Aging Epigenetics: Accumulation of Errors or Realization of a Specific Program?

V. V. Ashapkin*, L. I. Kutueva, and B. F. Vanyushin

Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, 119991 Moscow, Russia; fax: + 7 (495) 939-3181; E-mail: ashapkin@genebee.msu.ru

Received July 6, 2015
Revision received July 13, 2015

Abstract—Aging in mammals is known to be accompanied by a progressive loss of methylated cytosines from DNA. This loss is tissue-specific to a certain extent and affects mainly repeated sequences, transposable elements, and intergenic genome parts. Age-dependent DNA hypomethylation is correlated with and perhaps partly caused by a diminished activity of DNA methyltransferases. Along with the global DNA demethylation during aging, hypermethylation of certain genes occurs. On the whole-genome scale, an age-dependent hypermethylation is typical for genes associated with promoter CG islands, whereas hypomethylation mostly affects CG-poor genes, besides the repeated sequences, transposable elements, and intergenic genome parts mentioned above. The methylation levels of certain CG sites display strict correlation to age and thus could be used as a molecular marker to predict biological age of cells, tissues, and organisms. Epigenetic cell reprogramming, such as induced pluripotent stem cell production, leads to complete resetting of their epigenetic age.

DOI: 10.1134/S0006297915110024

Key words: epigenetics, aging, DNA methylation, epigenetic age, epigenetic reprogramming

Participation of epigenetic mechanisms in cell and organismal aging has been a matter of investigations and discussions for many years. The first experimental data concerning age-dependent changes in DNA methylation were obtained in our lab nearly half a century ago. The 5-methylcytosine (5mC) content in DNA samples isolated from different organs of humpback salmon was found to be decreased with age [1]. Studies of 5mC content in mammals showed, first, that it is tissue specific [2], and second, that the DNA methylation levels are decreased upon aging in some rat tissues (brain, heart, spleen) and are essentially unchanged (liver, lung) or even slightly increased (kidney) in others [3]. Besides, a study of methylation of the variously repeated genome sequences showed that the repeated sequences are more heavily methylated as compare to the unique ones [4]. This difference is probably a consequence of the higher content of CG dinucleotides in the repeated genome compartment, these dinucleotides being the main, if not the sole methylation targets. The age-dependent hypomethylation mainly affects these heavily methylated repeated DNA sequences. Since this genome fraction contains transposable elements, their repression being regarded as one of the main defensive functions of DNA methylation, one can suggest that their demethylation upon aging should result in general genome destabilization. The DNA methylation levels in different organs of mice show a correlation with chronological age [5]. The rates of 5mC loss upon aging in DNA of two species of mice (C57 BL/GJ and white-footed) are inversely correlated to their maximal lifespans. Similar results were obtained for human bronchial epithelial cells isolated from differently aged donors. A progressive decrease in the DNA methylation levels was observed in cultured mouse, hamster, and human fibroblasts [6]. The maximal rate of the 5mC loss was observed in mouse cells, which could undergo 4-8 population doublings maximum before senescence, the medium rate – in hamster cells, that could undergo ~25 doublings, and last, the minimal rate – in human cells, capable to 40 doublings minimum without decrease in growth rate. When human fibroblasts are treated with 5-azacytidine (5azaC), a significant loss of 5mC and a decrease in maximal number of cell doublings occur [7, 8]. It could be suggested that the progressive 5mC loss in dividing cells is mainly caused by errors in the maintenance of DNA methylation, its

Abbreviations: aDMR, aging-associated DMR; DMR, differentially methylated region; hyper-aDMR, hypermethylated aDMR; hypo-aDMR, hypomethylated aDMR; 5mC, 5-methylcytosine; SNP, single-nucleotide polymorphism.

* To whom correspondence should be addressed.
fidelity being ~95% per cell generation [9, 10]. However, it was noted that no changes of DNA methylation levels are observed in dividing immortal cell lines [6]. Fidelity of the maintenance of DNA methylation could hardly be higher in these immortal cell lines compared with their normal counterpart cells. Besides, in mice similar losses of 5mC upon aging occur in tissues widely different in proliferative activity (liver, brain, small intestine mucosa) [5]. Thus, the age-related loss of 5mC could not be explained by maintenance methylation errors alone. On the other hand, an insufficient activity of the maintenance DNA methyltransferase DNMT1 can be one of the causes of DNA hypomethylation upon aging. It was shown, for example, that DNMT1 activity significantly diminishes with age [11, 12]. Even more surprising is that along with the global DNA hypomethylation, selective hypermethylation of multiple genes takes place upon aging [13]. Among these hypermethylated genes are often the tumor suppressor genes, known to be hypermethylated in tumor cells. For example, the promoter CG island methylation levels of the lysyl oxidase (LOX), cyclin-dependent kinase inhibitor 2A (CDKN2A, also known as p16, INK4a or p16INK4a), RUNX3, and tazarotene-induced gene 1 (TIG1) genes in normal stomach epithelial cells were found to display a strict positive correlation with the patient's age [14]. These genes were practically unmethylated before 50 years of age, whereas between 50 and 80 years their methylation levels progressively increased. This increase could well explain the increase in tumor occurrences in older people. Since the age-dependent increase in DNA methylation levels is nonlinear, in this case a specific process of some kind, not just accumulation of stochastic errors, evidently must cause it.

AGE-DEPENDENT CHANGES IN GENOME METHYLATION PATTERNS

An analysis of DNA methylation at 1413 CG sites associated with 773 genes in various tissue samples showed an evident dependence of the methylation patterns on tissue source of DNA and on age [15]. A rather distinct correlation between age and methylation was observed in solid tissues, the methylation levels of CG sites inside CG islands being increased with age, whereas the methylation levels of CG sites outside CG islands are decreased. Quite similar results were obtained in a study of methylation of the unique gene-associated CG islands in mouse small intestine cells at 3 to 35 month ages [16]. Of 3627 genes studied, 774 (21%) displayed hypermethylation with age and 466 (13%) displayed hypomethylation. Among genes devoid of CG islands in their 5′-untranslated region, 7% displayed hypermethylation and 11% hypomethylation. Among differentially methylated genes were Myod1, Cdkn2a, and Tmeff2. The hypermethylated gene group was enriched with genes involved in development and differentiation, whereas the hypomethylated gene group was not enriched for any specific functional category. Detailed study of 11 loci displaying clear age-dependent methylation in small intestine cells showed similar age-dependence for seven of them in hepatic cells, six in spleen cells, and three in lung cells. The methylation levels of 8821 genes were analyzed in a similar study on human colon tissue. Of these genes, 884 (10%) appeared to be hypermethylated with age, whereas only 1% appeared to be hypomethylated. Of 276 genes, hypermethylated in human colon and having a distinct homology on their promoter sequences with mouse genes, 116 (42%) were among the hypermethylated ones in mouse intestine. Of 99 genes hypomethylated in human colon, only 3 (3%) were also hypomethylated in mouse intestine. Thus, age-dependent gene methylation is only partially conserved between humans and mice, and the conserved genes are mainly among the hypermethylated ones. A study of several representative genes from both hypermethylated and hypomethylated groups showed that age-dependent hypermethylation is correlated with downregulation of gene transcription, and the age-dependent hypomethylation is correlated with upregulation. No changes in transcription were observed for invariably methylated genes.

The Illumina Infinium Human Methylation 27 BeadChip hybridization assay was used in the first genome-wide study of DNA methylation patterns in peripheral blood leukocytes from 93 healthy women of 49 to 75 years of age (including 31 twin pairs). The assay used provided a quantitative estimation of the methylation levels for 27,578 CG-dinucleotides localized within promoter regions of 14,495 genes; 20,006 of these CGs are parts of CG islands, whereas the remaining 7572 are located in promoters devoid of CG islands [17]. Two hundred-thirteen CG sites were found that become more methylated with age (hyper-aDMRs) and 147 CG sites that become less methylated (hypo-aDMRs). More than 95% of these aging-associated differentially methylated regions (aDMRs) are located within 500 bp of the transcriptional start sites. A similar investigation was performed on a smaller test group using purified fractions of CD14+ monocytes and CD4+ T-cells as DNA sources instead of the total leukocyte fraction. CD14+ monocytes are relatively short-lifespan (weeks) cells of the myeloid lineage, whereas CD4+ T-cells are long-lifespan (months-years) cells of the lymphoid lineage. A significant share (131 of 213, or >60%) of the hyper-aDMRs found earlier was also present in both purified fractions. The hypo-aDMRs set was significantly reproduced in T-cells, but not in monocytes. It could be supposed that the majority of the hyper-aDMRs represent cell-intrinsic aging-associated epigenetic perturbations, whereas the hypo-aDMRs reflect aging-associated changes in relative proportions of cell subtypes in whole blood. The conservative hyper-aDMRs
group is enriched for CG islands of genes with tissue-specific functions, including neural cell-related processes. Most of these genes are moderately active or inactive in the peripheral blood cells. At least in T-cells, the hyper-aDMR-associated promoters are poor in histone modification marks associated with transcriptional activity (H3K4me3, H2AZ) and RNA polymerase II, but enriched for histone marks associated with transcriptional repression (H3K9me3 and H3K27me3). These promoters significantly overlap those located in the chromatin bivalent domains in the embryonal stem cells, as well as those hypermethylated in various human cancers. The significant share of the hyper-aDMRs found in the peripheral blood cells was observed in a similar study on epithelial buccal cells. Thus, these hyper-aDMRs are multi-tissue hyper-aDMRs. Aberrant DNA methylation at bivalent chromatin domain promoters in cell culture is associated with decreased capacity to differentiate and increased ability to proliferate [18]. Similarly, in cancer cells such methylation leads to permanent silencing of genes required for differentiation, thereby stimulating cell proliferation. Hence, aging must also decrease cell differentiation potential and increase stem cell self-renewal. This may be one of the mechanisms that enhance cancer frequency at advanced ages.

In a similar study of DNA methylation in four brain divisions (frontal cortex, temporal cortex, pons, and cerebellum) in variously aged donors, somewhat different results were obtained [19]. In total, 1141 CG sites were found that showed age-dependent methylation. Of these sites, 589 showed age-dependence in one brain division, 167 – in two, 86 – in three, and only 10 sites – in all four brain divisions studied. Of all age-related CG sites found, 932 were located within CG islands, 129 were not within CG islands, and 80 were in regions that could not be unequivocally defined as CG islands or non-islands. All 10 CpG sites that showed significant correlation with age in four brain regions were located within CpG islands, and their methylation levels were increased with age. One of these sites, located in the MYOD1 gene, is among the loci showing age-dependent methylation in brain and pleura found earlier [15], whereas three others were earlier identified as differentially methylated with age in leukocytes [17]. A positive correlation of methylation levels with age was observed also for the majority (95.4%) of the remaining CG sites. Of the age-related CG sites that are parts of CG islands, 98% were hypermethylated. Frontal and temporal cortex divisions are most similar by the methylation patterns of age-related CG sites, pons being somewhat more divergent from both cortex divisions, but not as much as cerebellum. The positively age-correlated CG site group is most enriched for loci involved in transcription regulation and morphogenesis. As compared with other tissues, brain contains mainly sites hypermethylated with age, most of them located in the promoter-associated CG islands of genes coding for transcription factors. This part of age-dependent DNA methylation could be involved in an epigenetic program affecting genome expression in the aging brain cells. And vice versa, the age-dependent loss of total DNA methylation observed in aging brain is probably stochastic in nature, which explains its non-selectiveness in relation to different CG sites.

This view gained support from a methylation study in monozygotic (MZ) twin pairs during early life (from birth to 18 month age) [20]. It was found that intergenic regions and CG-poor promoters are most likely to undergo changes in methylation during early life, whereas methylation of CG islands and CG-rich promoters is most stable. Since environmental and lifestyle differences between MZ twins can be excluded at these ages, their discordance in DNA methylation must be a result of stochastic methylation errors. Methylation of CG islands and CG rich promoters seems to be most stringently maintained. Thus, age-dependent variations of their methylation are likely to be nonrandom.

The authors of another study using the same platform on more various organs (brain, blood, kidney, skeletal muscle) arrived at a similar conclusion [21]. In all cases studied, the sites located within CG islands predominated among those hypermethylated with age, whereas hypomethylation with age was more characteristic of CG sites located beyond CG islands, and their methylation levels were more variable between tissues. Hence, age-related variations in methylation common for different tissues are mainly observed in CG islands and are usually represented by increases in methylation levels with age. On the other hand, tissue-specific variations in methylation are more characteristic of CG sites located beyond CG islands and are usually represented by decreases in methylation levels with age. Gene loci functionally connected with transcription regulation and morphogenesis dominate among those hypermethylated with age in various tissues. The age-related loci common for all four organs studied above were represented by only 29 CG sites; tissue-specific age-related sites (not displaying correlation with age in other tissues) were observed in each organ, 28 loci in blood, 32 – in brain, 144 – in kidney, and 132 – in skeletal muscle. Among the common age-related loci, 90% are those hypermethylated with age and CG-island associated, whereas among the tissue-specific age-related sites only 30% are represented by such loci. The common age-related loci are frequently associated with chromatin bivalent areas containing both repressive (H3K27me3) and activating (H3K4me1, H3K4me3, H3K9Ac) epigenetic marks. These chromatin domains are readily convertible between reversibly repressed (H3K27me3 associated) and active (H3K4me1, H3K4me3, and H3K9Ac associated) states. The loci displaying hypomethylation with age are more associated with active “weak” promoters and enhancers mainly enriched with H3K4me1 marks.
LIFESPAN: EPIGENETIC ASPECTS

Dietary restrictions have enormous importance in health generally and in aging specifically. Caloric restriction is one of the most thoroughly studied ways to increase the lifespan in mammals and other organisms [22]. Besides increasing lifespan, caloric restriction decreases frequency of a large number of aging-associated diseases (cancers, cardiovascular and neurodegenerative diseases, atherosclerosis, diabetes). Culturing on a glucose-free medium enhanced lifespan (the period of active proliferation before senescence) of human lung diploid fibroblasts, whereas the lifespan and proliferation of immortalized fibroblasts were greatly diminished [23].

The expression level of the hTERT (human telomerase reverse transcriptase) gene encoding a key component of telomerase was several-fold increased on glucose-free medium compared with the control in normal fibroblasts and several-fold decreased in immortalized fibroblasts. The opposite effects were found for expression of the tumor suppressor gene p16, which was downregulated by glucose restriction in normal fibroblasts and upregulated in immortalized ones. The changes in the hTERT gene expression levels were correlated with an increase in active histone marks (acetyl-H3, acetyl-H4, and H3K4me2) and a decrease of repressive marks (H3K9me3 and HDAC1) associated with its promoter, whereas the opposite was observed in immortalized cells. Changes in the H3K9me3 levels were rather small, whereas those in other marks were many-fold. Essentially opposite results were obtained for the p16 gene promoter, the levels of active histone marks being decreased in normal cells and increased in immortalized ones. The methylation pattern of the hTERT promoter CG-island was not significantly changed by glucose restriction. In the p16 gene, two promoter CG sites located inside the transcription factor E2F-1 binding sequence showed changed methylation on glucose-free medium in normal fibroblasts, but not in immortalized cells. A DNA methylation inhibitor, AzaC, was found to stimulate activity of the p16 gene promoter and to suppress activity of the hTERT gene promoter in normal fibroblasts. These effects were stronger on glucose-free medium compared with the control. The effects of AzaC on promoter activity of both genes were insignificant in immortalized cells. Thus, immortalization could be suggested to somehow disturb the epigenetic control of gene activity on glucose restriction and lead to rapid cell death. And vice versa, the methylation of the p16 gene promoter, that suppress E2F-1 binding, cause its repression in normal fibroblasts, thus ensuring escape from early senescence and apoptosis. The growth on a glucose-free medium enhanced the total DNA methyltransferase activity 4.5-fold and the histone deacetylation activity 1.7-fold in normal fibroblasts. These effects could well be responsible for the p16 gene promoter methylation described above. The effects of glucose-free medium on both activities were insignificant in immortalized cells. In total, these results show that the normal cell response to a caloric restriction includes activity of epigenetic systems modulating gene expression to optimize cell growth and survival.

Among the genes found to be modulated by caloric restriction in various organisms is the gene for S-adenosyl-L-methionine synthetase (SAM synthetase), which is responsible for the synthesis of S-adenosyl-L-methionine, a universal methyl donor for most methylation reactions in cells [24]. Among the proteins found to interact with those involved in realization of the caloric restriction effects is the enzyme S-adenosylhomocysteinase, which catalyzes hydrolysis of S-adenosylhomocysteine (SAH), a competitive methyltransferase inhibitor. The maintenance of SAM to SAH ratio inside certain limits is one of the crucial factors of normal DNA methylation, whereas its disturbance in aging cells could be partially responsible for global DNA hypomethylation [25].

There is very little direct evidence of existing links between DNA methylation and lifespan. In contrast to mammals, Drosophila has a single cytosine DNA methyltransferase, Dnmt2. In several independent transgenic lines, on overexpressing the Dnmt2 gene lifespan was found to be increased by 16-58% compared with control flies [26]. Conversely, in a mutant line having a P-element insertion inside the first exon of Dnmt2, the lifespan was decreased by 27%. The expression of several classical longevity genes, such as InR (insulin-like receptor), chico (InR substrate protein), mth (a G-protein-coupled receptor), and SOD (Cu/Zn superoxide dismutase) appeared to be unchanged in all fly lines studied. However, the expression levels of several genes encoding small heat shock proteins (Hsp22, Hsp23, and Hsp26) were increased about 3-fold in the Dnmt2 superexpressing lines, and decreased 2-3-fold in the mutant line. These proteins are chaperones affecting aging. Overexpression of Hsp22 (which is a mitochondrial protein) and the cytosolic proteins Hsp23 and Hsp26 was found to increase lifespan in Drosophila. Decrease in the Hsp22 expression due to a P-element insertion conversely was found to shorten Drosophila lifespan. Overexpression of Hsp22 and Hsp26 also increased resistance to oxidative and heat stress. Thus, modulation of Hsp22, Hsp23, and Hsp26 expression may be mediating, at least partially, the effects of Dnmt2 on lifespan. It is interesting that expression of a foreign DNA methyltransferase gene, Dnmt3a, in Drosophila leads to developmental derangement and lethality at pupal stages, coexpression of the Dnmt3a and Dnmt1 genes aggravating this effect [27]. Whether these findings are applicable to mammals and to what extent is presently unknown. The exact biological role of Dnmt2 in mammals is still unknown. No apparent phenotypic defects were observed in embryonic stem cells homozygous for a null allele of Dnmt2 [28], and in mice produced from these cells and propagated for seven generations as a homozygous line.
A popular view is that Dnmt2 functions not as a DNA methyltransferase, but rather as an RNA methyltransferase methylating the 38th cytosine residue in the anticodon loop of the asparaginic acid tRNA [30]. Dnmt2 appears to be the only known enzyme capable of RNA methylation by a catalytic mechanism characteristic of DNA methyltransferases. Whichever activity of Dnmt2 is the primary one, both could be supposed to have a certain biological significance. A reasonable view from the relevant accumulated evidence is that either of these two activities could gain a predominant role in the evolution of various species.

Possible significance of DNA methylation as a factor affecting lifespan in humans was studied in a comparative investigation of DNA methylation patterns in leukocytes of female centenarians (>100 years age), their daughters of about 70-years of age selected from pairs, where the father was also long-lived (died at age >77 years), females of about 70 years ages whose parents were both non-long-lived (mothers died at ≤72 years, fathers at ≤67 years), and, last, a control group of young (17-34 years) women [31]. Clinical histories showed the centenarians’ daughters to have much better health status (as judged by the prevalence of several age-related diseases, constant use of prescribed drugs, etc.) compared with daughters of non-long-lived parents. Evidently, the probability to become long-lived is inheritable to a very significant degree. Just as it could be expected, the global DNA methylation levels were significantly decreased in all three aged groups compared with the control group of young women, but to different extents. Maximum hypomethylation was observed in daughters of non-long-lived parents, minimal — in centenarians’ daughters, intermediate — in centenarians themselves. The methylation of Alu family repeated sequences in centenarians’ daughters was unchanged, whereas in offspring of non-long-lived parents and in centenarians it was decreased to similar degrees. The methylation levels of various loci were investigated by microchip hybridization on the Infinium Human Methylation 27 BeadChip platform. In 607 genes, 709 CG sites were detected that were hypermethylated compared with the control in all three aged groups, to similar extents in daughters of centenarians and non-long-lived parents, and largely in centenarians themselves. A large number of genes among this hypermethylated group were functionally related to organ development, cell differentiation, and transcription regulation. On the other hand, 330 CG sites (located in 326 genes) were found that were hypomethylated in aged subjects, to similar extents in both daughters’ groups, and largely in the centenarians’ group. Large numbers of genes involved in defense responses, acute inflammation, and signal transduction were found among these hypomethylated loci. A detailed comparison of methylation patterns revealed 150 CG sites (located in 124 genes) that were significantly hypermethylated in centenarians’ daughters compared with daughters of non-long-lived parents. Genes functionally related to nucleotide metabolism and nucleic acids synthesis were enriched in this group. On the other hand, 67 CG sites (located in 65 genes) were found that showed significant hypomethylation in centenarians’ daughters compared with daughters of non-long-lived parents. The genes related to signal transduction were predominant in this group. Most strongly pronounced differences in methylation levels were found in twelve CG sites (10 hypermethylated and two hypomethylated in centenarians’ daughters) located in nine genes. Six of the hypermethylated genes (SLC38A4, SLC22A18, MGC3207, ECRG4, ATP13A4, and AGPAT2) are involved in metabolic processes, one hypermethylated, DUSP22, is a tumor-suppressor gene, still another hypomethylated gene, ZNF169, encodes a zinc finger DNA-binding protein with unknown function, and last, the function of the only hypomethylated gene, FLJ32569, is also unknown. Overall, the genome methylation in centenarians’ daughters is obviously more stable compared with daughters of non-long-lived parents. The epigenome stability, as well as a more robust epigenetic control for nucleotide metabolism, nucleic acids synthesis, and signal transduction may contribute to increase in lifespan and healthy aging in centenarians.

METHYLATION ERRORS OR A SPECIFIC PROGRAM?

To understand what part of the age-related changes in DNA methylation represents a cumulative result of stochastic errors and what part results from realization of a hypothetical specific aging program, discordances in DNA methylation patterns between monozygotic twins were studied [32]. Since MZ twins originate from the same zygote, they are regarded to be genetically identical. Thus, phenotypic differences between them could be explained by either accumulation of stochastic events or specific responses to environmental factors. In female MZ twin pairs, 81% were found to have identical patterns of the X chromosome loci methylation and 19% — discordant. Hence, the epigenetic differences could arise at early developmental stages. The methylation levels of total lymphocyte DNA were practically identical between MZ twins in 65% of the pairs and significantly different in 35%. Identical DNA methylation levels were observed in young pairs, whereas aged pairs had the most differences. The differences were not a simple consequence of age-dependent increase in general DNA methylation variability, since DNA methylation levels were variable to similar extents in young (<28 years) and aged (>28 years) twin groups. Thus, the methylation discordance between MZ twins gradually increases with age. An analysis of the differentially methylated genome loci in the most epigenetically discordant twin pairs showed that 43% of these loci...
are located in the Alu family repeated sequences, 9% – in repeated sequences of other families (LINE, MER, MIR), 34% – in unidentified transcribed sequences, and, last, 13% – in known unique genes. Different methylation levels were correlated with differences in expression levels, hypermethylated alleles usually being expressed less or not expressed at all.

Generally, nearly identical methylation patterns were characteristic of young MZ twin pairs that lived together for most of their life and had similar lifestyles, whereas most discordant methylation patterns were characteristic of older twin pairs that lived separately and had different lifestyles. In the most epigenetically discordant 50-year-old twin pairs, differences in the methylation patterns of CG islands associated with unique genes were observed 2.5 times more often compared with young epigenetically close twin pairs. No differences in gene expression were detected in 3-year-old twin pairs, whereas such differences were frequently observed in 50-year-old pairs. Similar results were obtained for other cell types studied (epithelial mouth cell, intraabdominal fat, and skeletal muscle biopsies). Thus, many phenotypic discordances in MZ twin pairs may be caused by the accumulated epigenetic differences. Causes of the age-dependent increase in epigenetic differences could be both various effects of external and internal factors (smoking, physical activity, dietary preferences, etc.) and stochastic methylation errors, “epigenetic drift”, accompanying aging. Methylation errors could be suggested to occur much more often than mutations, since systems of their correction are far less effective compared with the DNA repair systems. The existing evidence show that both accumulation of stochastic errors and epigenetic responses to external and internal factors contribute to the net epigenetic variability. The first being the main source of age-dependent increase in epigenetic variations, whereas the second being responsible for increase in variations caused by separate living and different lifestyles.

A comparative study of DNA methylation in a newborn infant and a 103-year-old centenarian showed that the total number of methylated CG sites in lymphocyte DNA was 3% less in the centenarian than in the newborn, and their average methylation level was also less in the centenarian compared with the newborn (73 and 80.5%, respectively) [33]. In a 26-year-old man, the total number of methylated sites was ~0.4% less than in the newborn, whereas the average level of their methylation was intermediate between levels found in newborn and centenarian (77.8%). The correlation between the methylation statuses of neighboring CG sites decreases with age. The age-related methylation loss affects sequences of all chromosomes and of all types (promoters, exons, introns, intergenic sequences). Compared with newborn, the centenarian DNA was found to be less methylated at CG-poor promoters and more methylated – at CG-rich promoters. In total, nearly 18,000 differentially methylated regions (DMRs) were found between newborn and centenarian DNA methylomes, 10.2% at promoters, 10.1% at exons, 45.3% at introns, and 34.4% at intergenic regions. Among the latter, the most frequent DMR-associated sequences were LINE 1 repeats (3777), Alu repeats (3242), interspersed repeats (MIRs) (1890), LTR-retrotransposons (1448), and LINE 2 repeats (1321). Of all DMRs found, 87% were more methylated in newborn, whereas only 13% were more methylated in centenarian. DMRs hypomethylated in centenarian were associated not only with repeated sequences, but also with promoters containing repressive histone marks (H3K27me3, H3K9me2, H3K9me3), whereas hypermethylated DMRs were associated with CG-rich promoters.

EPIGENETIC CLOCK AND EPIGENETIC AGE CONCEPT

The aging rates are not equal for different persons. Thus, women are known to have a longer average lifespan compared with men. Nearly all supercentenarians, people who have reached the 110 years age, are women. Aging can be accelerated or slowed by such life habits as smoking or physical training. Thus, searches are encouraged for robust molecular markers of aging that could be used to estimate biological, rather than chronological (passport) age, which corresponds more precisely to age-dependent changes in physiological functions. The age-related epigenetic variations, especially methylation of certain DNA loci, seem to be good contenders for this role. The main obstacle in searches for aging epigenetic markers is that age-related methylation loci are masked by a plethora of other epigenetic variations in our epigenome that are caused by other factors, including stochastic methylation errors.

An analysis of DNA methylation in saliva samples of 34 pairs of identical twins between 21 and 55 years of age revealed 88 sites in or near 80 genes for which the degree of cytosine methylation was significantly correlated with age [34]. Of these, 19 sites were negatively correlated, and 69 were positively correlated with age. Most (83%) age-correlated sites were within CpG islands and on average 283 bp upstream of the transcription start sites. The 80 age-related genes were highly enriched for genes involved in cardiovascular and neurological disease, the most enriched cellular function being molecular transport. Three genes for which the methylation status showed most clear correlation with age and which had the widest distribution of methylation values were tested in additional populations of both genders with a larger age range (18-70 years). All three genes tested, Edaradd (Edar associated death domain), NPTX2 (Neuronal Pentraxin II), and TomIL1 (target of myb1 like 1 membrane trafficking protein) showed a clear correlation with age in men, but only two of them (Edaradd and TomIL1) – in women. The
methylation levels of Edaradd and TomIL1 decreased linearly with age, whereas the methylation level of NPTX2 increased. Based on methylation levels of just two CG sites (located in the Edaradd and NPTX2 genes), the age of the test subjects could be predicted with 5-6-year accuracy, whereas addition of one more site (located in ELN gene) reduced the average error to 3.5 years. The methylation of NPTX2 is known to increase in pancreatic cancer, and its expression is increased in Parkinson’s disease. Its methylation status was shown to correlate with age in blood cells as well [17]. Mutations in the Edaradd gene can cause loss of hair, sweat glands, and teeth, and reduce the speed of wound healing. It is interesting that no epigenetic drift was detected in promoter studied CG sites. This finding corroborates the view that stochastic methylation errors are mainly accumulated in repeated sequences and intergenic regions, whereas gene and promoter methylation is under more robust control.

A comparative analysis of age-related methylation patterns in various tissues shows these patterns to be highly tissue-specific. Nevertheless, a small group of loci exists whose methylation is significantly correlated with age in various tissues. The assumption that these methylated loci are most related to the aging per se seems quite reasonable. Thus, the methylation patterns of such loci could be used as an epigenetic signature to estimate biological age [35]. An assessment of results obtained in several independent studies on the Infinium Human Methylation 27 BeadChip platform revealed 19 CG sites that showed correlation between methylation level and age by most stringent criteria ($p < 10^{-13}$), the methylation degree of all these sites being increased with age. More than 400 age-correlated sites were found by a less stringent criteria ($p < 10^{-5}$), most of them also being hypermethylated with age and only 15 being hypomethylated. To construct an epigenetic aging signature, CG sites were chosen that correlated with age across all studies analyzed and had largest methylation variation ranges. The most accurate age predictions were obtained by use of four hypermethylated loci, TRIM58 (tripartite motif-containing 58), KCNQ1DN (KCNQ1 downstream neighbor), NPTX2 (neuronal pentraxin II), and GRIA2 (glutamate receptor ionotropic AMPA 2). To further enhance the prediction accuracy, a hypomethylated locus was added to the set, namely BIRC4BP (XIAP associated factor-1).

Interestingly, two groups [17, 35] have chosen NPTX2 as a diagnostic locus using different bioinformatic approaches, whereas TRIM58 and GRIA2 loci were among those most significantly correlating with age in both cases. When all five loci were used, the prediction accuracy was about 13 years, whereas the use of only the three most reliable (NPTX2, GRIA2, KCNQ1DN) enhanced the accuracy to ~11 years. One should take into account that in the described work the age prediction was applicable to various tissues and was gender-independent, whereas in previous study [34] prediction was based only on saliva samples containing leukocytes and buccal epithelial cells. When blood samples were investigated, the number of CG sites with high predictive capabilities could be narrowed to just three, located in the ITGA2B, ASPA, and PDE4C genes [36]. The accuracy of age prediction was 5 years or better.

In larger-scale investigation on the Illumina Infinium Human Methylation 450 BeadChip assay platform (485,577 CG sites), blood DNA samples from more than 650 volunteers 19 to 101 years of age were used [37]. A correlation with age was found for 70,387 (15%) CG sites (excluding those located on sex chromosomes). Based on analyses by various bioinformatic methods, an aging predictive model including methylation of 71 sites was constructed. The mean error of age prediction by this model was less than four years (correlation of 96% between the predicted and passport ages).

Nearly all CG sites used in the model are associated with genes with known functions in aging-related conditions (Alzheimer’s disease, cancer, tissue degradation, DNA damage, and oxidative stress). For example, two sites are within the somatostatin gene (SST), a regulator of endocrine and nervous system function, and six sites are within the gene encoding transcription factor KLF14, a “master regulator” of obesity and other metabolic diseases. The model not only predicts age, but also reveals factors affecting the prediction accuracy. For example, the body mass index (BMI) appeared not to affect the aging rate, whereas gender affects it very substantially, “aging” of the methylome being ~4% faster in men compared with women. Since there is an evident dependency of lifespan and aging characteristics on genotype, a whole-exome sequencing was carried out for 252 test subjects that revealed 10,694 single-nucleotide polymorphisms (SNPs), of which 303-affected methylation of CG sites included in the epigenetic age-predicting model. Among them is SNP rs140692 affecting methylation of a CG site within the MBD4 (methyl-CpG binding domain protein 4) gene, these SNPs being located within an intron of the gene, whereas the affected CG site was just upstream of its coding sequence. Other SNPs affect methylation of some loci related to changed rates of aging. Genetic marker rs2230534 is a synonymous mutation in the NEK4 gene and affects its methylation; the encoded NEK family kinase plays a role in cell-cycle regulation and cancer. The rs2818384 marker is a synonymous mutation in the GTPBP10 gene and affects its methylation; variants of this gene have been noted in association with glioblastoma. Finally, marker rs42663 is a missense mutation in the GTPBP10 gene and affects methylation of the STEAP2 gene. The latter is known to play a role in maintaining homeostasis of iron and copper.

The age prediction model appeared to work with samples of other organs (breast, lung, kidney, skin) with the same accuracy as with blood samples when a linear offset was used (intercept of vertical axis and slope of the
calibration line) that was specific for each organ. Hence, the model reflects some processes common for various organs, not only blood. When epigenetic predictive models were constructed with the same algorithm based on age-related methylation data from other organs (breast, lung, and kidney), the main differences were in the sets of most informative CG sites chosen. Only two CG sites near the ELOVL2 gene, involved in skin cell aging, appeared to be common.

When the models constructed were used to estimate the age of tumor and respective normal cells, the tumors appeared to be 40% more aged than the normal cells of the same person. Not only the methylation levels per se of age-related CG sites were changed with age, but also the variation limits of these methylation levels between different persons became larger for most sites. For any one person, the extent of deviation in these values from population averages appeared to be an accurate measure of individual aging rate. Studies of the aging peculiarities using predictive epigenetic models could have many practical implications, from health assessment to forensic analysis. Similarly to the analysis of the influence of gender on the aging rates described above, the effects of various environmental and lifestyle factors, such as smoking, alcohol consumption, dietary preferences, and many others, could be assessed. As the predictive accuracy of the model improves, it seems quite probable that biological age measured by epigenetic and other molecular markers might become more useful than passport age in clinical practice.

To build a universal age predictor, bioinformatic analysis was performed with all publicly available data sets concerning age-related variations of DNA methylation in various normal tissues and cell lines (in total nearly 8000 samples, 51 tissue and cell types) [38]. The developed multi-tissue age predictor system estimates with high accuracy and reproducibility epigenetic age (DNAm age) and can be used in various aging studies. For example, it shows that premature aging diseases (progerias) do not look like accelerated normal aging from the epigenetic point of view, and that the epigenetic age is reset when induced pluripotent stem cells (iPSCs) are created.

To build this system, from the total 21,369 CGs assessed by both widely used Illumina platforms (Illumina Infinium 450K and 27K), 353 sites were chosen that allow for the most accurate age prediction for various tissues and cells (96% correlation to passport age, error ±3.6 years). The age predictor is particularly accurate when used for age assessment in tissues of children and adolescents. Blood consists of cells that have very different lifespans. Thus, while CD14⁺ monocytes (myeloid lineage) live only several weeks, CD4⁺ T-cells (lymphoid lineage) live from months to years. Nevertheless, epigenetic ages of these cells in blood samples of healthy male subjects are identical. Hence, epigenetic age reflects some internal methylome features related to chronological age, not just age-dependent peculiarities of the blood cell composition. A similar correlation to chronological age is observed when neurons and glial cells are used, and in different brain divisions. The system is less accurate (has highest errors) in breast tissue, uterine endometrium, dermal fibroblasts, skeletal muscle tissue, and heart tissue. This may be due to hormonal effects or peculiarities of cell renewal in these tissues. Mean epigenetic age is highly correlated (correlation value 0.99) with mean chronological age in most tissues. Breast tissue displays epigenetic features of accelerated aging. The variations in epigenetic ages between different tissues of the same person are rather small. Notable exceptions found are breast tissue in women (epigenetically older compared with other tissues) and sperm in men (epigenetically younger compared with other tissues). Surprisingly, the developed system could be used to predict with fair accuracy the age of chimpanzee tissues, whereas its performance in gorillas was significantly worse.

As could be expected, embryonic stem cells appeared to have an epigenetic age close to zero. Induced pluripotent cells do not differ from embryonic stem cells by their epigenetic age; therefore, these cells are much younger than their source primary cells. Unfortunately, due to technological limitations of the platforms used (Illumina Infinium 450K and 27K), DNA methylation at non-CG sites, characteristic of embryonic stem cells and induced pluripotent cells, remained beyond the scope of the elaborated predictive models. When cells are maintained in culture (e.g. embryonic stem cells or induced pluripotent cells), their epigenetic age increases with each passage.

A functional analysis of 353 predictive CG sites showed most of them to be associated with genes involved in cell death-survival, growth-proliferation, organ or organism development, and cancer. Of the 353 sites, 193 are positively correlated (hypermethylated) and 160 are negatively correlated (hypomethylated) with age. Similar to findings of other studies, the methylation of the former were least variable, whereas the methylation of the latter was most variable in different tissues.

The ticking rate of the epigenetic clock is maximal in the first 18 years of life, slowing to a constant rate thereafter. Epigenetic age is clearly not a reflection of mitotic age, since it tracks chronological age in tissue widely different in proliferative potential, including postmitotic neurons. It is also not related to cell senescence, since its correlation to chronological age is observed in immortal cell lines, such as embryonic stem cells. As the author of the work described has proposed [38], epigenetic age could be regarded as a quantitative measure of the cumulative work done by the epigenome maintenance systems. Thus, its value equals zero in “newborn” embryonic stem cells. The epigenome of these cells, just formed, could be considered an epigenetic starting point. The early developmental stages involve massive differentiation events of various cell lines based on epigenome conversions. These
events evidently demand high activity of the epigenome maintenance systems. It could well be that large epigenome rearrangements are accompanied by large numbers of epigenetic errors, whose correction further enhance the demands for activity of the epigenome maintenance systems. All these on the whole would increase the ticking rate of the epigenetic clock. When the early developmental stages end, the number of epigenetic rearrangements (mainly coupled to tissue renewal) and of the accompanying epigenetic errors will stabilize, and thus the rate of ticking of the epigenetic clock will become constant.

One of the predictions from the model is that the rate of the epigenetic aging must be increased by epigenome destabilizing agents. Indeed, tumor tissue cells show evidence of aging acceleration by 36 years on average. Since that acceleration is caused by enhanced activity of the epigenome maintenance systems, it should have a genome stabilizing effect, i.e. decrease the somatic mutation frequency. Such correlation was found in several tumor tissue samples (bone marrow, breast carcinoma, kidney carcinoma, ovarian cancer, prostate, and thyroid cancer). The TP53 gene mutations were associated with lower epigenetic aging in five different cancer types. This is in a good agreement with the role of p53 as a triggering factor in the protective action of epigenome maintenance systems. In breast cancers, mutations of the estrogen receptor (ER) or progesterone receptor (PR) genes accelerate epigenetic aging, whereas amplification of the HER2/neu gene does not affect the aging rate. Colorectal cancer samples with BRAF gene mutations are associated with a higher aging acceleration compared to those with the K-RAS gene mutations. Promoter hypermethylation of the mismatch repair gene MLH1 leads to the most significant acceleration of aging. In glioblastoma samples, mutations of the H3F3A gene encoding replication-independent histone variant H3.3 variously affect the rate of epigenetic aging. The lysine 27 to methionine replacement (K27M) causes a much smaller aging acceleration compared with glycine 34 to arginine replacement (G34R). This finding is in a good accord with distinct effects of these mutations on the global genome methylation patterns. The levels of aging acceleration vary significantly across other known glioblastoma subtypes.

The epigenetic age concept can be used in various developmental, aging, and cancer studies. The epigenetic age can evidently serve as a convenient marker in the testing of rejuvenating treatments. It could help to test the aging rates of different tissues of the same person to identify those with evidence of accelerated aging probably caused by a disease, especially cancer. It remains to be clarified whether the epigenetic age is just an aging marker, or it may be somehow related to the causes of aging. An interesting example of application of epigenetic age was estimation of mortality risk [39]. It was found that an acceleration of epigenetic aging to a five-year difference between epigenetic age and chronological age results in a 16% increase in mortality risk, irrespective of general health, lifestyle, and genetic factors.

Predictive accuracy can be enhanced when less universal models are used. For example, for blood cells it reached a value of 2.6 years when only 17 marker CG sites were used [40]. In a follow-up study of the same persons 8 years later, the predicted increases in methylation levels of hypermethylated sites and decreases in methylation levels of hypomethylated sites were observed.

**IS DECELERATION, ARREST, OR RESTART OF THE EPGENETIC AGING CLOCK POSSIBLE?**

An example of the biological aging clock restart is fertilization; in humans, the fusion of an oocyte and a sperm cell, each of them having an age of about twenty years or higher, results in formation of a zygote that has an age equal to zero. Since there are no reasons to believe germinal cells to be specifically protected from chronological aging, the species would age with every generation, suffering the burden of the effects of aging of all antecedent generations in their cells if there is no such biological clock reset. A similar clock reset takes place when the nucleus of a somatic cell is transferred to the cytoplasm of an enucleated oocyte (SCNT), resulting in development of a new normal individual. Thus, the oocyte cytoplasm possesses the capability to “erase” all the aging features accumulated in the nucleus of a somatic cell. These investigations prove the age-related features of the somatic cell nuclei, whatever is their nature, to be in principal reversible. The mechanisms of this reversibility are understandable if the age is viewed as an epigenetic one.

The epigenome of any differentiated cell (including germinal cells) is a complex mosaic of epigenetic marks, where age-related marks are intermixed with invariant ones, developmentally changed, arisen as responses to external factors, or as a result of spontaneous errors, and, probably many other factors. In this situation, the easiest way to reset the epigenetic clock is to fully erase all existing epigenetic information and then rebuild it from scratch in a form corresponding to zero age. Something like this takes place during the first hours following fertilization [41-43]. However, this view does not exclude the accumulation of some irreversible changes during chronologic aging, provided that these changes are not the specific causes of the aging per se. Evidently, gradual accumulation of mutations increase the genetic variability of humans as a biological species, but they are not directly related to aging. On the other hand, similarly to any other cells, harmful mutations eventually arise in germinal cells, including those accelerating aging. It could be suggested that these mutations are contra-selected at fertilization (defective sperm disposal), prenatal develop-
agement (defective embryo lethality), or early postnatal development (early postnatal lethality) stages. Nevertheless, some harmful mutations succeed in passing all these contra-selection filters, as evidently follows from the existence of a few thousand genetic diseases in the human population. Still another example of experimental resetting of the biological clock is production of induced pluripotent stem cells [44].

In all cases described above, resetting of the epigenetic clock was coupled to cell dedifferentiation. Is it possible to rejuvenate a cell without disturbing its differentiation status? By an experimental procedure known as heterochronic parabiosis (joining together of the circulatory systems of two differently aged mice), the skeletal muscle and liver cells of the old mouse were shown not only to adopt a more youthful phenotype, but also to restore molecular signatures of aging to a more youthful state [45], whereas opposite changes were observed in the younger mouse [46]. The systemic environment seems to influence the biological age of cells to a significant degree. Most likely, its effects are caused by changes in activity of the main signaling pathways (Notch, Wnt, and TGFβ) due to changed concentrations of various cytokines and chemokines.

Molecular profiles of biological age obtained by transcriptome analyses showed certain NF-κB pathway-regulated genes to be more actively expressed in aged cells [47]. In transgenic mice with NF-κB gene conditionally expressed in skin cells, inhibition of its expression was accompanied by indications of skin rejuvenation; markers of cell senescence (p16) disappeared, proliferative activity of progenitor cells was regained, skin constitution and general condition improved. In hematopoietic stem cells, mTOR activity increases with age, and rapamycin not only decreases the content of aging markers in these cells in aged mice, but also enhances their activity to the level of young stem cells in heterochronic transplantation experiments [48].

In all cases described, the rejuvenation of differentiated or committed cells was achieved without disturbing their differentiation programs. The epigenetically reprogrammed cells obtained and young cells could hardly be identical, since the former should contain some irreversible mutations accumulated during their previous aging. However, although there is little doubt concerning the role of mutations in aging, there is no evidence of their being a direct cause of aging. Perhaps accumulated mutations are responsible for the low efficiency of somatic cell reprogramming, limiting successful reprogramming to a small share of cells.

The possibility of resetting the aging clock suggests that the epigenetic signals are not the mere correlates of the aging process, but rather a substantial part of the aging mechanism. The link between epigenome and aging can be mutual; epigenome is changed by age factors, but it also affects aging. One can suppose the epigenome to be a general sensor of cellular dysfunction, responsive to any changes in the genome and internal milieu, including those related to aging. On the other hand, the epigenome affects aging by determining differential gene expression patterns in cells at variable ages. These patterns underlie a substantial part of the aging phenotype features, such as cell senescence and metaplasia [49]. Similarly to other cell functional systems, the epigenome is prone to gradual degradation due to genome damage, stressful agents, and other aging factors. Rejuvenating interventions could not reverse the breach in genome integrity, but age-related epigenetic changes could be fully or partially reversed to a naive state. It is interesting that many of the rejuvenating agents act on stem cells; perhaps there are some common genetic and biochemical pathways affecting stem cell function and lifespan [50]. Probably, many aspects of organismal aging depend on changed capability of stem cells to self-renew. Another important proposition is that the epigenome could mediate the action of all genetic, internal, and external influences that change lifespan.

This work was supported by a grant from the Russian Science Foundation (project No. 14-50-00029).

REFERENCES


