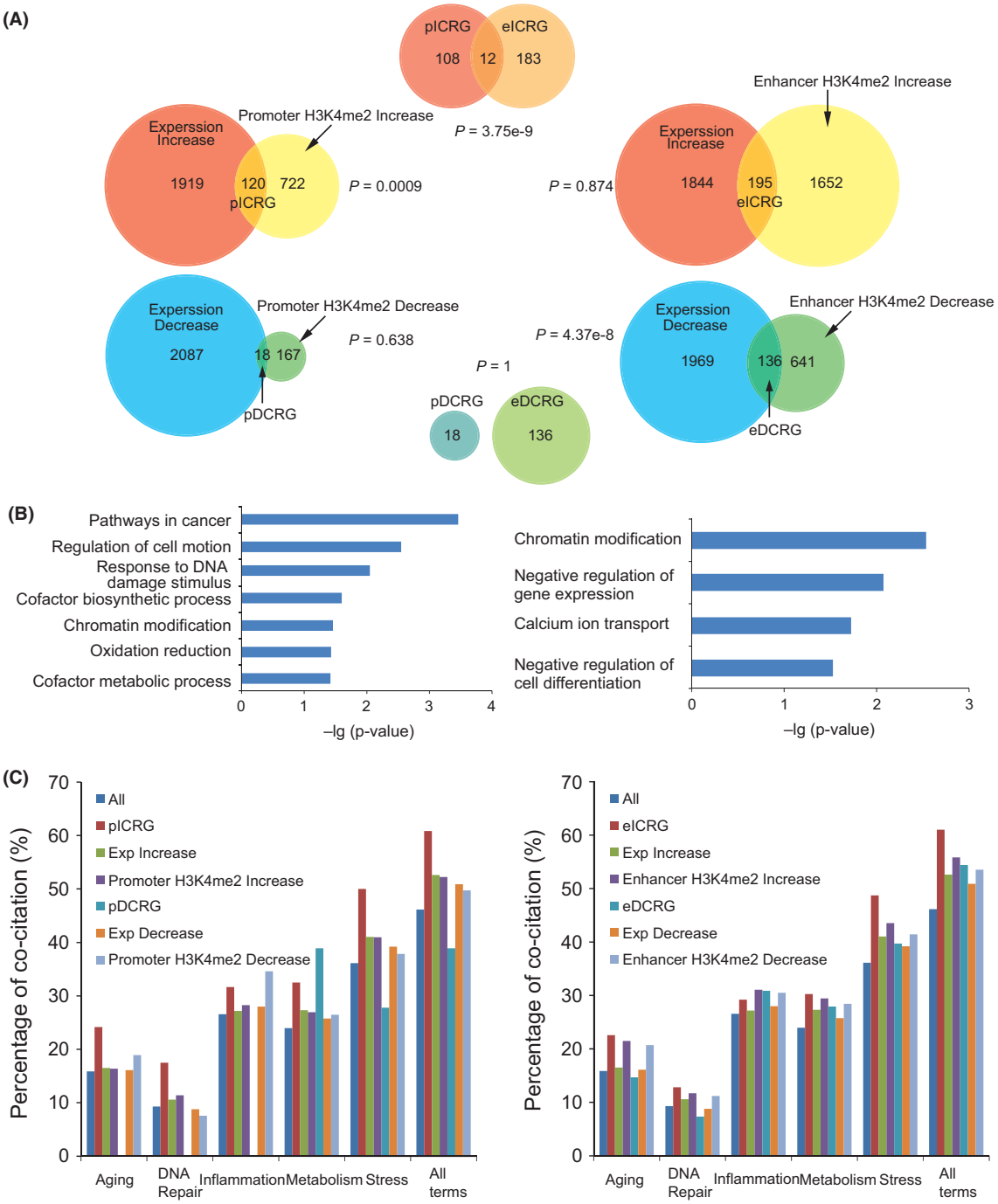

Fig. 3 Functional enrichment and literature co-citation analysis of the H3K4me2-regulated genes. (A) Overlaps for the number of genes that show age-related increase/decrease in gene expression or promoter/enhancer-like region H3K4me2 (pI/DCRGs and eI/DCRGs). (B) GO terms enriched among pICRGs (left) and eICRGs (right). The *y*-axis lists the GO terms, and the *x*-axis indicates the significance *P*-values of enrichment. (C) The percentages of pICRGs (left) and eICRGs (right) co-cited with the indicated terms in the PubMed database. (D) and (E) Venn diagram visualization of pICRG (D) and eICRG (E) co-cited with aging, DNA repair, inflammation, metabolism, stress, and their union, together with the functional interactions among the genes found in the KEGG, HPRD, IntNetDB, or STRING databases. Nodes represent genes, and edges represent functional interactions.

Therefore, we hypothesize that at the whole genome level, H3K4me2 marks at functional elements such as TSSs and enhancers induce the progressive opening of chromatin structure during development and aging, as a consequence of stress responses to various insults from the cellular environment, such as DNA damage and oxidative stresses.

Age-dependent increase in the transcription of two H3K4 methyltransferases

RNAi of H3K4 methyltransferases in *C. elegans* extends worm lifespan (Greer *et al.*, 2010), and therefore, H3K4 methyltransferase gene

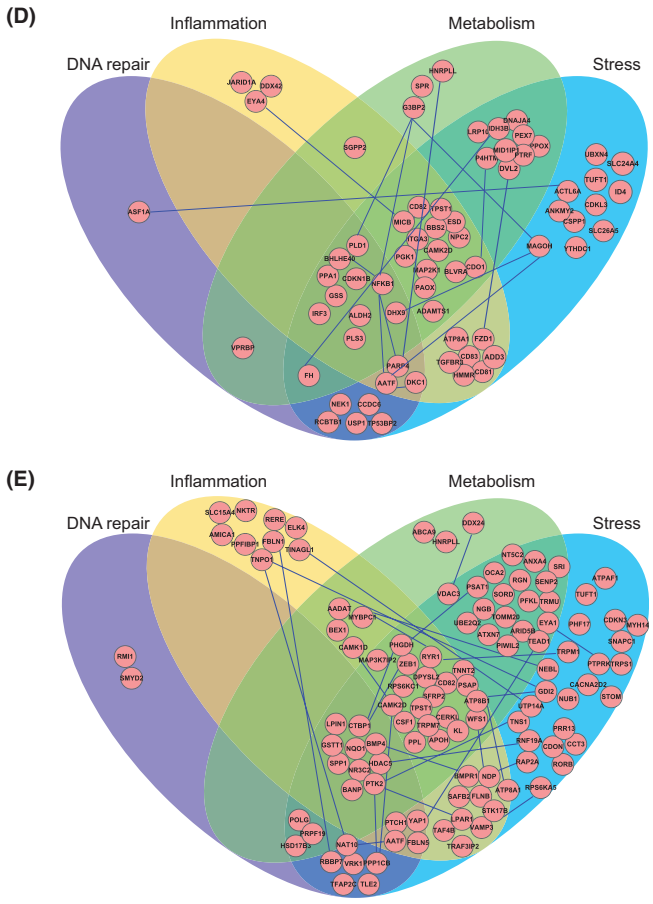

Fig. 3 (Continued)

expression and H3K4 methylation levels have been proposed to increase with age (Rando & Chang, 2012). Using the macaque brain microarray data, we identified two H3K4 methyltransferases, SETD7 and DPY30, that significantly increase with age in macaque brains ($P = 0.0002$ and 0.0004, respectively), and another H3K4 methyltransferase, MLL3, that marginally increases ($P = 0.0395$, Fig. 5). In addition, the H3K4me2 level in the promoter region of SETD7 also increases with age ($P = 0.00013$).

Taken together, these results support the hypothesis that the overall increase in H3K4me2 at gene promoters and functional elements can be elicited by various insults from the cellular environment.

Discussion

Here, we have investigated the influence of H3K4me2 changes in gene expression alterations during rhesus macaque brain postnatal development and aging through integrative analysis of ChIP-seq-generated genome-wide H3K4me2 profiles and microarray-generated transcriptome profiles. We observed a global increase in H3K4me2 during aging, implying a trend toward a more open and active chromatin structure at the TSS and enhancers throughout life history. Such increases, in particular those that contribute to *cis* regulation of gene expression, are highly enriched for stress response functions, leading us to hypothesize that the progressive opening of chromatin structure is a consequence of cellular stress. In support of this, we show here that H3K4me2 methyltransferases, SETD7 and DPY30, are significantly up-regulated in old versus young individuals.

A previous study in budding yeast demonstrated that H3K4me3 and its methyltransferase set1p (which also catalyzes the formation of H3K4me2) are increased upon DNA double-strand break damage (Faucher & Wellinger, 2010). Similarly, we also found a significant increase in the gene expression level of SETD7 and DPY30, but not MLL3, in response to the DNA damage agent 4-nitroquinoline 1-oxide (4NQO, Fig. S8, qPCR primers are listed in Table S10). Our observations implicate that responses to DNA damage might be potential triggers for the passive opening of chromatin after postnatal cellular and environmental challenges, and support the theory that environmental and developmental history can be 'remembered' through epigenetic modifications to the genome (Berger, 2007; Li *et al.*, 2007; Yu *et al.*, 2008). Consistent with this notion, high level of SETD7, which we found to increase in both promoter H3K4me2 and expression level with age (Fig. 5), has been reported to promote inflammation as a coactivator of NFKB (Li *et al.*, 2008).

Histone modifications play important roles in shaping the epigenetic landscape and chromatin structure, and it is hoped that future studies will address whether increased H3K4 methylation in the chromatin (Jiang & Pugh, 2009) is also associated with additional epigenetic modifications, such as histone acetylation, repressive histone modifications, and DNA methylation. It will also be interesting to determine how these combined modifications translate to changes in the high-dimensional structure of chromatin. Our finding that chromatin modifiers are major targets of age-related gene expression up-regulation by H3K4 methylation suggests a potentially extensive concerted change to the epigenome during postnatal development and aging. For example, among the genes up-regulated by H3K4me2, the anti-silencing factor ASF1A may induce further chromatin opening elsewhere in the genome; whereas KDM5A, a demethylase for H3K4 methylation, might be induced as a feedback mechanism to decrease global H3K4me2 after DNA repair (Table S5).

As previously observed for the transcriptome (Somel *et al.*, 2010), changes in the epigenome are largely continuous from the postnatal to aging stage, indicating that these changes are not unique to aging, but perhaps due to general stress experienced by cells and tissues throughout the lifespan of an organism. Even though the number of genes that show a decrease in H3K4me2 with age is much smaller than those that show an increase, H3K4me2-associated expression decrease may also be important in aging as the overlap (eDCRGs) between genes with decreased H3K4me2 at enhancer-like regions and the genes with decreased expression is highly significant and is also highly enriched for neural functions (Table S9). This suggests that in response to stress, or as a secondary effect, there is a decline in the normal functions of cells and tissues. Finally, our study further underlines the theory that reducing cellular and environmental stresses might protect the 'naïve' state of chromatin and hence contribute to slower or delayed aging.

Experimental procedures

Rhesus macaque samples

The rhesus macaque tissue samples were obtained from the Suzhou Experimental Animal Center (Suzhou, China). The origin of the animals is China. The animals lived in a family setting in a spacious (H × L × W: 3 × 5 × 3 m^3) and clean environment with exercise equipment and were adequately fed three times (rice, corn, and vegetable and nutritional supplements) per day. All of the animals were raised and sacrificed as described previously (Somel *et al.*, 2010).

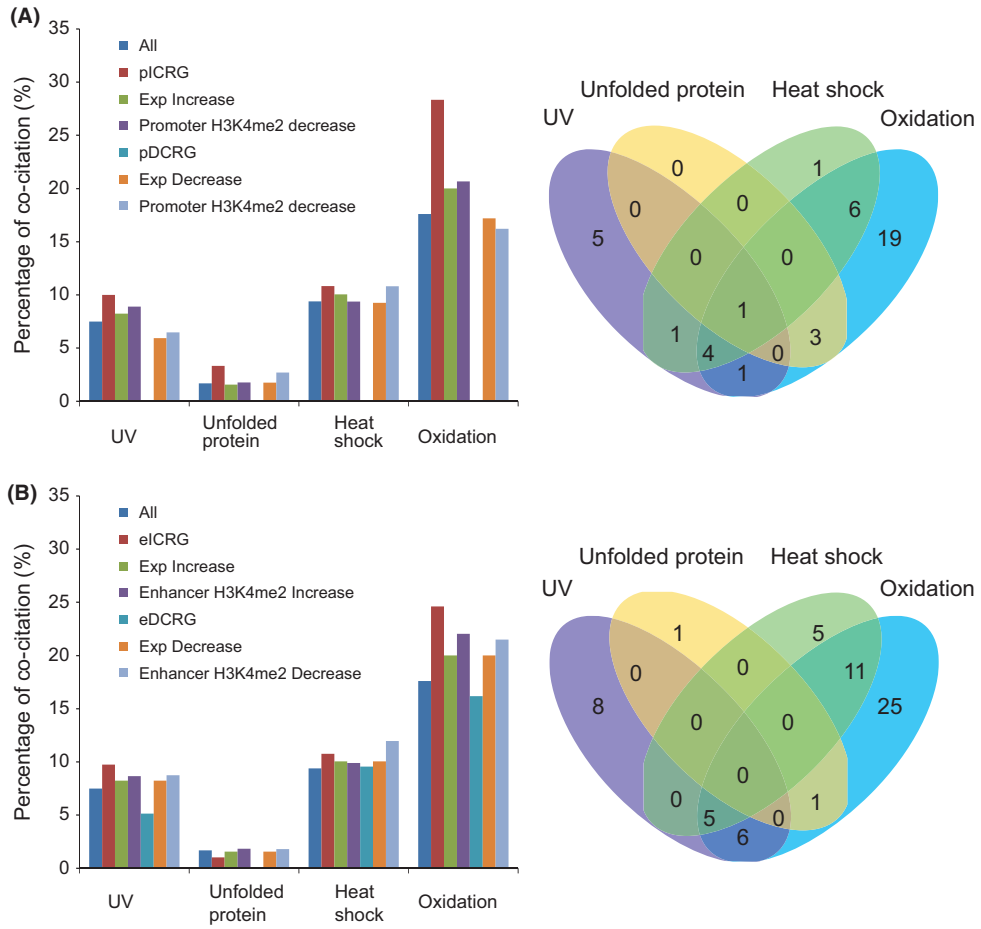

Fig. 4 Co-citation of pICRG or eICRG with specific stresses. Percentages of pICRG (A) or eICRG (B) co-cited with terms corresponding to the indicated stresses (Table S6) in at least one reference are shown in the left panels, while the number of these co-cited genes and their overlaps illustrated by Venn diagrams are shown in the right panels.

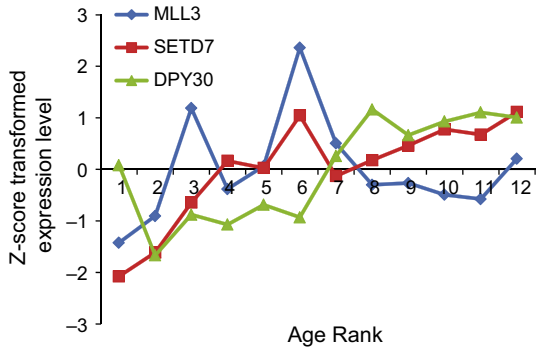

Fig. 5 Change in H3K4 methyltransferases expression levels with age and upon DNA damage. *Z*-score transformed expression level of H3K4 methyltransferases SETD7, DPY30, and MLL3 against the ages of 12 macaque (ranging from 0.04 to 28 years old as shown in Fig. S1A) prefrontal cortex samples.

ChIP-seq

Around 200 mg of rhesus macaque PFC tissues were used for each ChIP reaction using an antibody raised against H3K4me2 (07-030; Merck Millipore, Beijing, China). The tissues were ground into powder in liquid nitrogen, and the ChIP, and subsequent library construction before Illumina Genome Analyzer (GA II) (Illumina, San Diego, CA, USA) sequencing, were performed as described previously (Fei *et al.*, 2010; Liu *et al.*, 2011). Genomic DNA extracted from a pool of seven macaque individuals including the four used for ChIP-seq was also sequenced as a control 'input' sample, yielding 14 148 769 unique reads. Sequencing data generated in this study have been submitted to NCBI Gene Expression Omnibus under accession no. GSE41128.

Gene annotations

Computationally predicted genes (Non-Rhesus RefSeq Genes) were downloaded from UCSC genome browser on October 13, 2009. Only human homologue genes were used for analysis.

Gene expression and H3K4me2 level at gene promoters

Microarray data for rhesus macaque PFC were obtained from Somel *et al.*, 2010 and preprocessed using robust multi-array analysis (RMA, Irizarry *et al.*, 2003) in the Bioconductor package.

To examine the correlation of gene expression with H3K4me2, ChIP-seq reads surrounding promoters (defined as +/−2000 bp of TSS) were summed for each gene and normalized against all aligned reads in each sample and expressed as *per million aligned reads*. Genes with multiple promoters, or multiple copies, were discarded.

H3K4me2 peak detection

Summary islands were detected by SICER (Zang *et al.*, 2009) using the combined sequence reads from all four samples. Enhancer-like regions were obtained by excluding H3K4me2 summary islands overlapping with annotated or predicted (Liu *et al.*, 2011) TSSs +/−10 kb.

Age-related changes in H3K4me2 and gene expression

Polynomial regression models (Somel *et al.*, 2009) were used to detect age-related changes in expression or H3K4me2 levels. For each gene, the expression (or H3K4me2) level and log2-transformed ages were used to fit all possible 2 or 3 degree polynomial regression models and evaluated by the 'adjusted r^2'. The significance of the best regression model was estimated using *F*-test. Genes with *P*-value < 0.05 were defined as age-related genes. For H3K4me2 in promoter regions, the H3K4me2 level was normalized against total reads overlapping with the summary islands identified by SICER to increase signal-to-noise ratio, while for H3K4me2 in enhancer-like regions, the H3K4me2 level was normalized against the raw total reads to avoid artificial reduction of enhancer signals relative to the strong TSS signals.

H3K4me2 *cis*-regulated genes

cis-Regulated genes are the genes that gradually increased or decreased with age in both gene expression and H3K4me2. To identify promoter CRGs (pICRG/pDCRG), we separately clustered the age-related genes for expression or H3K4me2 values through *k*-means with $k = 4$ and took the overlapping genes between the expression and H3K4me2 clusters that showed consistent age-dependent changes. To identify enhancer CRGs (eICRG/eDCRG), we first identified enhancer-like region–associated genes using the 'basal plus extension association rule' (McLean *et al.*, 2010) before selecting the overlapping genes.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA, Subramanian *et al.*, 2005) was used to test the significance of the overlap between CRGs and the human genes marked by aged versus young, up- or down-regulated H3K4me3. To generate a preranked gene list for GSEA, RankProd (Hong *et al.*, 2006) was used to compare the differences between promoter H3K4me3 reads per million total reads (RPM) for the two older samples and the three younger samples, and then, the genes were ranked with RP/Sum. pICRG/pDCRG (or H3K4me2 increased/decreased genes) was used as a gene set for GSEA. The *P*-value was generated by 1000 permutations.

Functional enrichment analysis

Functional enrichment analyses were performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID, Huang da *et al.*, 2009a,b).

Literature co-citation

Literature co-citation was based on the PubMed database of October 2011. Only *Homo sapiens, Mus musculus*, and *Rattus norvegicus* publications were used for co-citation analysis. Terms used are listed in Table S6 (Supporting information). Genes and terms co-cited $\geq$ 1, 2, or 3 times were examined.

Functional interactions information

The functional interactions annotated in the KEGG (Kanehisa *et al.*, 2010), HPRD (Keshava Prasad *et al.*, 2009), IntNetDB (Xia *et al.*, 2006), and STRING (Jensen *et al.*, 2009) databases were combined as functional interactions.

Acknowledgments

We are indebted to Drs Teng Fei and Yeguang Chen for the use of equipments. This work was supported by grants from the China Natural National Science Foundation (Grant #30890033 and 91019019), Chinese Ministry of Science and Technology (Grant #2011CB504206), and Chinese Academy of Sciences (Grant #KSCX2-EW-R-02, KSCX2-EW-J-15, and XDA01010303) to J.D.J.H.

Author contributions

J.-D.J.H. and P.K. designed the study. Y.H. performed the experiments. D.H. performed computational analysis. Z.Y. and P.K. provided tissue samples. J.-D.J.H., Y.H., D.H., J.D.B.-K., and C.D.G. analyzed the results and wrote the manuscript.

References

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837.

Berger SL (2007) The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412.

Campisi J (2005) Senescent cells, tumor suppression, and organismal aging: good citizens, bad neighbors. *Cell* **120**, 513–522.

Cheung I, Shulha HP, Jiang Y, Matevossian A, Wang J, Weng Z, Akbarian S (2010) Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. *Proc. Natl Acad. Sci. USA* **107**, 8824–8829.

Colman RJ, Anderson RM, Johnson SC, Kastman EK, Kosmatka KJ, Beasley TM, Allison DB, Cruzen C, Simmons HA, Kemnitz JW, Weindruch R (2009) Caloric restriction delays disease onset and mortality in rhesus monkeys. *Science (New York, N.Y.)* **325**, 201–204.

Faucher D, Wellinger RJ (2010) Methylated H3K4, a transcription-associated histone modification, is involved in the DNA damage response pathway. *PLoS Genet.* **6**, e1001082.

Fei T, Xia K, Li Z, Zhou B, Zhu S, Chen H, Zhang J, Chen Z, Xiao H, Han JD, Chen YG (2010) Genome-wide mapping of SMAD target genes reveals the role of BMP signaling in embryonic stem cell fate determination. *Genome Res.* **20**, 36–44.

Fraser HB, Khaitovich P, Plotkin JB, Paabo S, Eisen MB (2005) Aging and gene expression in the primate brain. *PLoS Biol.* **3**, e274.

Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, Batzer MA, Bustamante CD, Eichler EE, Hahn MW, Hardison RC, Makova KD, Miller W, Milosavljevic A, Palermo RE, Siepel A, Sikela JM, Attaway T, Bell S, Bernard KE, Buhay CJ, Chandrabose MN, Dao M, Davis C, Delehaunty KD, Ding Y, Dinh HH, Dugan-Rocha S, Fulton LA, Gabisi RA, Garner TT, Godfrey J, Hawes AC, Hernandez J, Hines S, Holder M, Hume J, Jhangiani SN, Joshi V, Khan ZM, Kirkness EF, Cree A, Fowler RG, Lee S, Lewis LR, Li Z, Liu YS, Moore SM, Muzny D, Nazareth LV, Ngo DN, Okwuonu GO, Pai G, Parker D, Paul HA, Pfannkoch C, Pohl CS, Rogers YH, Ruiz SJ, Sabo A, Santibanez J, Schneider BW, Smith SM, Sodergren E, Svatek AF, Utterback TR, Vattathil S, Warren W, White CS, Chinwalla AT, Feng Y, Halpern AL, Hillier LW, Huang X, Minx P, Nelson JO, Pepin KH, Qin X, Sutton GG, Venter E, Walenz BP, Wallis JW, Worley KC, Yang SP, Jones SM, Marra MA, Rocchi M, Schein JE, Baertsch R, Clarke L, Csuros M, Glasscock J, Harris RA, Havlak P, Jackson AR, Jiang H, Liu Y, Messina DN, Shen Y, Song HX, Wylie T, Zhang L, Birney E, Han K, Konkel MK, Lee J, Smit AF, Ullmer B, Wang H, Xing J, Burhans R, Cheng Z, Karro JE, Ma J, Raney B, She X, Cox MJ, Demuth JP, Dumas LJ, Han SG, Hopkins J, Karimpour-Fard A, Kim YH, Pollack JR, Vinar T, Addo-Quaye C, Degenhardt J, Denby A, Hubisz MJ, Indap A, Kosiol C, Lahn BT, Lawson HA,

Marklein A, Nielsen R, Vallender EJ, Clark AG, Ferguson B, Hernandez RD, Hirani K, Kehrer-Sawatzki H, Kolb J, Patil S, Pu LL, Ren Y, Smith DG, Wheeler DA, Schenck I, Ball EV, Chen R, Cooper DN, Giardine B, Hsu F, Kent WJ, Lesk A, Nelson DL, O'Brien WE, Prufer K, Stenson PD, Wallace JC, Ke H, Liu XM, Wang P, Xiang AP, Yang F, Barber GP, Haussler D, Karolchik D, Kern AD, Kuhn RM, Smith KE, Zwieg AS (2007) Evolutionary and biomedical insights from the rhesus macaque genome. *Science (New York, N.Y.)* **316**, 222–234.

Greer EL, Maures TJ, Hauswirth AG, Green EM, Leeman DS, Maro GS, Han S, Banko MR, Gozani O, Brunet A (2010) Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in *C. elegans*. *Nature* **466**, 383–387.

He HH, Meyer CA, Shin H, Bailey ST, Wei G, Wang Q, Zhang Y, Xu K, Ni M, Lupien M, Mieczkowski P, Lieb JD, Zhao K, Brown M, Liu XS (2010) Nucleosome dynamics define transcriptional enhancers. *Nat. Genet.* **42**, 343–347.

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, Ching KA, Antosiewicz-Bourget JE, Liu H, Zhang X, Green RD, Lobanenkov VV, Stewart R, Thomson JA, Crawford GE, Kellis M, Ren B (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112.

Hirabayashi Y, Gotoh Y (2010) Epigenetic control of neural precursor cell fate during development. *Nat. Rev. Neurosci.* **11**, 377–388.

Hong F, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J (2006) RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics (Oxford, England)* **22**, 2825–2827.

Huang da W, Sherman BT, Lempicki RA (2009a) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13.

Huang da W, Sherman BT, Lempicki RA (2009b) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics (Oxford, England)* **4**, 249–264.

Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P, von Mering C (2009) STRING 8–a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* **37**, D412–D416.

Jiang C, Pugh BF (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.* **10**, 161–172.

Jin C, Li J, Green CD, Yu X, Tang X, Han D, Xian B, Wang D, Huang X, Cao X, Yan Z, Hou L, Liu J, Shukeir N, Khaitovich P, Chen CD, Zhang H, Jenuwein T, Han JD (2011) Histone demethylase UTX-1 regulates *C. elegans* life span by targeting the insulin/IGF-1 signaling pathway. *Cell Metab.* **14**, 161–172.

Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38**, D355–D360.

Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, Balakrishnan L, Marimuthu A, Banerjee S, Somanathan DS, Sebastian A, Rani S, Ray S, Harrys Kishore CJ, Kanth S, Ahmed M, Kashyap MK, Mohmood R, Ramachandra YL, Krishna V, Rahiman BA, Mohan S, Ranganathan P, Ramabadran S, Chaerkady R, Pandey A (2009) Human Protein Reference Database–2009 update. *Nucleic Acids Res.* **37**, D767–D772.

Kirkwood TB (2005) Understanding the odd science of aging. *Cell* **120**, 437–447.

Krishnan V, Chow MZ, Wang Z, Zhang L, Liu B, Liu X, Zhou Z (2011) Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice. *Proc. Natl Acad. Sci. USA* **108**, 12325–12330.

Li B, Carey M, Workman JL (2007) The role of chromatin during transcription. *Cell* **128**, 707–719.

Li Y, Reddy MA, Miao F, Shanmugam N, Yee JK, Hawkins D, Ren B, Natarajan R (2008) Role of the histone H3 lysine 4 methyltransferase, SET7/9, in the regulation of NF-kappaB-dependent inflammatory genes. Relevance to diabetes and inflammation. *J. Biol. Chem.* **283**, 26771–26781.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293.

Liu Y, Han D, Han Y, Yan Z, Xie B, Li J, Qiao N, Hu H, Khaitovich P, Gao Y, Han JD (2011) Ab initio identification of transcription start sites in the Rhesus macaque genome by histone modification and RNA-Seq. *Nucleic Acids Res.* **39**, 1408–1418.

Lu T, Pan Y, Kao SY, Li C, Kohane I, Chan J, Yankner BA (2004) Gene regulation and DNA damage in the ageing human brain. *Nature* **429**, 883–891.

Margulies DS, Vincent JL, Kelly C, Lohmann G, Uddin LQ, Biswal BB, Villringer A, Castellanos FX, Milham MP, Petrides M (2009) Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc. Natl Acad. Sci. USA* **106**, 20069–20074.

Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **6**, 838–849.

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560.

Morgan M, Anders S, Lawrence M, Aboyoun P, Pages H, Gentleman R (2009) ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics (Oxford, England)* **25**, 2607–2608.

Peleg S, Sananbenesi F, Zovoilis A, Burkhardt S, Bahari-Javan S, Agis-Balboa RC, Cota P, Wittnam JL, Gogol-Doering A, Opitz L, Salinas-Riester G, Dettenhofer M, Kang H, Farinelli L, Chen W, Fischer A (2010) Altered histone acetylation is associated with age-dependent memory impairment in mice. *Science* **328**, 753–756.

Rando TA, Chang HY (2012) Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock. *Cell* **148**, 46–57.

Shulha HP, Cheung I, Whittle C, Wang J, Virgil D, Lin CL, Guo Y, Lessard A, Akbarian S, Weng Z (2012) Epigenetic signatures of autism: trimethylated H3K4 landscapes in prefrontal neurons. *Arch. Gen. Psychiatry* **69**, 314–324.

Somel M, Franz H, Yan Z, Lorenc A, Guo S, Giger T, Kelso J, Nickel B, Dannemann M, Bahn S, Webster MJ, Weickert CS, Lachmann M, Paabo S, Khaitovich P (2009) Transcriptional neoteny in the human brain. *Proc. Natl Acad. Sci. USA* **106**, 5743–5748.

Somel M, Guo S, Fu N, Yan Z, Hu HY, Xu Y, Yuan Y, Ning Z, Hu Y, Menzel C, Hu H, Lachmann M, Zeng R, Chen W, Khaitovich P (2010) MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.* **20**, 1207–1218.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550.

Verzi MP, Shin H, He HH, Sulahian R, Meyer CA, Montgomery RK, Fleet JC, Brown M, Liu XS, Shivdasani RA (2010) Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Dev. Cell* **19**, 713–726.

Xia K, Xue H, Dong D, Zhu S, Wang J, Zhang Q, Hou L, Chen H, Tao R, Huang Z, Fu Z, Chen YG, Han JD (2006) Identification of the proliferation/differentiation switch in the cellular network of multicellular organisms. *PLoS Comput. Biol.* **2**, e145.

Xue H, Xian B, Dong D, Xia K, Zhu S, Zhang Z, Hou L, Zhang Q, Zhang Y, Han JD (2007) A modular network model of aging. *Mol. Syst. Biol.* **3**, 147.

Yu H, Zhu S, Zhou B, Xue H, Han JD (2008) Inferring causal relationships among different histone modifications and gene expression. *Genome Res.* **18**, 1314–1324.

Zang C, Schones DE, Zeng C, Cui K, Zhao K, Peng W (2009) A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952–1958.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1 Age-related gene expression changes and sample selection.

Fig. S2 Saturation curve of peak calling.

Fig. S3 The ChIP-Seq profiles around TSS region.

Fig. S4 Genome browser profiles of normalized macaque brain H3K4me2 ChIP-seq reads (left panels) and gluteal muscle ChIP-qPCR validation of increased H3K4me2 at PPARGC1B, SIRT1, NFKB2, PRDX2, IRS1/2 and WNT11 promoters (right panels) in the old (20, 21 and 22 years old) compared to the young (8, 9 and 9 years old) samples.

Fig. S5 Heatmap of the k-means clustering of genes that show age-related H3K4me2 changes at promoters (A) or enhancer-like regions (B).

Fig. S6 Robustness of the co-citation result over different gene sets or number of co-cited papers.

Fig. S7 Percentage of eICRGs co-cited with the indicated terms with at least 2 (A) or 3 (B) co-citation references.

Fig. S8 RT-qPCR analysis of the expression levels of H3K4 methyltransferases SETD7, DPY30 and MLL3 before or after DNA damage induced by 0.25 mg/mL 4NQO in HeLa cells.

Table S1. Summary of ChIP-seq samples.

Table S2. Genes whose promoter H3K4me2 levels increase in old versus young macaque brain samples determined by ChIP-seq were validated by ChIP-qPCR in young and old gluteal muscle samples.

Table S3. List of genes with promoter H3K4me2 levels that increase or decrease with age.

Table S4. List of promoter H3K4me2 *cis*-regulated genes (pCRGs) that show age-related increase (pICRG) or decrease (pDCRG).

Table S5. GO terms enriched among pI/DCRGs.

Table S6. List of terms used in co-citation analysis.

Table S7. List of enhancer-like regions with increased or decreased H3K4me2 levels with age.

Table S8. List of enhancer-like region H3K4me2 *cis*-regulated genes (eCRGs) that show age-related increase (eICRG) or decrease (eDCRG).

Table S9. GO terms enriched among eI/DCRGs determined by DAVID.

Table S10. qPCR primers for HeLa cells.

Data S1. Experimental procedures.