Clinical Epigenetics

RESEARCH

Open Access

Epigenetics of the far northern Yakutian population

Alena Kalyakulina[1,2*†], Igor Yusipov[1,2†], Elena Kondakova[2,3†], Maria Giulia Bacalini[4], Cristina Giuliani[5], Tatiana Sivtseva[6], Sergey Semenov[6], Artem Ksenofontov[7], Maria Nikolaeva[6], Elza Khusnutdinova[8], Raisa Zakharova[6], Maria Vedunova[3], Claudio Franceschi[1,2] and Mikhail Ivanchenko[1,2]

Abstract

Background Yakuts are one of the indigenous populations of the subarctic and arctic territories of Siberia characterized by a continental subarctic climate with severe winters, with the regular January average temperature in the regional capital city of Yakutsk dipping below −40 °C. The epigenetic mechanisms of adaptation to such ecologies and environments and, in particular, epigenetic age acceleration in the local population have not been studied before.

Results This work reports the first epigenetic study of the Yakutian population using whole-blood DNA methylation data, supplemented with the comparison to the residents of Central Russia. Gene set enrichment analysis revealed, among others, geographic region-specific differentially methylated regions associated with adaptation to climatic conditions (water consumption, digestive system regulation), aging processes (actin filament activity, cell fate), and both of them (channel activity, regulation of steroid and corticosteroid hormone secretion). Further, it is demonstrated that the epigenetic age acceleration of the Yakutian representatives is significantly higher than that of Central Russia counterparts. For both geographic regions, we showed that epigenetically males age faster than females, whereas no significant sex differences were found between the regions.

Conclusions We performed the first study of the epigenetic data of the Yakutia cohort, paying special attention to region-specific features, aging processes, age acceleration, and sex specificity.

Keywords Yakutia, DNA methylation, EWAS, GSEA, Aging, Age acceleration, Cold environment, Climate, Sex specificity

†Alena Kalyakulina, Igor Yusipov, Elena Kondakova: Co-first authors.

*Correspondence:
Alena Kalyakulina
kalyakulina.alena@gmail.com
[1] Institute of Information Technologies, Mathematics and Mechanics, Lobachevsky State University, Nizhny Novgorod 603022, Russia
[2] Institute of Biogerontology, Lobachevsky State University, Nizhny Novgorod 603022, Russia
[3] Institute of Biology and Biomedicine, Lobachevsky State University, Nizhny Novgorod 603022, Russia
[4] ISNB Institute of Neurological Sciences of Bologna, 40139 Bologna, Italy
[5] Laboratory of Molecular Anthropology and Centre for Genome Biology, Department of Biological, Geological and Environmental Sciences, University of Bologna, 40126 Bologna, Italy
[6] Research Center of the Medical Institute of the North-Eastern Federal University M.K. Ammosova, Yakutsk 677013, Russia
[7] State Budgetary Institution of the Republic of Sakha (Yakutia) Republican Center for Public Health and Medical Prevention, Yakutsk 677001, Russia
[8] Institute of Biochemistry and Genetics, Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia 450054

Background

The Yakuts (Sakha) are people living in the subarctic and arctic territories of eastern Siberia. Anatomically modern humans inhabited the region of the modern Republic of Sakha about 30,000 years ago [1], moving from the west, where Sakha connects to the inner Eurasian steppe belt through southern Siberia [2, 3]. The northeastern part of Yakutia is located on one of the main migration routes from the southern regions of the Yenisei, Amur and Baikal coasts to the Arctic coast and to America [4]. Thus, the native inhabitants of Siberia, including Yakuts, are also viewed with regard to American colonization [5, 6]. There is evidence bringing Yakuts closer to Amerindians in terms of genetic variability, which may indicate a common ancestry of Siberians and Native Americans [7]. The most recent wave of migration to the territory of modern Yakutia involved the Tungus of Transbaikalia and Turkic-speaking Yakuts of the Western Baikal region and took place about 2000 years ago [8].

Eastern Siberia is one of the permanently inhabited regions with an extremely cold climate. From the evolutionary perspective, populations surviving in such an extreme climate for many years should have accumulated genetic changes adapting them to the cold and to other local factors such as seasonal extremes in daylight, food availability, etc. [9]. Indeed, low serum lipid levels were identified and related to the increased energy metabolism [10], and higher blood pressure was also observed [11, 12]. A large study of Siberian populations identified candidate genes for adaptation to the cold, associated with energy and metabolic regulation, as well as contraction of vascular smooth muscle [9]. Besides, leptin and irisin, which play an important role in the processes of adaptation to the cold, become the focus of several other studies of the Yakutian cohort [13–15]. However, genetic studies for Siberian populations are often limited to mitochondrial DNA, Y chromosome or single-nucleotide polymorphisms [8, 9, 16, 17], while epigenome-wide studies have not been performed.

Epigenome-wide association studies (EWAS) usually analyze DNA methylation and focus on differentially methylated CpG sites (differentially methylated positions or DMPs), similar to single-nucleotide polymorphisms in genome-wide association studies (GWAS) [18]. EWAS detects epigenetic differences between human cohorts by quantitatively comparing methylation levels of thousands of CpG sites [19]. Individual studies have shown that significant ethnic differences in DNA methylation are reflected in cell composition and risk of some non-communicable diseases [20] and are associated with blood lipid levels [19], body mass index [21], liver function [22], and blood pressure [23]. There are also studies showing environmental [24, 25] and socioeconomic [26–28] influence on DNA methylation. The influence of race/ethnicity has been also studied in the context of health status, mortality, and susceptibility to diseases [29, 30], as well as epigenetic aging speed [31, 32]. Different frequencies of certain genetic variants can lead to epigenetic differences between ethnicities [33–35] and contribute to specific patterns of epigenetic aging.

One of the best-known mechanisms for tracking changes in DNA methylation with age are epigenetic clocks that are models aggregating information about a limited set of CpG sites and used to construct estimates of epigenetic age and mortality risk. The most common clocks are Hannum DNAmAge [36], Horvath DNAmAge [37], GrimAge [38], DNAmPhenoAge [39]; the metric of accelerated epigenetic aging as a deviation of chronological age from that is predicted by the clock is also often considered.

Another interesting aspect would be addressing sex-specific differences in DNA methylation for the cohorts of different ethnicities that live in different climates. EWAS reported differences in DNA methylation associated with sex differences in genes on the autosomes [40, 41]. Moreover, regions of different methylation in males and females are not limited to the sex chromosomes but are genome-wide and can be tracked in various tissues, e.g., brain, pancreas, blood [42–44]. Epigenetic profiles undergo profound changes during aging, with differences between males and females either remaining [45, 46] or changing in different regions of the genome [40, 47]. It can be hypothesized that such sex-specific DNA methylation trajectories may contribute to sex differences in survival [48].

This work is the first epigenetic study focused on the Yakutian population. Since the distinctive feature of this ethnic group is living in severe climatic conditions, it can shed light on the epigenetic features of the adaptation mechanisms to the cold. Further, since severe climate can affect both health and aging, it is interesting to assess age acceleration and aging rate for the Yakutian population. Another aspect of the study is the sex-specific methylation patterns in the Yakutian cohort.

The whole-blood DNA methylation data of the residents of Yakutia and Central Russia collected in this work allow us to perform a comparison between these two cohorts living in very different environmental and climatic conditions. The wide age range and the balanced number of representatives of both regions allow us to compare them and identify region-specific features of DNA methylation. Region-specific CpG sites can be detected by searching for differentially methylated positions (DMPs), and EWAS can analyze them in terms of the corresponding biological processes and pathways. Epigenetic clocks allow us to determine the

age acceleration in the considered regions and compare the aging rate between them. Also, specialized approaches for blood cell deconvolution by whole-blood methylation allow us to compare the estimations of blood cell counts between the regions. Similar age and quantitative distributions of males and females from Yakutia and Central regions in our data enable even deeper analysis and investigation of sex specificity between the regions. Following the same pipelines as for the region specificity analysis, DMPs between males and females in both regions can be found and analyzed using EWAS; epigenetic ages, age acceleration, blood cell composition can also be investigated for males and females. This could reveal whether there are sex differences specific to particular regions.

Results

Participants and study design

This study involved whole-blood DNA methylation data from 245 healthy participants collected in 2020–2022 in the Central region of Russia (Nizhny Novgorod, Vladimir, and Moscow regions, highlighted yellow in Fig. 1A) and Yakutia (Republic of Sakha, highlighted gray in Fig. 1A). All participants from Yakutia are indigenous people, born and living in Yakutsk or in the nearby uluses (villages). The Central region includes 131 samples (78 females and 53 males), and the Yakutia region includes 114 samples (63 females and 51 males). Information about participants is presented in Additional file 1: Table S1. Figure 1B shows the age distributions in the two regions. In the

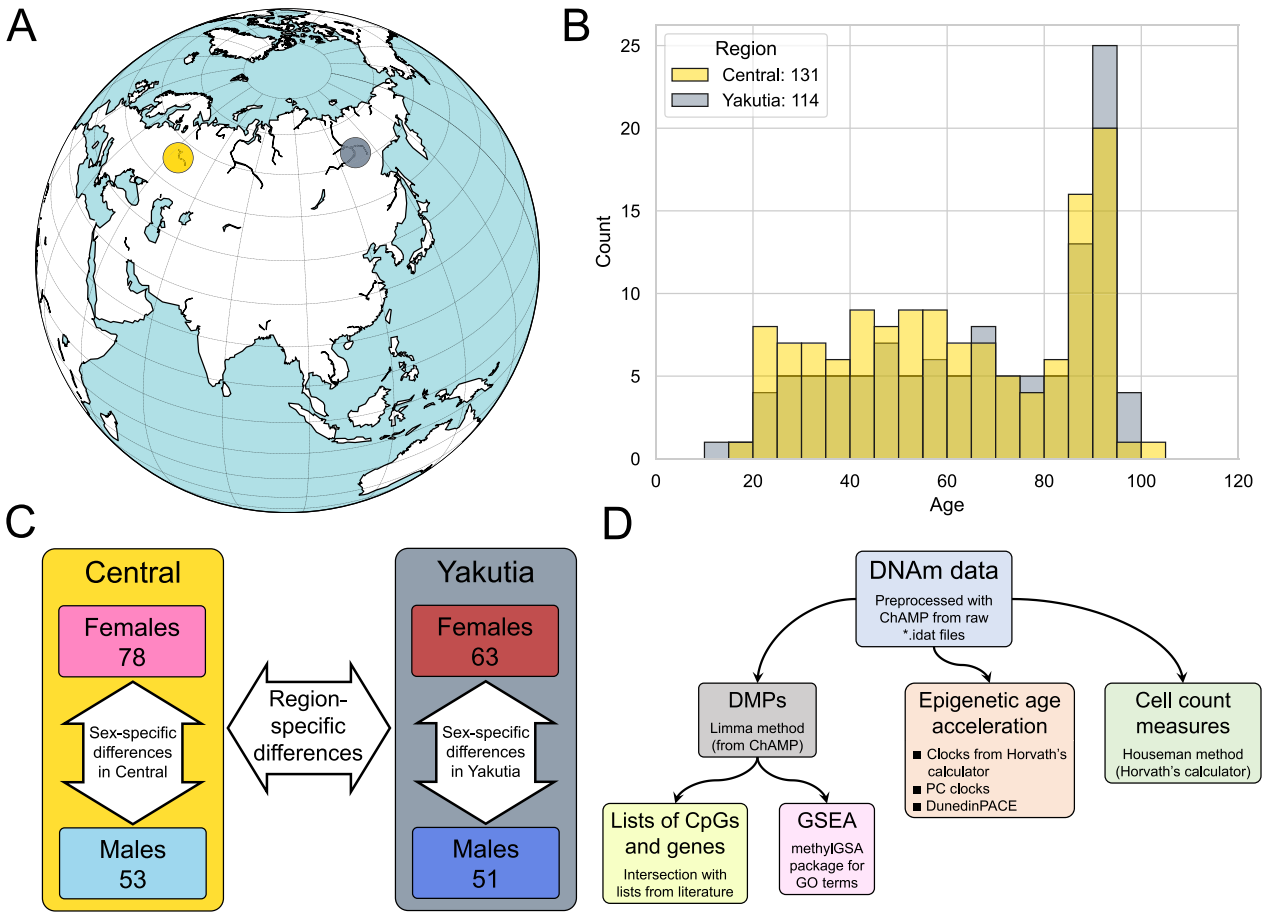

Fig. 1 Participants and study design. **A** The globe with the highlighted spots where participants were recruited for the study. Central region is highlighted in yellow; Yakutia is highlighted in gray. **B** Histogram of the age distribution in the two regions. **C** The scheme of EWAS experiments in this study: first, epigenetic differences between regions are studied, then sex-specific differences in both regions are sought and compared between the regions. **D** The basic workflow that the EWAS experiments follow. It involves searching for DMPs—CpGs that are statistically different between the phenotypes in question—which are used for Gene Set Enrichment Analysis (GSEA) and comparison with lists of CpGs and genes from other works. In addition, the analysis of various biomarkers obtained from DNAm data (epigenetic ages, their derivatives, and blood cell count measures)

Central region, the age of participants ranged from 15 to 101 years, in the Yakutia region from 11 to 99 years.

The compared Central region of Russia and Yakutia differ significantly in their climatic conditions. The average winter temperature in Nizhny Novgorod is −13 °C (most participants from the Central Russia live in Nizhny Novgorod and Nizhny Novgorod region), while Yakutia is one of the coldest geographical regions on Earth with a permanent population, and the average winter temperature in Yakutsk reaches −42 °C. The duration of the negative temperature period is also different: from November to March in Nizhny Novgorod and from October to April in Yakutsk. The difference in average temperatures between the warmest and coldest periods is extremely high in Yakutsk, and it reaches 70 °C, while in Nizhny Novgorod it is about twice less [49].

Nizhny Novgorod region and the Republic of Sakha (Yakutia) differ significantly in demographics. The average age of the population is 35.0 years in Yakutia and 42.9 years in Nizhny Novgorod region. There are also significant differences between the sexes: The average age of men in Yakutia is 33.2 years, in Nizhny Novgorod—39.7 years; the average age of women in Yakutia is 36.6 years, in Nizhny Novgorod—45.5 years. Life expectancy at birth in the Republic of Sakha in 1990 was 66.2 years, which is lower than life expectancy in Nizhny Novgorod, which was 69.8 years. However, recently there has been a tendency to reduce this gap in life expectancy, which may be due to many reasons, in particular, migration, changes in the mortality rate from natural and unnatural causes, and many others. It is worth noting that life expectancy of men and women differs significantly in both regions. Male life expectancy at birth in Yakutia in 1990 is 61.4 years (65.65 years in 2021), in Nizhny Novgorod—64.0 years (63.81 years in 2021); female life expectancy at birth in Yakutia in 2021 is 71.4 years (74.47 years in 2021), in Nizhny Novgorod—75.0 years (73.97 years in 2021) [50].

DNA methylation data for 245 samples were collected using Illumina Infinium MethylationEPIC BeadChip technology that measures DNA methylation levels from a total number of 866,836 genomic sites with single-nucleotide resolution. After all preprocessing procedures, 739,168 CpG sites remained.

EWAS was performed in three different settings (Fig. 1C): EWAS between two regions, Central Russia and Yakutia (Sect. "Region-specific differences"), as well as two EWAS investigating sex differences in these two regions (Sect. "Sex-specific differences in regions"). The results of these two sex-specific studies were also then compared with each other.

Each EWAS study followed the workflow shown in Fig. 1D. Differentially methylated positions (DMPs) analysis allows to perform a contrast comparison between two phenotypes for the considered covariate (in our case, region or sex, depending on the task). The top of the most significant CpG sites and corresponding genes were considered: Enrichment of chromosomes, genomic regions, and CpG islands were studied, and overlaps with similar lists of specific CpG sites and genes from published works were examined. The resulting FDR-corrected [51] p values of the statistical test for the difference in methylation levels between the two considered groups for all CpGs were used to perform Gene Set Enrichment Analysis (GSEA). The resulting significant terms from the Gene Ontology (GO) library [52, 53] were then analyzed and interpreted from a biological perspective.

Next, methylation data were uploaded to Horvath's online calculator [54]—it allowed to obtain values for the 4 most common estimators of epigenetic age: Horvath DNAm age [37], Hannum DNAm age [36], DNAm PhenoAge [39], and GrimAge [38] and also blood cell count measures for the following cell types: CD8T, CD4T, NK, B cells, monocytes, granulocytes, according to Houseman algorithm [55]. In addition to the classical versions of epigenetic clocks, we considered their retrained with principal components variations, which significantly increase the reliability of classical epigenetic models [56]. Another epigenetic biomarker used in the work, which characterizes the rate of aging, is DunedinPACE—estimation of the rate of aging by DNA methylation. This metric can be interpreted as the number of biological years per chronological year (with a mean value of 1) [57].

Region-specific differences

In this section, epigenetic differences between the two regions—Central Russia (131 samples) and Yakutia (114 samples)—are analyzed according to the workflow shown in Fig. 1C.

Region-specific CpGs, DMPs, and GSEA

To investigate the epigenome-wide differences between the two regions, we first applied the limma method [58] to search for region-specific CpGs (DMPs). Figure 2 shows the results of this algorithm.

The Manhattan plot (Fig. 2A) illustrates the ordering by chromosomal position of CpG sites with different statistical significance (FDR-corrected p value) in the methylation level between the two regions. Volcano plot (Fig. 2B) in addition to statistical significance contains absolute value of logarithmic fold change, which indicates relative methylation level: Yellow dots with negative log2(FoldChange) values correspond to CpG sites hypermethylated in Central region relative to Yakutia (Fig. 2C), whereas gray dots with positive log2(FoldChange) values are hypermethylated in Yakutia region relative to

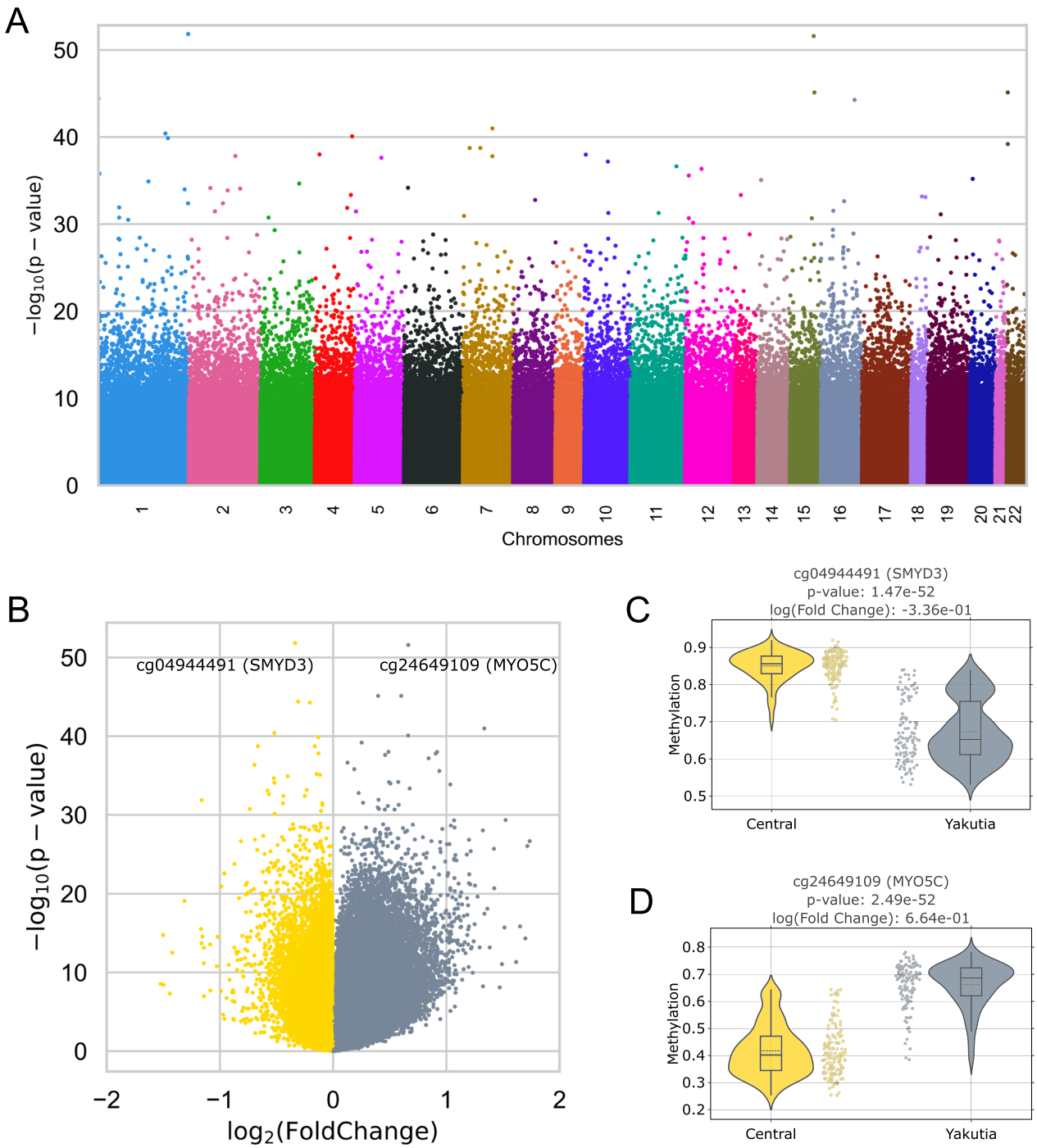

Fig. 2 DMP analysis (limma) of regional specificity. **A** Manhattan plot for the distribution of corrected *p* values of region-specific CpGs distributed by location in chromosomes; **B** Volcano plot of limma results for all CpGs. Hypermethylated CpGs in the Central region are highlighted in yellow, and hypermethylated CpGs in Yakutia are highlighted in gray. The two most significant CpGs are highlighted in the plot; the distributions of methylation levels between regions are shown in **C** cg04944491 (SMYD3 gene) and **D** cg24649109 (MYO5C gene). The solid line on the boxplot corresponds to the median value; the dotted line corresponds to the mean value

Central region (Fig. 2D). SMYD3 gene (Fig. 2C) is highly expressed in platelets and testis, moderately expressed in CD8+T cells and plays a crucial role in carcinogenesis and tumorigenesis [59, 60]. MYO5C gene (Fig. 2D) is associated with adiposity [61, 62] and serum C-reactive protein levels [63, 64].

We identified the Top-1000 CpG sites with the most statistically different methylation levels between regions with the lowest p values (Additional file 1: Table S2) and considered the distribution and enrichment analysis of these CpGs for chromosomes, genomic regions, and CpG islands (Fig. 3).

These region-specific probes were statistically significantly overrepresented only in chromosome 16, and no statistically significant features were detected in all other chromosomes. Overrepresentation of such CpGs was also found in the regulatory genomic region TSS1500 and underrepresentation in the gene body. It is noteworthy that the selected Top-1000 CpGs belong mainly to islands, regions with increased CpG density (overrepresentation with p value $< 1e-25$), and CpGs in OpenSea (regions with the lowest density) are statistically underrepresented.

It is interesting that chromosome 16 contains numerous erythroid-specific alpha-globin genes, and hypermethylation of certain differentially methylated regions at different stages of erythropoiesis has been shown [65]. In the context of differences between the considered regions, it would be interesting to consider the relationship between red blood cell count (RBC) and adaptation to cold temperatures; however, very few such studies have been performed. In [66], it was shown that exposure to extremely low temperatures for 50 days can lead to a decrease in RBC.

It is interesting to compare the obtained Top-1000 CpGs, corresponding to 656 genes, with the list of 791 genes from the work Cardona et al. [9], analyzing single-nucleotide polymorphism data for 10 indigenous Siberian populations, thereby describing the genetic influence/impact on epigenetic variation. These 791 genes are associated with biological processes and pathways hypothesized to be involved in cold adaptation and related to basal metabolic rate, non-shivering thermogenesis, response to temperature, smooth muscle contraction, blood pressure, and energy metabolism. The intersection of the two lists (our list and list from [9]) results in 33 genes (Additional file 1: Table S3), which, according to [9], is associated with blood pressure (PDGFB, DDAH1, POMC, REN, SGK1, F2R, NAV2, LRP5, PLCB3), basal metabolic rate (PIK3CD, TG, CCNA1, STK11, PAX8, CDKN1A, CACNA1A), energy metabolism (STK11, DPP4, PLCB3, CACNA1A, GATA4, STX1A, GNB2), response to temperature (ADM, XYLT1, DNAJB6, TRPM2, IL6, MYOF, DNAJC7), smooth muscle contraction (ADM, F2R, PTPRM, KCNMA1, P2RY2), non-shivering thermogenesis (EPAS1, PRDM7, PARK2). Let us further explore more functions of some genes from the intersection.

PIK3CD gene is involved in the development and migration of natural killer cells to inflammation foci and is involved in NK cell receptor activation [67]. Another gene, IL6, is also found to be associated with the findings of differences in cellular composition between the regions: It is responsible for the differentiation of CD4T cells and takes part in the initiation of the immune response [68]. TG gene is responsible for thyroid hormone thyroglobulin production [69], and PAX8 gene is also associated with thyroid function and thyroid-specific gene expression [70]. As shown in [9], the PDGFB gene may be associated with blood pressure regulation. In the context of the difference between the considered regions, the appearance of this gene can be explained by the influence of the ambient temperature on blood pressure, the increase in which is associated with lower temperatures

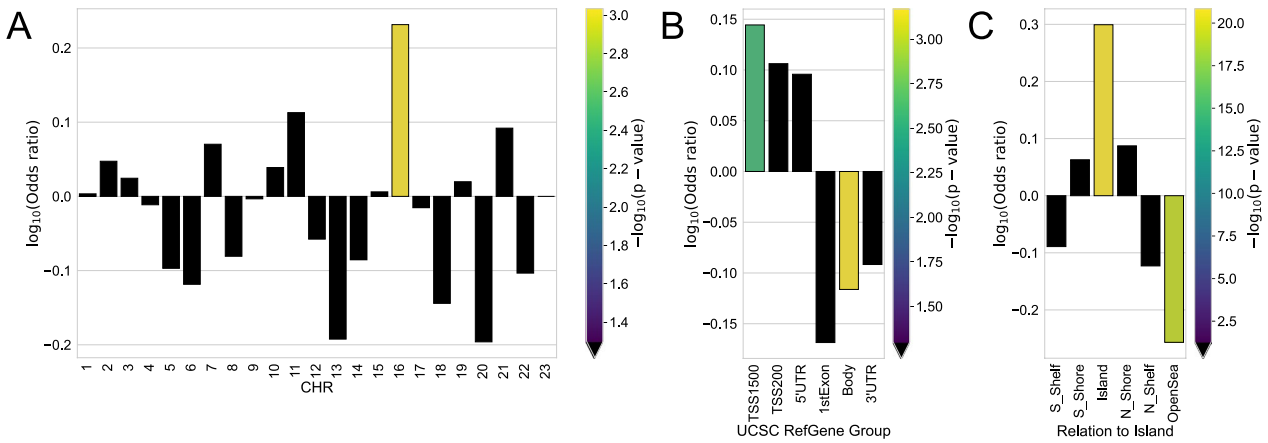

Fig. 3 Enrichment of **A** chromosomes, **B** genomic region, and **C** relation to CpG island for Top-1000 region-specific CpGs. Odds ratio values and corresponding p values (shown by color) were obtained from Fisher exact test. Black color indicates the absence of statistical significance (p value > 0.05)

[71]. This may also include the ADM gene, whose activity has also been associated with the production of hypotensive and vasodilator agents [72], as well as the REN gene, which initiates a response cascade to elevate blood pressure [73]. The found gene GATA4, which is included in the regulation of cardiac-specific gene expression, is also associated with cardiac development [74, 75]. Interestingly, the PRDM7 gene itself is presumably associated with epigenetic regulation of gene expression [76].

Indigenous Siberian populations were also studied in [77, 78]. However, there is no overlap between our list of genes and the genes from these works. The reason for this may be that [77, 78] considered a trio of Yakuts-Han Chinese-Europeans, whose differences in adaptation to climate and/or nutritional aspects are much higher than those between Yakutia and Central Russia.

For further GSEA analysis, we used the adaptive methylGSA method [79], which does not use pre-selected lists of CpG sites and genes according to some threshold value, but works with all available CpG sites and their p values. As a result, methylGSA identified 17 statistically significant terms (with adjusted p value < 0.05) from the Gene Ontology library [52, 53] (Additional file 1: Table S4). Among the found terms, we can highlight one related to actin filament activity, which regulate cellular behavior and are involved in the aging process and age-associated diseases [80, 81]. Another term related to drinking behavior reflects the mode of water consumption, and the difference between the regions may be related to water supply, water pretreatment, and sanitation problems in Yakutia [82]. The group of terms related to channel activity is especially interesting in the context of differences between Yakutia and Central Russia. The sensation of low temperature through the skin sensory nerve terminals and the propagation of action potentials in cold-sensitive nerve fibers affects a large number of ion channels [83], and thus, they are closely related to cold susceptibility [84, 85] and may differ in populations living in significantly different climates. The presence of terms concerning the regulation of steroid and corticosteroid hormone secretion is also legitimate, since they are related to endocrine function regulating many aspects of human life, likely playing a role in the developmental processes that lead up to age-specific early life-history transition and thus also in the aging process [86, 87] and adaptation to the cold exposure [88], which are most interesting in the context of our study. Cell fate commitment related to metabolism has also been shown to be an important term in the aging process [89, 90]. Nutritional differences between regions and the increased risk of digestive system diseases in the Yakutian population [91] may be the reason for identifying a term related to digestive system regulation.

It is interesting that among the found terms there are not only those related to adaptation to climatic conditions (water consumption, digestive system regulation), but also those related to aging processes (actin filament activity, cell fate), and both of them (channel activity, regulation of steroid and corticosteroid hormone secretion). Therefore, we further pay attention to aging analysis in the considered cohorts.

Epigenetic age accelerations

To investigate the differences in aging processes between the two regions, we considered 9 types of epigenetic ages: the 4 most common clocks from Horvath's online calculator [54] (Horvath DNAm age [37], Hannum DNAm age [36], DNAm PhenoAge [39], and GrimAge [38]) and 5 of their PC improvements [56]: PCHorvath1, PCHorvath2, PCHannum, PCPhenoAge, PCGrimAge.

We investigated the correlation of all epigenetic ages with chronological age in the two regions, and it was found that the correlation coefficient is close to 1 in all cases (Fig. 4A). Scatter plots illustrating dependence between chronological age and epigenetic ages reveal that Hannum and GrimAge demonstrate underestimation in older participants, while PCGrimAge demonstrates overestimation in younger participants (Fig. 4A).

Region-specific age acceleration was determined as follows: First, a linear regression was built between the different epigenetic ages and the chronological age for the Central region group. Region-specific age acceleration values were defined as residuals relative to this linear approximation. Thus, it follows from the definition that in the Central region group, the average age acceleration value is 0, and a nonzero value in the Yakutia group allows us to conclude about accelerated aging in this group relative to the Central region (region specificity of age acceleration). In addition, the DunedinPACE value also allows us to make conclusions about the differences in age acceleration between regions, since this metric characterizes the pace of aging [57]. Figure 4B shows the distributions of region-specific age acceleration values for all considered types of epigenetic clocks with FDR-corrected [51] p values of Mann–Whitney U Test [92], which reflect the measure of statistical significance of the found differences. Figure 4C shows the DunedinPACE values with the corresponding p value.

For all types of epigenetic ages, we observe statistically significant (p value < 0.05) positive region-specific age acceleration in Yakutia relative to the Central region. Horvath DNAm age shows the lowest p value, and the median acceleration relative to the Central region is 5.36 years. For other epigenetic ages, the median acceleration does not exceed 3 years. It is also interesting that the average DunedinPACE value

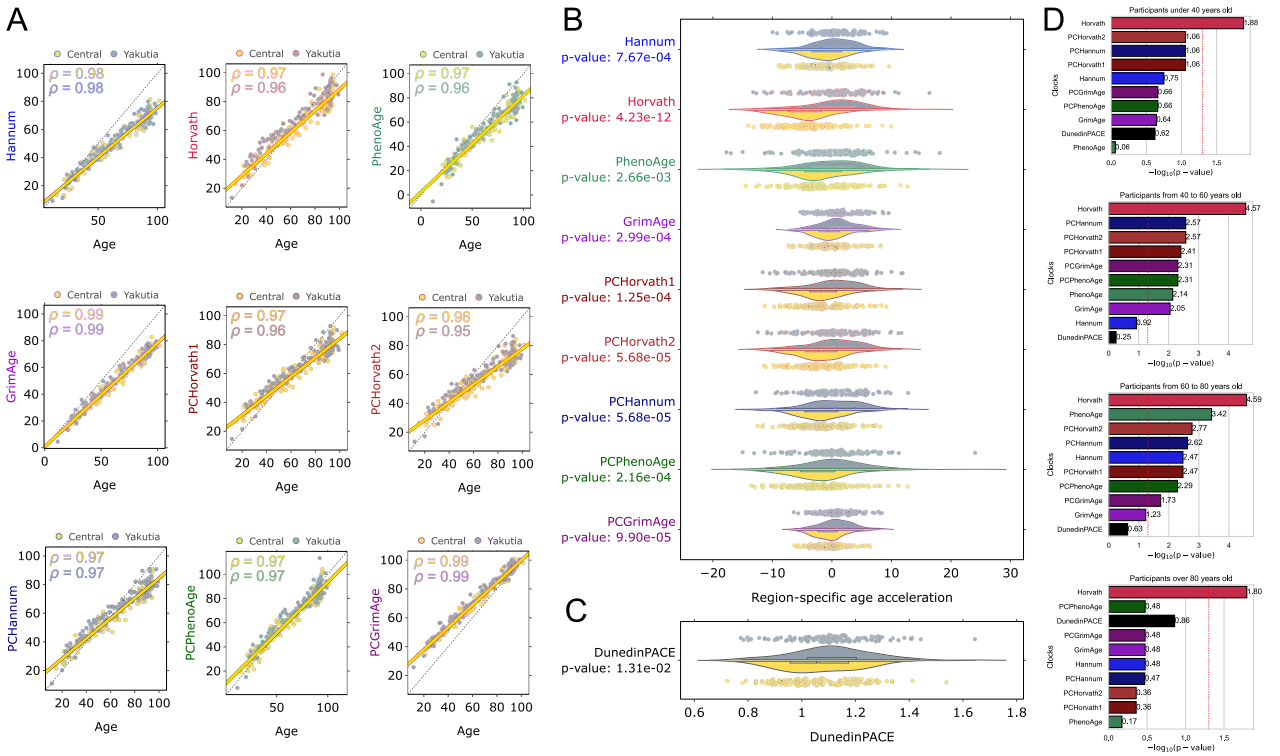

Fig. 4 Region-specific age acceleration in Central and Yakutia regions for different epigenetic ages: Hannum, Horvath, PhenoAge, GrimAge, PCHorvath1, PCHorvath2, PCHannum, PCPhenoAge, PCGrimAge. **A** Scatter plots demonstrate dependence between chronological age and epigenetic ages for Central (yellow) and Yakutia (gray). The black dotted line y = x in each scatter plot corresponds to the equality of the plotted ages; the yellow bold line represents the regression plotted on samples from the Central region. Pearson correlation coefficients between the corresponding epigenetic age and chronological age are given for Central (yellow) and Yakutia (gray). **B** Violin plots show region-specific age acceleration for different epigenetic ages. Region-specific age acceleration values were determined as residuals from a linear regression model constructed on samples from the Central region group (yellow bold lines in Fig. 4A). Mann–Whitney U test was applied to analyze the statistically significant difference in epigenetic age acceleration between groups. The obtained *p* values were FDR-corrected. **C** DunedinPACE values with the *p* values of Mann–Whitney U test. **D** Bar plots illustrate *p* values of Mann–Whitney U test analyzing the statistically significant difference in epigenetic age acceleration between regions in different age groups (under 40 years, 40–60 years, 60–80 years, over 80 years). The red dotted line corresponds to the significance level of *p* value = 0.05

(Fig. 4C) exceeds 1 for both regions, which may indicate that the Russian population ages faster overall than the Dunedin Study participants, whose data were used to develop this metric [57]. However, the same effect for DunedinPACE that was found for all epigenetic clocks persists: In the Yakutia group, the region-specific age acceleration is statistically higher with respect to the Central region (median DunedinPACE is 1.11 in Yakutia, 1.05 in Central Russia). Region-specific age acceleration in the Yakutia group may also be related to the fact that the original epigenetic clock models were built on data for different populations. To train the Horvath model, 39 datasets were used including participants from different African, Asian, Hispanic, and Caucasian populations. The Hannum model used participants of Caucasian and Hispanic origin. Participants from NHANES for the PhenoAge model were of African American and Caucasian ancestry. GrimAge was built using Caucasian participants from the Framingham heart study Offspring Cohort. For PC clock modifications, data of participants from the London Life Sciences Prospective Population study, of South Asian origin, were used.

Analysis of the statistical significance of the difference in region-specific age acceleration for different age groups shows that the lowest p value for all groups is demonstrated by the Horvath clock (Fig. 4D). Interestingly, for younger participants (under 40 years) and for older participants (after 80 years), only Horvath age acceleration is statistically significant between the regions. The difference in DunedinPACE values is not statistically significant between regions in any of the age groups, GrimAge acceleration is statistically significant only in the 40–60 years group, and Hannum age acceleration is statistically significant only in the 60–80 years group.

Blood cell counts estimation differences

Horvath's online calculator [54] allows to obtain the distributions of blood cell count measures: CD8T, CD4T, NK, B cells, Monocytes, and Granulocytes using Houseman algorithm [55]. The quantitative composition of blood cell populations can vary by race/ethnicity. For example, it has been shown that there is a difference in the levels of naïve CD8+T cells and naïve CD4+T cells, as well as different estimated proportions of neutrophils, B cells, and natural killer cells in different races [20, 31].

Figure 5 shows the distribution of blood cell counts for individuals from the Central Russia and Yakutia regions with FDR-corrected [51] p values of Mann–Whitney U Test [92].

Statistically significant increased numbers of CD8T and NK are observed in the Yakutia region relative to Central. In contrast, the values of CD4T and Monocytes are statistically lower in Yakutia than in the Central region. CD8T cells are responsible for the response to the impact of external pathogens unfamiliar to the body [93], their increased number may be associated with an increased viral load in Yakuts compared to the residents of Central Russia. This is consistent with an increase in natural killer cells, which, along with their main function, can produce high levels of cytokines [93]. A decrease in the CD8T cell count is observed with age [94], which can be associated with underestimation of Hannum age in the elderly (Fig. 4A, top left). Differences in CD4T levels, as well as in CD8T and NK levels, may also be due to ethnic differences [95]. For monocytes, European ancestry has been shown to have higher levels of monocytes compared to other populations [96].

Sex-specific differences in regions

This section examines epigenetic differences between males and females independently in the two regions and then compares the results to highlight sex-specific differences between the regions. In the Central region, the sex distribution among participants was 78 females and 53 males, and in Yakutia, 63 females and 51 males (Fig. 1B). The sex-specific analyses in each region follow the workflow shown in Fig. 1C.

Sex-specific CpGs, DMPs, and GSEA

Further, following the general pipeline (Fig. 1C), we studied epigenome-wide differences between males and females in both regions using limma [58]. The resulting distribution of p values of sex-specific CpGs in both regions (Fig. 6A–C for Central region and Fig. 6D–F for Yakutia) differs significantly from the similar distribution in the region-specific task (Fig. 2A, B): The number of statistically significant CpG sites in both sex-specific tasks is significantly lower than in the region-specific task. However, we should note the similarity in the distributions of

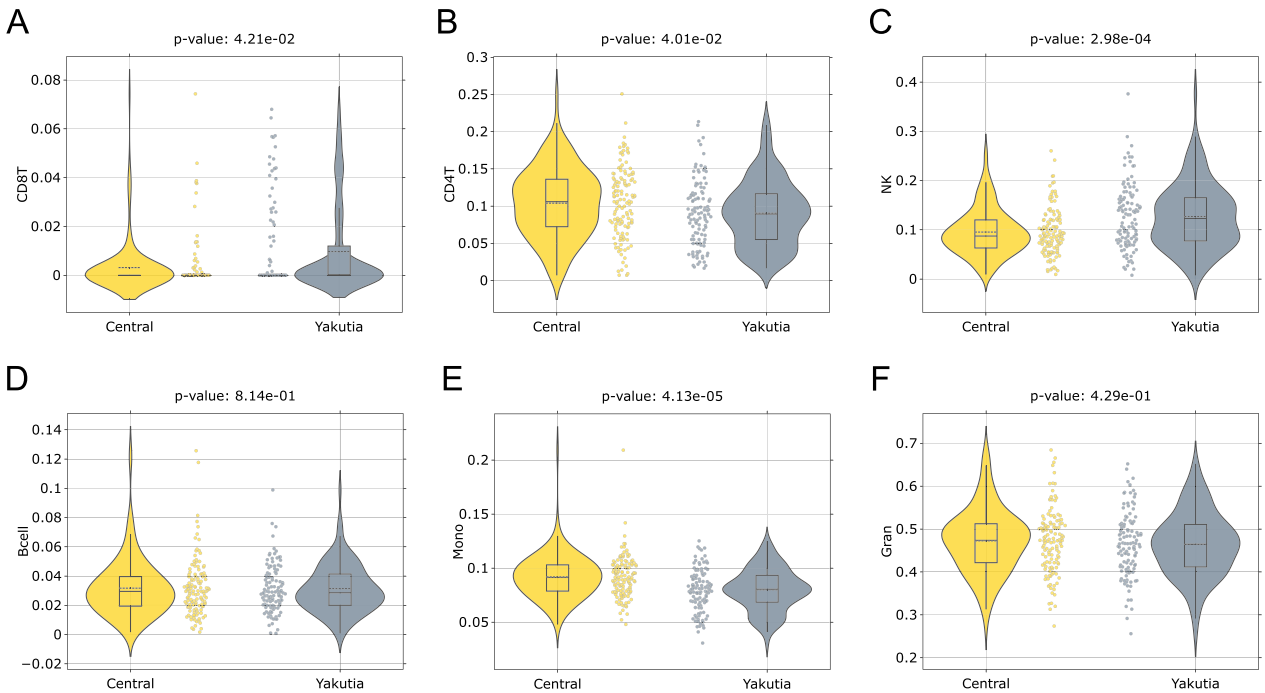

Fig. 5 Blood cell count distribution in Central and Yakutia regions: **A** CD8T, **B** CD4T, **C** NK, **D** B cells, **E** Monocytes, **F** Granulocytes. Mann–Whitney U test was applied to analyze the statistically significant difference in distribution between the groups. The resulting p values were FDR-corrected. The solid line on each boxplot corresponds to the median value, dashed line—to mean value

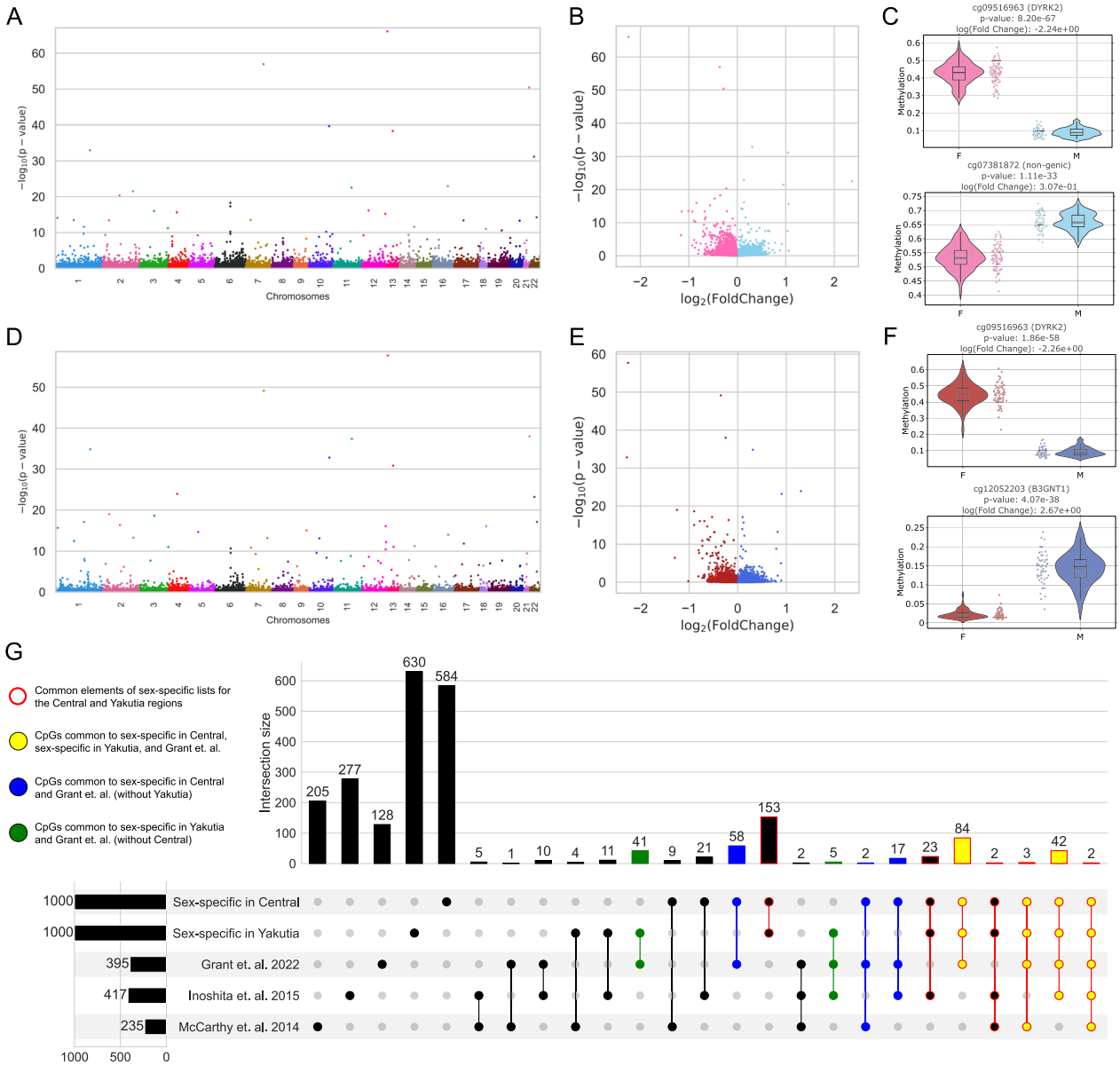

Fig. 6 DMP analysis of sex specificity in the Central and Yakutia regions. **A** Manhattan plot for the distribution of adjusted *p* values of sex-specific CpGs distributed by location in chromosomes for the Central region. **B** Volcano plot of limma results for all CpGs in the Central region, hypermethylated in females CpGs are highlighted in pink and hypermethylated in males CpGs are highlighted in light blue. **C** Examples of the sex-specific DMPs distribution for the Central region: top—hypermethylated in females; bottom—hypermethylated in males. **D** Manhattan plot for the distribution of adjusted *p* values of sex-specific CpGs distributed by location in chromosomes for the Yakutia region. **E** Volcano plot of limma results for all CpGs in the Yakutia region, hypermethylated in females CpGs are highlighted in red and hypermethylated in males CpGs are highlighted in dark blue. **F** Examples of the sex-specific DMPs distribution for the Yakutia region: top—hypermethylated in females; bottom—hypermethylated in males. **G** UpSet plot showing the intersections of the Top-1000 most statistically significant sex-specific CpGs in the Central and Yakutia regions with the lists of sex-specific CpGs from other published studies. Each column corresponds to the intersection of all lists. The sum of all elements in a row is the total value of the elements in the selected CpG list. (This number is given on the left bar plot: 395 for Grant et al. [44], 417 for Inoshita, et al. [97], 235 for McCarthy et al. [98], 1000 for both our sex-specific CpG lists.) All subsets that include common elements of sex-specific lists for the Central and Yakutia regions from this study are highlighted in red. Yellow indicates subsets of CpGs common to the three sets: sex-specific in Central, sex-specific in Yakutia, and from Grant et al. [44]. Blue indicates subsets of CpGs common to the two sets: sex-specific in Central and Grant et al. [44] (without Yakutia). Green indicates subsets of CpGs common to the two sets: sex-specific in Yakutia and Grant et al. [44] (without Central)

sex-specific CpGs between the two regions, which can be observed especially clearly in the *p* value and fold change ranges in Fig. 6B, E. Examples of the distributions of hypo- and hypermethylated sex-specific CpGs for both regions are shown in Fig. 6C, F.

Selecting the Top-1000 sex-specific CpGs (Additional file 1: Table S5) with the lowest *p* values in both regions, it appears that these lists overlap by about a third (309 CpGs, which are the sum of all bars with red highlighting in Fig. 6G). In addition, we considered the intersections with the lists of sex-specific CpGs from other papers, which also investigated various aspects of sex specificity in epigenetics. One of the first meta-analyses of autosomal chromosome DNA methylation to identify sex-specific CpG sites was performed in McCarthy et al. [98]. 235 CpG sites were detected after correction for multiple testing. Using full-genome DNA methylation profiling, Inoshita et al. [97] investigated the effect of sex using multiple linear regression analysis corrected for age and inferred blood cell proportions. 417 CpG sites showed significant gender differences in DNA methylation. One of the most recent studies of sex-specific DNA methylation patterns by Grant et al. [44] used data from the Illumina EPIC standard and identified 395 sex-associated CpG sites.

A detailed representation of all possible intersections of the lists of sex-specific CpGs is illustrated in Fig. 6G. The largest overlap between the two lists proposed in this paper for Central and Yakutia regions was found with the chronologically most recent work of Grant et al. [44]. All possible subsets involving all 3 lists, Central, Yakutia, and Grant et al., are highlighted in yellow (131 CpGs in total). If we consider the lists of the two regions separately, the Central region overlapping with Grant et al. gives 208 CpGs (sum of yellow and blue bars in Fig. 6G), and for Yakutia region—177 CpGs (sum of yellow and green bars in Fig. 6G). Thus, the lists of both regions include almost half of all sex-specific CpGs from the work [44]. CpG sites located in noteworthy genes according to [44] were found in the intersection. In particular, CpGs located in the DDX43 gene, which are involved in spermatogenesis and male fertility [98], or CpG located in the GABPA gene, are associated with early onset Alzheimer's disease, Parkinson's disease, and breast cancer [99]. All the overlaps with other works turned out to be relatively smaller and are also presented in Fig. 6G.

The enrichment analysis of the Top-1000 sex-specific CpGs for the both regions was also performed (Fig. 7). CpGs for the Central region were overrepresented in the 6th chromosome, and for Yakutia in the 4th chromosome. Interestingly, the work [44] also observed the largest number of sex-specific CpGs in chromosome 6. (This work considers the population from the UK, which is just more similar in terms of ethnicity to Central Russia than to Yakutia.) Statistically significant underrepresentation in the gene body is observed in both lists. Also, for both lists there is overrepresentation in high-density CpG islands and shores and underrepresentation in low-density OpenSea regions.

Further GSEA, performed independently for both regions, revealed 40 terms in the GO library for the Central region and 39 terms for Yakutia (Additional file 1: Table S6). Moreover, the lists of terms for the two regions are almost identical: 39 common terms and only 1 specific for the Central region (GO:0004713 protein tyrosine kinase activity). Among the common terms appear ones related to fat cell proliferation, whose differences between the sexes have been well studied [100, 101]. Next, a whole group of terms related to the processes of glycogen, glucose, and carbohydrate metabolism is worth mentioning. The processes associated with glycogen metabolism during exercise differ significantly in men and women, with less muscle glycogen being depleted in women [102, 103]. In the resulting list, 6 of 40 terms are related to glycogen. The processes of glucose metabolism are also related to it. In women, glucose metabolism is affected by the menstrual phase, and glucose production in the liver is lower and glucose appearance rates are higher [103]. Because of this, the incidence of different types of diabetes differs in men and women [104]. Terms related to carbohydrate processes contribute to 10% of the total number of sex-specific terms. Women tend to oxidize less total carbohydrates than men in response to physical activity [102, 105]. The term manganese ion binding may reflect the fact that men in general absorb less manganese than women [106].

Epigenetic metrics: age accelerations and blood cell counts

Similar to the analysis of between-region differences (Sect. "Region-specific differences"), 9 types of epigenetic ages (4 types of classical clocks [36–39] and their PC variations [56]), DunedinPACE values [57], and blood cell counts measures [55] were considered.

Sex-specific age acceleration within each region was determined as follows: A linear regression was built between chronological age and corresponding epigenetic age on the female group only, and the sex-specific age acceleration values were derived from these linear models. As a result, in both regions, the mean acceleration value for the female group will be 0 (by construction), while for the male group a nonzero value will be obtained, based on which we can conclude about sex-specific age acceleration in the two regions. The difference in the distribution of sex-specific age acceleration values between males and females was tested using the Mann–Whitney U Test with FDR-corrected

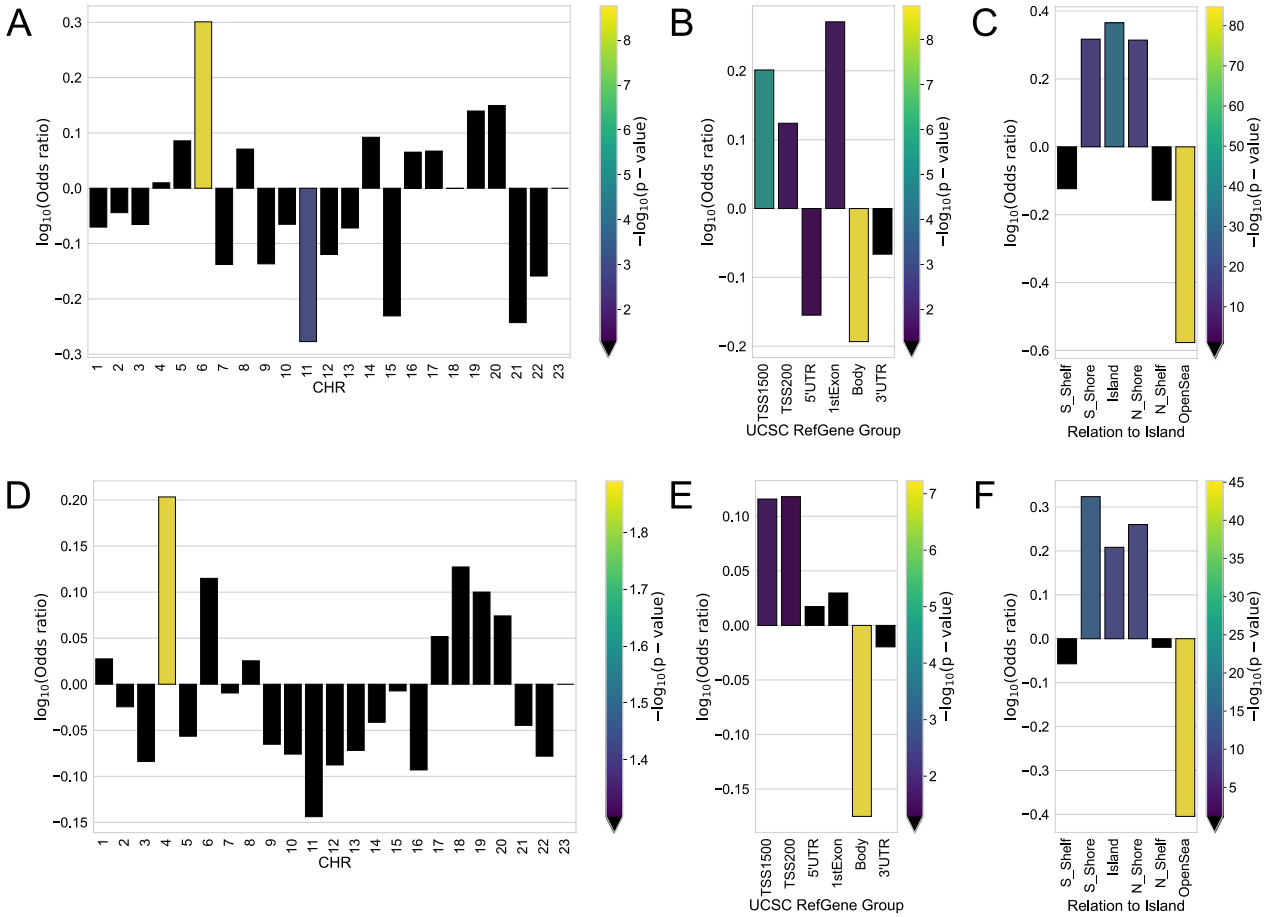

Fig. 7 Enrichment for the Top-1000 sex-specific CpGs in Central region of: **A** chromosomes, **B** genomic region, **C** relation to CpG island; in Yakutia of: **D** chromosomes, **E** genomic region, **F** relation to CpG island. Odds ratio values and corresponding *p* values (shown by color) were obtained from Fisher exact test. Black color indicates the absence of statistical significance (*p* value > 0.05)

p values < 0.05. An additional characteristic that allows us to make conclusions about sex-specific age acceleration in the two regions is DunedinPACE with the corresponding *p* value. Figure 8 shows the summary of statistically significantly different epigenetic biomarkers between males and females in both regions.

In the Central region, a positive statistically significant sex-specific age acceleration in the male group was found in 5 of the 9 types of epigenetic ages, while in Yakutia it was found only for 2 types, and they are GrimAge and its PC modification, which are also statistically significant in the Central region. Several studies have shown that DNAm GrimAge shows more accurate results compared to other models, as well as confirming the thesis of slower aging rates in females compared to males [107–109]. The distribution of DunedinPACE values is not statistically different between males and females in different regions, nor are blood cell counts estimations.

Discussion

Conclusion

This work is the first epigenetic study focused on the Yakutian population that try to elucidate different dimensions of the variability observed addressing: (1) epigenetic features of the adaptation mechanisms to the cold; (2) pattern of age acceleration and aging rate in a population characterized by extreme environmental conditions; and (3) sex-specific methylation patterns.

In this work, we examined the epigenetic features of the 114 individuals from the Yakutian population in comparison with 131 participants from Central Russia. The Yakutian population live in the extreme environment and experience an extended period of severe cold temperatures, dramatic variation in photoperiod, and variability in food resources. Participants from Central Russia live in a milder environment, with higher temperatures in a shorter winter period compared to Yakutia. The considered populations also differ in terms of genetic history:

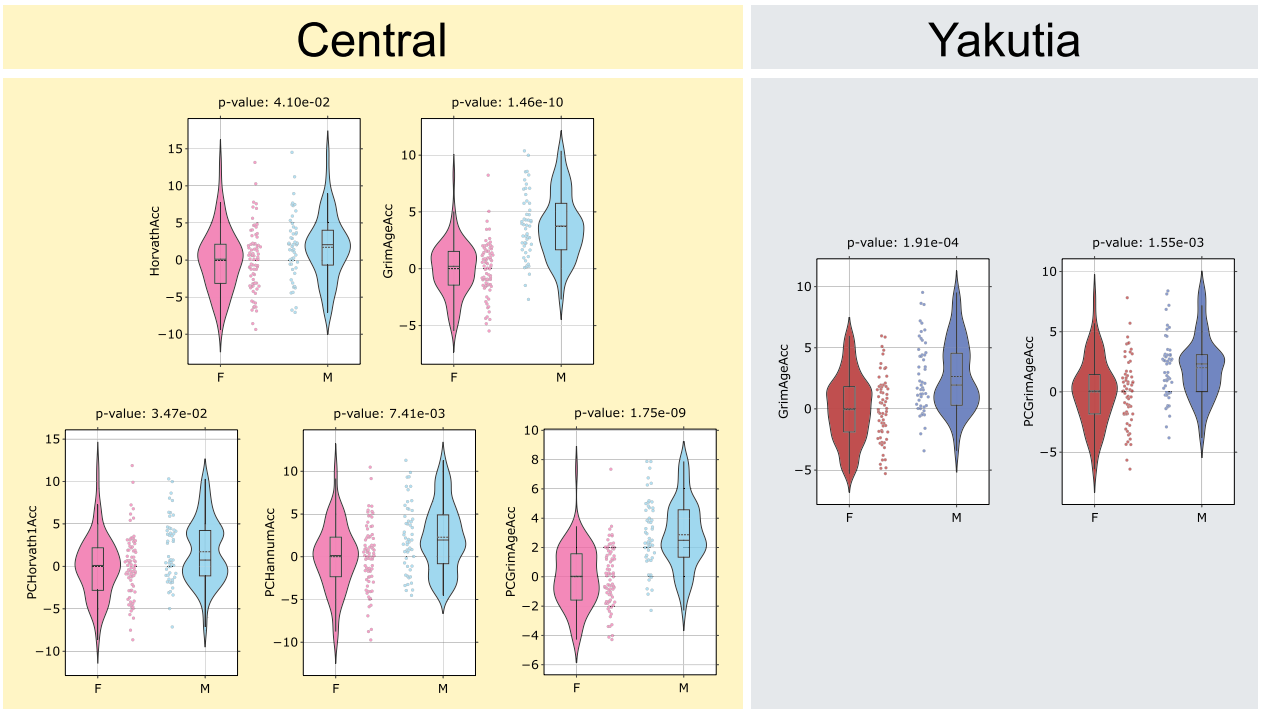

Fig. 8 Summary of statistically significantly different epigenetic biomarkers between males and females in the Central region (left yellow column) and Yakutia (right gray column). Mann–Whitney U test was applied to analyze the statistically significant difference in the distribution between the groups separately for all epigenetic ages. The obtained *p* values were FDR-corrected. Only significant biomarkers are shown. The solid line in each diagram corresponds to the median value, the dotted line to the mean value

Representatives of Central Russia are close to the peoples of Northeast Europe, while Yakutia shows genetic components that characterized Siberian populations and East Asian populations.

First, we focused specifically on region-specific differences between the considered regions, and next, we studied sex-specific differences in both regions and compared them.

We selected 1000 most significant region-specific CpGs, for which enrichment analysis showed statistically significant overrepresentation in chromosome 16, in the regulatory genomic region TSS1500 and in CpG islands. We intersected the list of genes we obtained from the list of the most statistically significantly different CpGs with the list from Cardona et al. [9], a notable work analyzing single-nucleotide polymorphism data from Siberian populations, including Yakutian, and identifying genes associated with biological processes hypothetically associated with adaptation to the cold. 33 genes appeared in the intersection, most of which are associated with blood pressure, basal metabolic rate, energy metabolism, and response to temperature, suggesting that part of the epigenetic variability observed could be ascribable to genetic background that characterized this population.

Among the overlapping genes, there are also those whose functions are closely related to immune responses and to the considered blood cells, such as NK cells and CD4T cells (PIK3CD and IL6 genes). Interesting correlation is found between the TG gene responsible for thyroid hormone thyroglobulin production [69] and the terms related to endocrine system functioning and hormone production found by GSEA. Glucocorticoids regulate thyroid function, thereby influencing the production of thyroglobulin [110]. The KCNMA1 gene, which is associated with the regulation of the contraction of smooth muscle, also has channel activity among its functions, which appeared among important GO terms [111, 112]. CACNA1A [113] and STX1A [114] genes are also associated with channel activity.

GSEA of statistically significantly different CpGs by methylation levels between the two regions revealed 17 Gene Ontology terms. Among the found terms, we can highlight the ones related to actin filament activity, channel activity, regulation of steroid, and corticosteroid hormone secretion, as well as a term related to regulation of the digestive system. Among the found terms, there are those related to adaptation processes to the cold climate, to aging processes, and to both aspects.

To analyze the aging process, we considered the 4 most common epigenetic clocks (Horvath DNAm age, Hannum DNAm age, DNAm PhenoAge, and GrimAge) and their 5 PC modifications (PCHorvath1, PCHorvath2, PCHannum, PCPhenoAge, PCGrimAge). All these metrics were highly correlated with each other. When comparing the region-specific age acceleration of the studied models and DunedinPACE aging rate, we found that in Yakutia, for all considered epigenetic ages, there is a statistically significant positive region-specific age acceleration compared to Central region representatives. At the same time, the average DunedinPACE aging rate exceeds 1 in both regions. (In other words, the Russian population as a whole is aging accelerated, with more than one biological year per chronological year.) All considered epigenetic models are based on data from participants of different populations mainly of European ancestry (mostly industrialized ones), without including the indigenous populations, which may affect the results. Epigenetic age acceleration in Yakutia and increased aging rates in both considered populations may signal the need for increased attention to healthcare and public health problems.

Differences between regions are also observed for blood cell composition: Statistically significant increased number of CD8T and NK and statistically significant decreased number of CD4T and Monocytes are observed for the Yakutia region.

Next, we analyzed sex-specific differences in the two regions and selected 1000 most significant sex-specific CpGs in each region. We also considered the lists of sex-specific CpGs from Grant et al. [44], McCarthy et al. [98], Inoshita et al. [97] and compared them with the lists obtained for the Central and Yakutia regions. The largest overlap of our lists is obtained with [44] (almost half of the list), and CpGs corresponding to genes related to spermatogenesis and various diseases can be found in the overlap.

GSEA of statistically significantly different CpGs by methylation levels between the sexes in both regions revealed almost identical Gene Ontology terms for the Central region and Yakutia: 39 common terms and one specific term for the Central region. Among the found terms, we can highlight the ones related to the processes of glycogen, glucan, glucose, and carbohydrate metabolism, as well as fat cell proliferation and manganese ion binding. However, we did not find clear region-specific differences between the sexes in the highlighted biological functions.

As for regional differences, we next focused on the analysis of sex-specific epigenetic differences in the aging process in the two regions. We considered 9 types of epigenetic clocks (classical ones and PC modifications), DunedinPACE metric, and blood cell counts. Five out of 9 epigenetic age models showed statistically significant positive sex-specific age acceleration in men compared to women in the Central region, while in the Yakutia region only GrimAge with its PC modification showed the same result. At the same time, neither DunedinPACE nor blood cell counts had statistically significant differences between males and females in both regions.

Thus, we performed the first study of the epigenetic data of the Yakutian cohort, paying special attention to region-specific features, aging processes, age acceleration, and sex specificity. We revealed geographic region-specific differentially methylated regions associated with both, adaptation to climatic conditions and aging processes. We also showed that representatives of the Yakutia region show higher age acceleration compared to Central Russia, one of the reasons for which may be the more severe climatic conditions in which Yakuts live and the need to adapt to the cold. Other possible factors that we do not take into account in this study may include diet, pathogen load, socioeconomic conditions, and many others. However, a certain degree of age acceleration is found for Central Russia too. For both regions, we confirmed that men age faster than women (resulting in a shorter life expectancy), but no significant sex-specific difference was found between the regions.

Limitations

We would also like to address the limitations of this work. The most common limitation of many studies investigating biomedical data is sample size. Data such as DNA methylation often have small sample sizes compared to their dimensionality. Another limitation is the wide variety of methods available for performing GWAS and EWAS analyses, which take into account different background information, require a selection of thresholds, and can lead to varying results. Among several approaches to determining age acceleration, we implemented a linear regression-based method, which is recommended when the slope of the relationship between predicted age and real age differs from unity. In terms of interpretation of the obtained results, differences between regions can be caused by both genetics and environment, and it is difficult to disentangle their influence. There remains a challenge in identifying population-specific CpGs and climate adaptation associated CpGs; a deeper comparison of the Yakut population with other Asian populations may shed light on this.

Methods

Data collection

All study participants were explained the specifics of the procedure, possible inconveniences and risks. Each

participant signed an informed consent and filled out a consent for personal data processing, taking into account the principle of confidentiality (accessibility only to the research group and presentation of data in a common array). The study was approved by the local ethical committee of Nizhny Novgorod State University. All research procedures were in accordance with the 1964 Helsinki Declaration and its later amendments.

All study participants were healthy; exclusion criteria included chronic diseases in the acute stage, cancer and acute respiratory viral infections at the moment of biomaterial donation, as well as pregnancy in women.

The main criterion for inclusion of participants from Yakutia in this study was ethnicity. All participants are Yakuts (Sakha) in three generations. They were born and live on the territory of the Republic of Yakutia (Sakha), and their parents and ancestors are indigenous to Yakutia. All participants from the Central region are native Russian residents.

DNAm processing

Phenol Chloroform DNA extraction was used. DNA was quantified using the DNA Quantitation Kit Qubit dsDNA BR Assay (Thermo Fisher Scientific), and 250 ng was bisulfite-treated using the EpiMark Bisulfite Conversion Kit (NEB) with case and control samples randomly distributed across arrays. The Illumina Infinium MethylationEPIC BeadChip [115] was used according to the manufacturer's instructions. DNA methylation is expressed as β values, ranging from 0 for unmethylated to 1 representing complete methylation for each probe.

DNAm data preprocessing, normalization, and batch effect correction were performed with the standard pipeline in the ChAMP [116, 117] R package. The preprocessing was as follows: (1) Probes with a detection p value above 0.01 in at least 10% of samples were removed; (2) probes with a beadcount less than three in at least 5% of samples were removed; (3) all non-CpG probes [118], SNP-related probes [119], and multi-hit probes were removed [120]; (4) all probes located on chromosomes X and Y were filtered out. It is also worth noting that all remaining samples have less than 10% of probes with a detection p value above 0.01 and they do not need to be excluded. As a result, 739,168 CpGs remained for the analysis. Functional normalization of raw methylation data was performed using minfi [121] R package function. The ComBat method [122, 123] was used to correct for Slide and Array batch effects.

DMPs

DMPs analysis was performed using limma method [58] generalization in the ChAMP [116, 117] R package. Region and sex were used as categorical variables to perform a contrast comparison between the two phenotypes. This method provides adjusted p values [51] of the statistical test for the difference in methylation levels between the two considered groups, as well as fold change, which indicates the differences in mean values between the groups.

GSEA

Traditional approaches to testing a gene set can produce biased results because of differences in gene length, and the number of CpG sites can vary even among genes of the same length. EWAS can lead to multiple p value associations per gene, so it is necessary to consider the number of CpGs instead of gene length, and considering all p values provides an opportunity to take into account dependencies between genes in the set. MethylGSA R package [79] takes all these details into account, and therefore, we used this method. It allows testing of gene sets with length bias adjustment in DNA methylation data. Enrichment of GO annotations was calculated using the methylglm function, which takes p values of each CpG and implements logistic regression adjusted for the number of probes in the enrichment analysis. The minimum and maximum number of genes in gene sets were set to 10 and 1000, respectively. All other settings were defaults. The result contains gene sets ranked by p values. GO terms whose adjusted p values were less than 0.05 were considered statistically significant.

Epigenetic ages and estimates of blood cell counts

We used Horvath's online calculator [54] to obtain 4 epigenetic ages (Horvath DNAm age [37], Hannum DNAm age [36], DNAm PhenoAge [39], and GrimAge [38]). The Hannum DNAm Age model measures the rate of human methylome aging under the influence of sex and genetic variants and is able to highlight certain components of the aging process. The Horvath DNAm Age model can be applied to a wide range of tissues and cell types, allowing to compare the age of different tissues of the same person to identify signs of accelerated aging associated with different diseases. Using a two-step process involving incorporation of clinical measures of phenotypic age, the epigenetic aging biomarker DNAm PhenoAge was developed, and it correlates with age in all tissues and cells tested. This model is capable of capturing the risks of a variety of outcomes in different tissues and cells and highlighting the aging pathways. Another predictor of longevity, DNAm GrimAge, is a biomarker based on a limited number of DNAm surrogates that can predict time to death, heart attack, cancer, and other age-associated diseases.

The PC modifications of the epigenetic ages were calculated according to the original algorithm [56]. These

PC clocks can minimize noisy PCs and separate noise from age-associated signals, and they use information from multiple CpGs, diluting noise from individual CpGs.

To compare different epigenetic ages, we used Pearson correlation coefficient [124]. Age acceleration was determined as follows for regions and sexes. To compare the regions (region-specific age acceleration), first, a linear regression was built between the different epigenetic ages and the chronological age for the Central region group. Age acceleration values were defined as residuals relative to this linear approximation. To compare the sexes (sex-specific age acceleration), a linear regression was built for the females group. This linear regression approach is recommended when the slope of the dependence between predicted age and real age differs from one [125–128]. Specifically, for our data, we observe an underestimation of Hannum and GrimAge in older participants and an overestimation of PCGrimAge in younger participants.

Horvath's online calculator also allows to obtain the distributions of blood cell counts: CD8T, CD4T, NK, B cells, Monocytes, and Granulocytes using Houseman algorithm [55].

The distributions of age-accelerated values and blood cell composition between the considered groups were tested using the Mann–Whitney U test [92]. This is a nonparametric test for comparing results between two independent groups, which is used to test the probability that two samples come from the same population, with a two-sided null hypothesis that the two groups are not the same. All resultant p values were FDR-corrected according to the Benjamini–Hochberg procedure [51].

Additionally, it is worth mentioning that both the original versions of the epigenetic clock in the Horvath's calculator and their PC modifications were developed for the Illumina 450 k methylation data standard, not for Illumina EPIC, so the values of some CpGs may be missing. Also, some CpGs involved in the calculation of epigenetic clock values are excluded at the data preprocessing stages (filtering, quality control, exclusion of SNP-related probes); therefore, all values of such missing for various reasons CpGs were automatically imputed according to the algorithms in the corresponding articles.

Abbreviations

DMP	Differentially methylated position
DNAm	DNA methylation
EWAS	Epigenome-wide association study
FDR	False discovery rate
GO	Gene ontology
GSEA	Gene set enrichment analysis
GWAS	Genome-wide association study
PC	Principal component
RBC	Red blood cell

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-023-01600-y.

Additional file 1. Supplementary Tables S1–S6.

Acknowledgements
The authors acknowledge Tatyana Klimova and Vladimir Osakovsky for help in organizing and participation in the discussion.

Author contributions
AK, IY, EK, MGB, CG, MV, CF, and MI contributed to conceptualization; AK, IY, EK, RZ, TS, MV, CF, and MI were involved in investigation; EK, RZ, TS, SS, AK, MN, and MV contributed to resources; AK, IY, MGB, CG, and MI were involved in methodology; AK and IY provided software and contributed to visualization, writing—original draft, and formal analysis; MGB, CG, EK, MV, CF, and MI were involved in supervision; and AK, IY, EK, MGB, CG, EK, CF, and MI contributed to writing—review and editing.

Funding
This work was supported by Lobachevsky University funding in the framework of "Priority-2030" State program, project No. N-470–99 (A.Ka., I.Y., E.Ko., M.V., C.F., M.I.), and by the state task of the Ministry of Education and Science of the Russian Federation, project FSRG-2023–0003 "Genetic characteristics of the population of the North-East of Russia: reconstruction of genetic history, mechanisms of adaptation and aging, age-dependent and hereditary diseases" (R.Z., T.S., S.S., A.Ks., M.N.).

Availability of data and materials
DNA methylation profiles generated in this study will be available as the GEO public repository with accession number GSE234461 prior to publication.

Code availability
No new or unpublished methods were used in the study. The scripts used in the generation of the manuscript are available on GitHub: https://github.com/GillianGrayson/EWAS_Yakutia.

Declarations

Ethics approval and consent to participate
Each participant signed an informed consent and filled out a consent for personal data processing, taking into account the principle of confidentiality (accessibility only to the research group and presentation of data in a common array). The study was approved by the local ethical committee of Nizhny Novgorod State University. All research procedures were in accordance with the 1964 Helsinki Declaration and its later amendments.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Received: 22 September 2023 Accepted: 13 November 2023
Published online: 06 December 2023

References

1. Pitulko VV, Nikolsky PA, Girya EY, Basilyan AE, Tumskoy VE, Koulakov SA, et al. The Yana RHS site: humans in the Arctic before the last glacial maximum. Science. 2004;303(5654):52–6.
2. Goebel T. Pleistocene human colonization of Siberia and peopling of the Americas: an ecological approach. Evol Anthropol. 1999;8(6):208–27.

3. Kuzmin YV. Siberia at the last glacial maximum: environment and archaeology. J Archaeol Res. 2008;16(2):163–221.
4. Fedorova SA, Reidla M, Metspalu E, Metspalu M, Rootsi S, Tambets K, et al. Autosomal and uniparental portraits of the native populations of Sakha (Yakutia): implications for the peopling of Northeast Eurasia. BMC Evol Biol. 2013;13(1):127.
5. Dulik MC, Zhadanov SI, Osipova LP, Askapuli A, Gau L, Gokcumen O, et al. Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between Native Americans and Indigenous Altaians. Am J Hum Genet. 2012;90(2):229–46.
6. Malyarchuk B, Derenko M, Denisova G, Maksimov A, Wozniak M, Grzybowski T, et al. Ancient links between Siberians and native Americans revealed by subtyping the Y chromosome haplogroup Q1a. J Hum Genet. 2011;56(8):583–8.
7. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science. 2008;319(5866):1100–4.
8. Fedorova SA, Khusnutdinova EK. Genetic structure and genetic history of the Sakha (Yakuts) population. Russ J Genet. 2022;58(12):1409–26.
9. Cardona A, Pagani L, Antao T, Lawson DJ, Eichstaedt CA, Yngvadottir B, et al. Genome-wide analysis of cold adaptation in indigenous siberian populations. PLoS ONE. 2014;9(5):e98076.
10. Leonard WR, Snodgrass JJ, Sorensen MV. Metabolic adaptation in indigenous Siberian populations. Annu Rev Anthropol. 2005;34(1):451–71.
11. Bjerregaard P, Dewailly E, Young TK, Blanchet C, Hegele RA, Ebbesson SEO, et al. Blood pressure among the Inuit (Eskimo) populations in the Arctic. Scand J Public Health. 2003;31(2):92–9.
12. Snodgrass JJ, Leonard WR, Sorensen MV, Tarskaia LA, Mosher MJ. The influence of basal metabolic rate on blood pressure among indigenous Siberians. Am J Phys Anthropol. 2008;137(2):145–55.
13. Nikanorova AA, Barashkov NA, Nakhodkin SS, Pshennikova VG, Solovyev AV, Romanov GP, et al. The role of leptin levels in adaptation to cold climates. Int J Environ Res Public Health. 2020;17(6):1854.
14. Nikanorova AA, Barashkov NA, Pshennikova VG, Nakhodkin SS, Gotovtsev NN, Romanov GP, et al. The role of nonshivering thermogenesis genes on leptin levels regulation in residents of the coldest region of Siberia. Int J Mol Sci. 2021;22(9):4657.
15. Nikanorova AA, Barashkov NA, Pshennikova VG, Gotovtsev NN, Romanov GP, Solovyev AV, et al. Relationships between uncoupling protein genes UCP1, UCP2 and UCP3 and Irisin levels in residents of the coldest region of Siberia. Genes. 2022;13(9):1612.
16. Pakendorf B, Morar B, Tarskaia LA, Kayser M, Soodyall H, Rodewald A, et al. Y-chromosomal evidence for a strong reduction in male population size of Yakuts. Hum Genet. 2002;110(2):198–200.
17. Pakendorf B, Wiebe V, Tarskaia LA, Spitsyn VA, Soodyall H, Rodewald A, et al. Mitochondrial DNA evidence for admixed origins of central Siberian populations. Am J Phys Anthropol. 2003;120(3):211–24.
18. Breeze CE, Wong JYY, Beck S, Berndt SI, Franceschini N. Diversity in EWAS: current state, challenges, and solutions. Genome Medicine. 2022;14(1):71.
19. Jhun MA, Mendelson M, Wilson R, Gondalia R, Joehanes R, Salfati E, et al. A multi-ethnic epigenome-wide association study of leukocyte DNA methylation and blood lipids. Nat Commun. 2021;12(1):3987.
20. Elliott HR, Burrows K, Min JL, Tillin T, Mason D, Wright J, et al. Characterisation of ethnic differences in DNA methylation between UK-resident South Asians and Europeans. Clin Epigenetics. 2022;14(1):130.
21. Chen Y, Kassam I, Lau SH, Kooner JS, Wilson R, Peters A, et al. Impact of BMI and waist circumference on epigenome-wide DNA methylation and identification of epigenetic biomarkers in blood: an EWAS in multi-ethnic Asian individuals. Clin Epigenet. 2021;13(1):195.
22. Song MA, Seffernick AE, Archer KJ, Mori KM, Park SY, Chang L, et al. Race/ethnicity-associated blood DNA methylation differences between Japanese and European American women: an exploratory study. Clin Epigenetics. 2021;11(13):188.
23. Kazmi N, Elliott HR, Burrows K, Tillin T, Hughes AD, Chaturvedi N, et al. Associations between high blood pressure and DNA methylation. PLoS ONE. 2020;15(1):e0227728.
24. Grönniger E, Weber B, Heil O, Peters N, Stäb F, Wenck H, et al. Aging and chronic sun exposure cause distinct epigenetic changes in human skin. PLoS Genet. 2010;6(5):e1000971.
25. Vandiver AR, Irizarry RA, Hansen KD, Garza LA, Runarsson A, Li X, et al. Age and sun exposure-related widespread genomic blocks of hypomethylation in nonmalignant skin. Genome Biol. 2015;16(1):80.
26. Mackenbach JP, Valverde JR, Artnik B, Bopp M, Brønnum-Hansen H, Deboosere P, et al. Trends in health inequalities in 27 European countries. Proc Natl Acad Sci U S A. 2018;115(25):6440–5.
27. Artazcoz L, Rueda S. Social inequalities in health among the elderly: a challenge for public health research. J Epidemiol Community Health. 2007;61(6):466–7.
28. Fiorito G, McCrory C, Robinson O, Carmeli C, Rosales CO, Zhang Y, et al. Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis. Aging (Albany NY). 2019;11(7):2045–70.
29. Levine ME, Crimmins EM. Evidence of accelerated aging among African Americans and its implications for mortality. Soc Sci Med. 2014;118:27–32.
30. Diez Roux AV, Ranjit N, Jenny NS, Shea S, Cushman M, Fitzpatrick A, et al. Race/ethnicity and telomere length in the multi-ethnic study of atherosclerosis. Aging Cell. 2009;8(3):251–7.
31. Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. Genome Biol. 2016;17(1):171.
32. Gensous N, Garagnani P, Santoro A, Giuliani C, Ostan R, Fabbri C, et al. One-year Mediterranean diet promotes epigenetic rejuvenation with country- and sex-specific effects: a pilot study from the NU-AGE project. GeroScience. 2020;42(2):687–701.
33. Fraser HB, Lam LL, Neumann SM, Kobor MS. Population-specificity of human DNA methylation. Genome Biol. 2012;13(2):R8.
34. Heyn H, Moran S, Hernando-Herraez I, Sayols S, Gomez A, Sandoval J, et al. DNA methylation contributes to natural human variation. Genome Res. 2013;23(9):1363–72.
35. Galanter JM, Gignoux CR, Oh SS, Torgerson D, Pino-Yanes M, Thakur N, et al. Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. Elife. 2017;6:e20532.
36. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359–67.
37. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
38. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY). 2019;11(2):303–27.
39. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573–91.
40. McCartney DL, Zhang F, Hillary RF, Zhang Q, Stevenson AJ, Walker RM, et al. An epigenome-wide association study of sex-specific chronological ageing. Genome Med. 2019;12(1):1.
41. Singmann P, Shem-Tov D, Wahl S, Grallert H, Fiorito G, Shin SY, et al. Characterization of whole-genome autosomal differences of DNA methylation between men and women. Epigenet Chromatin. 2015;8(1):43.
42. Numata S, Ye T, Hyde TM, Guitart-Navarro X, Tao R, Wininger M, et al. DNA methylation signatures in development and aging of the human prefrontal cortex. Am J Human Genet. 2012;90(2):260–72.
43. Hall E, Volkov P, Dayeh T, Esguerra JLS, Salö S, Eliasson L, et al. Sex differences in the genome-wide DNA methylation pattern and impact on gene expression, microRNA levels and insulin secretion in human pancreatic islets. Genome Biol. 2014;15(12):522.
44. Grant OA, Wang Y, Kumari M, Zabet NR, Schalkwyk L. Characterising sex differences of autosomal DNA methylation in whole blood using the Illumina EPIC array. Clin Epigenetics. 2022;14(1):62.
45. Pellegrini C, Pirazzini C, Sala C, Sambati L, Yusipov I, Kalyakulina A, et al. A meta-analysis of brain DNA methylation across sex, age, and alzheimer's disease points for accelerated epigenetic aging in neurodegeneration. Front Aging Neurosci. 2021. https://doi.org/10.3389/fnagi.2021.639428.
46. Yusipov I, Bacalini MG, Kalyakulina A, Krivonosov M, Pirazzini C, Gensous N, et al. Age-related DNA methylation changes are sex-specific: a comprehensive assessment. Aging. 2020;12(23):24057–80.

47. Masser DR, Hadad N, Porter HL, Mangold CA, Unnikrishnan A, Ford MM, et al. Sexually divergent DNA methylation patterns with hippocampal aging. Aging Cell. 2017;16(6):1342–52.
48. Iannuzzi V, Bacalini MG, Franceschi C, Giuliani C. The role of genetics and epigenetics in sex differences in human survival. Genus. 2023;79(1):1.
49. National Centers for Environmental Information (NCEI) [Internet]. [cited 2023 Mar 6]. Available from: https://www.ncei.noaa.gov/
50. Global Economic Data, Indicators, Charts & Forecasts [Internet]. [cited 2023 Mar 6]. Available from: https://www.ceicdata.com/en/
51. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57(1):289–300.
52. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000;25(1):25–9.
53. Gene Ontology Consortium. The gene ontology resource: enriching a GOld mine. Nucleic Acids Res. 2021;49(D1):D325–34.
54. DNA Methylation Age Calculator [Internet]. [cited 2023 Mar 7]. Available from: https://dnamage.clockfoundation.org/
55. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13(1):86.
56. Higgins-Chen AT, Thrush KL, Wang Y, Minteer CJ, Kuo PL, Wang M, et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. Nat Aging. 2022;2(7):644–61.
57. Belsky DW, Caspi A, Corcoran DL, Sugden K, Poulton R, Arseneault L, et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. Elife. 2022;11:e73420.
58. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7): e47.
59. Bernard BJ, Nigam N, Burkitt K, Saloura V. SMYD3: a regulator of epigenetic and signaling pathways in cancer. Clin Epigenetics. 2021;13(1):45.
60. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of the human proteome. Nature. 2014;509(7502):575–81.
61. Hedman ÅK, Mendelson MM, Marioni RE, Gustafsson S, Joehanes R, Irvin MR, et al. Epigenetic patterns in blood associated with lipid traits predict incident coronary heart disease events and are enriched for results from genome-wide association studies. Circ Cardiovasc Genet. 2017;10(1):e001487.
62. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature. 2017;541(7635):81–6.
63. Kvaløy K, Page CM, Holmen TL. Epigenome-wide methylation differences in a group of lean and obese women – a HUNT study. Sci Rep. 2018;8(1):16330.
64. Ligthart S, Marzi C, Aslibekyan S, Mendelson MM, Conneely KN, Tanaka T, et al. DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. Genome Biol. 2016;17(1):255.
65. Garoot NA, Kim BG. Stage-specific differential DNA methylation data analysis during human erythropoiesis in chromosome 16. Genet Res. 2018;100:e5.
66. Teległów A, Romanovski V, Skowron B, Mucha D, Tota Ł, Rosińczuk J, et al. The effect of extreme cold on complete blood count and biochemical indicators: a case study. Int J Environ Res Public Health. 2022;19(1):424.
67. Fung-Leung WP. Phosphoinositide 3-kinase delta (PI3Kδ) in leukocyte signaling and function. Cell Signal. 2011;23(4):603–8.
68. Kang S, Tanaka T, Narazaki M, Kishimoto T. Targeting interleukin-6 signaling in clinic. Immunity. 2019;50(4):1007–23.
69. Coscia F, Taler-Verčič A, Chang VT, Sinn L, O'Reilly FJ, Izoré T, et al. The structure of human thyroglobulin. Nature. 2020;578(7796):627–30.
70. Pasca di Magliano M, Di Lauro R, Zannini M. Pax8 has a key role in thyroid cell differentiation. In: Proceedings of the National Academy of Sciences. 2000 97(24):13144–9.
71. Fares A. Winter hypertension: potential mechanisms. Int J Health Sci (Qassim). 2013;7(2):210–9.
72. Wong HK, Cheung TT, Cheung BMY. Adrenomedullin and cardiovascular diseases. JRSM Cardiovasc Dis. 2012. https://doi.org/10.1258/cvd.2012.012003.
73. Nguyen G, Delarue F, Burcklé C, Bouzhir L, Giller T, Sraer JD. Pivotal role of the renin/prorenin receptor in angiotensin II production and cellular responses to renin. J Clin Invest. 2002;109(11):1417–27.
74. Sunagawa Y, Morimoto T, Takaya T, Kaichi S, Wada H, Kawamura T, et al. Cyclin-dependent kinase-9 Is a component of the p300/GATA4 complex required for phenylephrine-induced hypertrophy in cardiomyocytes. J Biol Chem. 2010;285(13):9556–68.
75. Gonzalez-Teran B, Pittman M, Felix F, Thomas R, Richmond-Buccola D, Hüttenhain R, et al. Transcription factor protein interactomes reveal genetic determinants in heart disease. Cell. 2022;185(5):794-814.e30.
76. Blazer LL, Lima-Fernandes E, Gibson E, Eram MS, Loppnau P, Arrowsmith CH, et al. PR domain-containing protein 7 (PRDM7) is a histone 3 lysine 4 trimethyltransferase. J Biol Chem. 2016;291(26):13509–19.
77. Hallmark B, Karafet TM, Hsieh P, Osipova LP, Watkins JC, Hammer MF. Genomic evidence of local adaptation to climate and diet in indigenous Siberians. Mol Biol Evol. 2019;36(2):315–27.
78. Hsieh P, Hallmark B, Watkins J, Karafet TM, Osipova LP, Gutenkunst RN, et al. Exome sequencing provides evidence of polygenic adaptation to a fat-rich animal diet in indigenous siberian populations. Mol Biol Evol. 2017;34(11):2913–26.
79. Ren X, Kuan PF. methylGSA: a Bioconductor package and Shiny app for DNA methylation data length bias adjustment in gene set testing. Bioinformatics. 2019;35(11):1958–9.
80. Manni E, Jeffery N, Chambers D, Slade L, Etheridge T, Harries LW. An evaluation of the role of miR-361-5p in senescence and systemic ageing. Exp Gerontol. 2023;1(174): 112127.
81. Kim YJ, Cho MJ, Yu WD, Kim MJ, Kim SY, Lee JH. Links of cytoskeletal integrity with disease and aging. Cells. 2022;11(18):2896.
82. Dudarev AA. Public health practice report: water supply and sanitation in Chukotka and Yakutia, Russian Arctic. Int J Circumpolar Health. 2018;77(1):1423826.
83. Lolignier S, Gkika D, Andersson D, Leipold E, Vetter I, Viana F, et al. New insight in cold pain: role of ion channels, modulation, and clinical perspectives. J Neurosci. 2016;36(45):11435–9.
84. Babes A. Ion channels involved in cold detection in mammals: TRP and non-TRP mechanisms. Biophys Rev. 2009;1(4):193–200.
85. Gruss F. Hot new structures of the cold sensor, TRPM8, reveal insights into the fundamentals of cold perception and adaptation. Cell Calcium. 2020;1(85): 102112.
86. Cruz-Topete D, Oakley RH, Cidlowski JA. Glucocorticoid signaling and the aging heart. Front Endocrinol (Lausanne). 2020;27(11):347.
87. Fiacco S, Walther A, Ehlert U. Steroid secretion in healthy aging. Psychoneuroendocrinology. 2019;1(105):64–78.
88. Pääkkönen T, Leppäluoto J. Cold exposure and hormonal secretion: a review. Int J Circumpolar Health. 2002;61(3):265–76.
89. Pouikli A, Tessarz P. Metabolism and chromatin: a dynamic duo that regulates development and ageing. BioEssays. 2021;43(5):2000273.
90. Galkin F, Mamoshina P, Kochetov K, Sidorenko D, Zhavoronkov A. DeepMAge: a methylation aging clock developed with deep learning. Aging Dis. 2021;12(5):1252–62.
91. Burtseva T, Uvarova T, Savvina M, Shadrin V, Avrusin S, Chasnyk V. Health status of native people living in the Republic of Sakha (Yakutia). Int J Circumpolar Health. 2013. https://doi.org/10.3402/ijch.v72i0.21166.
92. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat. 1947;18(1):50–60.
93. Mueller SN, Rouse BT. Immune responses to viruses. Clin Immunol. 2008. https://doi.org/10.1016/B978-0-323-04404-2.10027-2.
94. Nikolich-Zugich J, Li G, Uhrlaub JL, Renkema KR, Smithey MJ. Age-related changes in CD8 T cell homeostasis and immunity to infection. Semin Immunol. 2012;24(5):356–64.
95. Mandala WL, Ananworanich J, Apornpong T, Kerr SJ, MacLennan JM, Hanson C, et al. Control lymphocyte subsets: can one country's values serve for another's? J Allergy Clin Immunol. 2014;134(3):759-761.e8.
96. Patel AA, Yona S. Inherited and environmental factors influence human monocyte heterogeneity. Front Immunol. 2019;7(10):2581.
97. Inoshita M, Numata S, Tajima A, Kinoshita M, Umehara H, Yamamori H, et al. Sex differences of leukocytes DNA methylation adjusted for estimated cellular proportions. Biol Sex Differ. 2015;6(1):11.

98. McCarthy NS, Melton PE, Cadby G, Yazar S, Franchina M, Moses EK, et al. Meta-analysis of human methylation data for evidence of sex-specific autosomal patterns. BMC Genomics. 2014;15(1):981.
99. Perdomo-Sabogal A, Nowick K, Piccini I, Sudbrak R, Lehrach H, Yaspo ML, et al. Human lineage-specific transcriptional regulation through GA-binding protein transcription factor alpha (GABPa). Mol Biol Evol. 2016;33(5):1231–44.
100. Karastergiou K, Smith SR, Greenberg AS, Fried SK. Sex differences in human adipose tissues – the biology of pear shape. Biol Sex Differ. 2012;31(3):13.
101. Gavin KM, Bessesen DH. Sex differences in adipose tissue function. Endocrinol Metab Clin North Am. 2020;49(2):215–28.
102. Wismann J, Willoughby D. Gender differences in carbohydrate metabolism and carbohydrate loading. J Int Soc Sports Nutr. 2006;3(1):28–34.
103. Varlamov O, Bethea CL, Roberts CT. Sex-specific differences in lipid and glucose metabolism. Front Endocrinol (Lausanne). 2015;19(5):241.
104. Mauvais-Jarvis F. Gender differences in glucose homeostasis and diabetes. Physiol Behav. 2018;1(187):20–3.
105. Tarnopolsky MA, Ruby BC. Sex differences in carbohydrate metabolism. Curr Opin Clin Nutr Metab Care. 2001;4(6):521.
106. Balachandran RC, Mukhopadhyay S, McBride D, Veevers J, Harrison FE, Aschner M, et al. Brain manganese and the balance between essential roles and neurotoxicity. J Biol Chem. 2020;295(19):6312–29.
107. Engelbrecht HR, Merrill SM, Gladish N, MacIsaac JL, Lin DTS, Ecker S, et al. Sex differences in epigenetic age in Mediterranean high longevity regions. Front Ag. 2022. https://doi.org/10.3389/fragi.2022.1007098.
108. Kankaanpää A, Tolvanen A, Saikkonen P, Heikkinen A, Laakkonen EK, Kaprio J, et al. Do epigenetic clocks provide explanations for sex differences in life span? A cross-sectional twin study. J Gerontol Ser A. 2022;77(9):1898–906.
109. O'Shea DM, Maynard T, Tremont G. DNA methylation "GrimAge" acceleration mediates sex/gender differences in verbal memory and processing speed: findings from the health and retirement study. J Gerontol Ser A. 2022;77(12):2402–12.
110. Nadolnik L. Role of Glucocorticoids in Regulation of Iodine Metabolism in Thyroid Gland: Effects of Hyper-And Hypocorticism [Internet]. Glucocorticoids - New Recognition of Our Familiar Friend. IntechOpen; 2012 [cited 2023 Mar 13]. Available from: https://www.intechopen.com/chapters/41160
111. Li X, Poschmann S, Chen Q, Fazeli W, Oundjian NJ, Snoeijen-Schouwenaars FM, et al. De novo BK channel variant causes epilepsy by affecting voltage gating but not Ca2+ sensitivity. Eur J Hum Genet. 2018;26(2):220–9.
112. Liang L, Li X, Moutton S, Schrier Vergano SA, Cogné B, Saint-Martin A, et al. De novo loss-of-function KCNMA1 variants are associated with a new multiple malformation syndrome and a broad spectrum of developmental and neurological phenotypes. Hum Mol Genet. 2019;28(17):2937–51.
113. Spacey SD, Hildebrand ME, Materek LA, Bird TD, Snutch TP. Functional implications of a novel EA2 mutation in the P/Q-type calcium channel. Ann Neurol. 2004;56(2):213–20.
114. Condliffe SB, Zhang H, Frizzell RA. Syntaxin 1A regulates ENaC channel activity. J Biol Chem. 2004;279(11):10085–92.
115. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biol. 2016;17(1):208.
116. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k chip analysis methylation pipeline. Bioinformatics. 2014;30(3):428–30.
117. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. Bioinformatics. 2017;33(24):3982–4.
118. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the illumina infinium methylationEPIC BeadChip. Genomics Data. 2016;1(9):22–4.
119. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Res. 2017;45(4): e22.
120. Nordlund J, Bäcklin CL, Wahlberg P, Busche S, Berglund EC, Eloranta ML, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol. 2013;14(9): r105.
121. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363–9.
122. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118–27.
123. Leek JT, Johnson WE, Parker HS, Fertig EJ, Jaffe AE, Storey Y, et al. sva: Surrogate Variable Analysis [Internet]. Bioconductor version: Release (3.16); 2023 [cited 2023 Mar 22]. Available from: https://bioconductor.org/packages/sva/
124. Pearson K Notes on regression and inheritance in the case of two parents. In: Proceedings of the Royal Society of London. 1895, p 240–2
125. Higgins-Chen AT, Thrush KL, Levine ME. Aging biomarkers and the brain. Semin Cell Dev Biol. 2021;116:180–93.
126. Levine ME. Assessment of epigenetic clocks as biomarkers of aging in basic and population research. J Gerontol A Biol Sci Med Sci. 2020;75(3):463–5.
127. Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. J Infect Dis. 2015;212(10):1563–73.
128. El Khoury LY, Gorrie-Stone T, Smart M, Hughes A, Bao Y, Andrayas A, et al. Systematic underestimation of the epigenetic clock and age acceleration in older subjects. Genome Biol. 2019;17(20):283.

Publisher's Note