Complex sex-specific age-related changes in DNA methylation including variability, epimutations and entropy.

Igor Yusipov[#,1], Maria Giulia Bacalini[#*,2], Alena Kalyakulina[#,1], Mikhail Krivonosov[1], Chiara Pirazzini[2], Noémie Gensous[3], Francesco Ravaioli[3], Maddalena Milazzo[3], Maria Vedunova[4], Giovanni Fiorito[5], Amedeo Gagliardi[6], Silvia Polidoro[6], Paolo Garagnani[7,8,9], Mikhail Ivanchenko[$,1], Claudio Franceschi[$1]

# co-first authorship
$ co-senior authorship
*corresponding author

[1] Institute of Information Technologies, Mathematics and Mechanics, Lobachevsky University, Nizhniy Novgorod, Russia
[2] IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy
[3] Department of Experimental, Diagnostic and Specialty Medicine (DIMES), Alma Mater Studiorum – University of Bologna, 40126 Bologna, Italy
[4] Institute of Biology and Biomedicine, National Research Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russian Federation.
[5] Department of Biomedical Sciences, University of Sassari, Italy
[6] IIGM Italian Institute for Genomic Medicine, Torino, Italy
[7] Department of Laboratory Medicine,Clinical Chemistry, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden
[8] Applied Biomedical Research Center (CRBA), Policlinico S.Orsola-Malpighi Polyclinic, Bologna, Italy
[9] CNR Institute of Molecular Genetics "Luigi Luca Cavalli-Sforza", Unit of Bologna, Bologna, Italy

Abstract

Current epidemiological data indicate that, in humans, females live longer than males but experience a worse quality of life in advanced age. The reasons for this sex disparity are still unknown, but it is likely that it derives from a strict interplay between biological and cultural factors. Epigenetic modifications likely contribute to shape sex gap in aging and longevity, and genome-wide DNA methylation differences between males and females in autosomal chromosomes have been reported. Several studies showed that DNA methylation patterns are profoundly remodelled during aging, modulated in part by environmental exposures. However, few studies have specifically investigated if DNA methylation is differently affected by aging in males and females.

Here we performed a meta-analysis of 4 large whole blood datasets including males and females of different ages and we compared 4 aspects of epigenetic age-dependent remodelling between males and females: normative changes, variability, epimutations, and entropy. While we did not find differences in the age-associated increase in epimutations and in entropy, we reported a list of highly reproducible sex-specific age-associated differentially methylated positions (saDMPs) and sex-specific age-associated variably methylated positions (saVMPs). We investigated the enrichment in saDMPs and saVMPs in genomic regions, imprinted and sex hormone-related genes and Reactome pathways. Furthermore, we experimentally validated the most robust saDMPs, mapping in *FIGN* and *PRR4* genes, and showed sex-specific deviations of their methylation patterns in models of successful (centenarians) and unsuccessful (Down syndrome) aging.

In conclusion, we provided a comprehensive description of sex-differences in DNA methylation changes with aging in whole blood. Our results can pave the way to the identification of possible molecular triggers of the sex gap in aging and longevity.

Introduction

A profound and multifaceted remodeling of DNA methylation patterns occurs during human aging [1–3]. DNA methylation profiles tend to diverge among individuals during life course [4–6], shaped by an intricated combination of environmental exposures, random events and genetically-driven mechanisms. At the same time, several epigenome-wide association studies (EWAS) have shown that a subset of the about 28 million CpG sites of the genome undergoes age-associated normative changes, i.e. reproducible hypermethylation or hypomethylation events that characterize aging individuals [7,8]. Despite some controversial results [9,10], at least a fraction of normative changes is tissue specific, indicating that the cellular microenvironment affects the activity of the molecular writers of DNA methylation patterns during aging. The number of studies identifying age-associated DNA methylation changes at the level of single CpG sites has exponentially increased in the last 10 years, paving the way for the development of mathematical models, termed "epigenetic clocks", that predict the age of an individual on the basis of his/her epigenetic profile [11]. Epigenetic clocks are an appealing resource for chronological age estimation in forensic applications, but they have risen to the limelight particularly because multiple reports have shown that they are sensitive to the health status of an individual and are thus informative of his/her biological age. Although a conclusive association between epigenetic clocks predictions and risk of age-related diseases is still missing [12], several independent studies showed that epigenetic age acceleration (i.e., predicted epigenetic age higher than effective chronological age) is associated to age-related diseases like cancer, cardiovascular disease and neurodegenerative conditions and to all-cause mortality [13]. On the other side, epigenetic age deceleration was reported to be associated with successful aging and longevity [14,15].

Surprisingly, the research on the DNA methylation changes occurring during aging has largely neglected one of the hot topics in aging research, i.e. the sex differences in lifespan and health span. According to Global Health Observatory (GHO) data [16], global life expectancy at birth in 2016 was 74.2 years for females and 69.8 years for males and, although with different extent, this sex gap in longevity is worldwide [17]. At the same time, epidemiological data indicate that, in humans, females live longer than males but experience a worse quality of life in advanced age [18]. Sex disparity exists for several diseases: cardiovascular disease, cancer and Parkinson's disease have higher mortality rates in males than in females at a given age, while females are at higher risk of Alzheimer's disease and show an increased prevalence of disabling conditions like bone and joint problems and autoimmune diseases. The reasons of these differences are still unclear, but it is likely that they result from a strict interplay between nature (for example, differences in hormones, asymmetries in genetic inheritance, sexual dimorphism) and nurture (for example, different vulnerability to environmental hazards, sexual selection). Notably, sex-specific longevity loci have been recently identified [19], further pointing out the contribution of sex on aging trajectories.

DNA methylation is a key regulator of development and it is sensitive to several environmental exposures. Independent studies reported DNA methylation differences between males and females in various tissues [20–22], involving CpG sites widespread across the autosomal chromosomes. Notwithstanding, few studies have investigated whether and how aging affects sex-dependent epigenetic patterns, or if sex-specific epigenetic profiles emerge specifically during the life of an individual. Two independent studies reported that according to Horvath's clock males have accelerated epigenetic age compared to females [23,24]. Recently, Masser et al. analyzed genome-wide DNA methylation in mouse hippocampus and human frontal cortex and reported both CpG sites that show different DNA methylation levels between males and females lifelong (referred as sex differences) and CpG sites that are differently affected by aging in males and females (referred as sex divergence) [25]. The vast majority of EWAS studies on aging have been performed in whole blood, but to the best of our knowledge in these analysis sex has always been exiled as a confounding factor in statistical analysis. However, it is well known that immunological differences exist between males and females [26,27] and that age differently shapes the immune system in the two sexes [28].

Furthermore, the peripheral tissue can reflect age-related patterns that have functional and phenotypic effects occurring in other tissues.

On the basis of these considerations, question arises if DNA methylation differences exist in whole blood between males and females during aging and, if existing, whether they contribute to the sex gap in aging and longevity. In the present study we performed a meta-analysis of 4 large whole blood datasets including males and females of different ages, in order to identify sex-specific trends in age-related DNA methylation patterns. In order to provide a comprehensive view of this topic, we considered 4 aspects of epigenetic age: 1) age-related normative changes in DNA methylation levels [8,29]; 2) age-related increase in DNA methylation variability, as described by Slieker et al. [5]; 3) age-related increase in epimutations, i.e. rare or stochastic changes in DNA methylation levels that are not shared among subjects, as defined by Gentilini et al. [6] 4) age-related increase in entropy in DNA methylation profiles, as previously described by Hannum et al. [30].

Materials and Methods

Datasets

The Gene Expression Omnibus (GEO) Datasets repository [31] was interrogated using "GPL13534" (the accession code of the platform HumanMethylation450 BeadChip, Illumina) and "blood" as search terms, setting "tissue", "age", "gender" and "sex" as attributes and sorting the results by Number of Samples (High to Low). Only datasets including healthy subjects were considered. Based on these criteria, as to June 1st 2019 we selected the 3 datasets including the highest number of samples: GSE40279 [30], GSE87571 [32] and GSE55763 [33]. Furthermore, we analyzed a fourth dataset not uploaded in GEO that is part of the EPIC Italy study [34]. The total number of subjects included in each dataset, as well as the number of males and females, is reported in Supplementary Table 1. Supplementary Figure 1 reports, for each dataset, the number of males and females according to age.

Raw data (.idat files) were available only for GSE87571 dataset. We extracted raw data using minfi Bioconductor package and normalized them using the *preprocessFunnorm* function implemented in the same package. For the remaining datasets the analyses were performed on available beta value data, which according to authors' indications were normalized by quantile normalization of intensity values (GSE55763) or using an in-house software written for the R environment and extensively described in [35]. For GSE40279, available beta values were not normalized but adjusted for internal controls with Illumina's Genome Studio software.

Probes mapping on sex chromosomes and probes with internal SNPs, with non-unique mapping to the bisulfite-converted genome and with off-target hybridization, according to [36], were excluded from each dataset, leaving 414505 probes for GSE40279, 414950 probes for GSE87571, 349534 probes for EPIC and 382458 probes for GSE55763.

In each dataset, blood cell counts were estimated from methylation data using Horvath calculator [29,37]. Residuals were calculated regressing out from beta values the estimates of CD8T cells, CD4T cells, NK cells, B cells and Granulocytes.

Identification of probes having sex-specific trends in age-associated methylation changes

To identify CpG sites showing sex-specific trends in age-associated hyper- or hypo-methylation (sex-specific age-associated differentially methylated positions, saDMPs), each dataset was divided into two subsets (females and males) and a linear regression model by Ordinary Least Squares method was built for each probe, considering the dependence of beta values on age. The normality of residuals distribution was checked by Shapiro-Wilk test and only the probes having a Benjamini-Hochberg corrected p-value higher than 0.05 were retained. A CpG site was defined as having a sex-specific trend in age-associated hyper- or hypo-methylation if it fulfilled the following criteria (polygon approach): 1) the probe undergoes hyper- or hypo-methylation during aging at least in one sex: we selected probes having linear regression slope for

males or females greater than 0.001; 2) the probe has age-dependent beta values distribution only partially overlapping between males and females: for each probe we calculated the allowed region in males and females using confidence intervals for slope and intercept of linear regression and we selected only those probes having relative intersection area (i.e. area of intersection divided by area of union) between allowed regions for females and males less than 0.5. An example of allowed regions intersection area is shown in Figure 1A.

Identification of probes having sex-specific trends in age-associated methylation variability

To identify probes having sex-specific differences in age-dependent variability of methylation (sex-specific age-associated variably methylated positions, saVMPs), we first of all filtered out probes with bi-, three-modal distribution of beta values using data clustering algorithm DBSCAN – density-based spatial clustering of applications with noise [38], leaving 412618 probes for GSE40279, 412015 probes for GSE87571, 348564 probes for EPIC and 381006 probes for GSE55763.

For each age value present in each dataset, we applied a sliding window of 16 years, thus selecting all the subjects 8 years younger and 8 years older. Then in each subset, we calculated 5% and 95% percentiles of beta values distributions for males and females separately. We then fitted these values with linear regression in different scales: 1) untransformed age and untransformed beta values; 2) untransformed age and logarithmic beta values; 3) logarithmic age and logarithmic beta values. For each probe and for each sex, we selected the best fitting scale, which is the one with the highest R^2 value. For each probe, we had therefore four R^2 values (two associated to sex and two associated to 5% and 95% percentiles). We selected the minimal R^2 value to guarantee good fitting for both sexes and both quartiles at once. Then, for each dataset we chose probes having R^2 value greater than 75% percentile for the distribution of R^2 (Supplementary Figure 2), leaving 103155 probes for GSE40279, 103004 probes for GSE87571, 87141 probe for EPIC and 95252 probes for GSE55763. The quartiles values were 0.128 for GSE40279, 0.298 for GSE87571, 0.335 for EPIC and 0.357 for GSE55763. The R2 values were drastically lower in GSE40279 dataset compared to the other datasets, which may be due to normalization absence in the dataset.

Then, for each probe we used the specific fitting identified above and calculated the difference between 95% and 5% percentiles at the smallest (s) and the highest (h) age within each dataset for males and for females separately (Figure 1B). We calculated the ratio between these two values (h/s if h higher than s and s/h in the opposite case) in order to define the absolute value of age-dependent change in variance (v). Finally, we calculated I, that is the ratio of v in males and females (v^{male}/v^{female} if v^{male} higher than v^{female} and v^{female}/v^{male} in the opposite case) to quantify absolute sex-related differences in age-dependent variability. To identify probes with robust sex-specific differences in age-dependent variability in all datasets, we intersected the 4 lists of probes, leaving 10829 probes. Then, we selected only probes concordant between all datasets (probes with v^{male} greater than v^{female} or vice versa in all 4 datasets), leaving 1231 probes.

Value of I depends of the dataset age range and type of 5% and 95% percentiles fitting: it is greater for the larger age range and nonlinear fitting. To equalize the 4 datasets, for each probe we calculated mean I value between all datasets and selected the probes with mean I higher than 1.5, which corresponds to at least 50% difference between sexes in increase (or decrease) of age-dependent variability in all datasets.

When we applied the same analysis on the residuals, we identified 103155 probes for GSE40279, 103004 probes for GSE87571, 87141 probe for EPIC and 95252 probes for GSE55763 having R^2 value greater than 75% percentile for the distribution of R^2. Intersection of 4 lists of probes left 13231 probes and 1687 probes concordant between all datasets.

Identification of epimutations and Shannon entropy analysis

To identify epimutations (i.e., CpG probes for which one or few individuals show extremely different methylation levels compared to the rest of the cohort), for each probe we calculated the interquartile ranges

of beta values; we then selected the probes having one or more subjects (epimutated subjects) having a beta value exceeding three times interquartile ranges (Q1 – (3×IQR) and Q3 + (3×IQR)), as reported in [6].

To calculate Shannon entropy, we applied the following procedure, according to [39]: 1) we obtained residuals by filtering out the dependence of beta-values on age and blood cells count; 2) we recalculated beta-values according to the formula:

$$\beta_{i,j}^{adj} = residuals_{i,j} + mean(\beta_i) \quad (1)$$

where $mean(\beta_i)$ is the average methylation level for ith CpG site, i is the index of CpG and j is the index of subject. Then, we calculated Shannon entropy using the following formula, as indicated in [30]:

$$Entropy = 1 / N \cdot log\,(1/2) \sum_i [\beta_i^{adj} \cdot log(\beta_i^{adj}) + (1 - \beta_i^{adj}) \cdot log(1 - \beta_i^{adj})] \quad (2)$$

where β_i^{adj} is the recalculated methylation level for ith CpG site and N is the number of CpG sites.

Gene-targeted DNA methylation analysis

The EpiTYPER assay (Agena) was used to measure DNA methylation of FIGN and PRR4 in whole blood from 571 subjects belonging to 4 groups: 419 healthy controls of different ages, 49 centenarians, 48 centenarians' offspring and 44 persons with Down Syndrome. Age range and sex distribution of the 4 cohorts are reported in Supplementary Table 2. All subjects were recruited following approval by the Ethical Committee of Sant'Orsola-Malpighi University Hospital (Bologna, Italy).

Genomic DNA was extracted using the QIAamp 96 DNA Blood Kit (Qiagen) and 500 ng were bisulphite converted using EZ-96 DNA Methylation Kit (Zymo Research Corporation). Ten ng of bisulphite-converted DNA were amplified using the following bisulphite-specific primers, containing tag sequences for EpiTYPER protocol: FIGN forward aggaagagagTTTTTTGAAAAGAGAGAAAGAAGGA; FIGN reverse cagtaatacgactcactatagggagaaggctATAAACAATCAAACCATCCAATTTCTA; PRR4 forward aggaagagagTTTGTGTTTTGAGTTGAGTTTAGAG; PRR4 reverse cagtaatacgactcactatagggagaaggctCCTAAAAATAAAACTTCTATCATCCA. Primers for FIGN and PRR4 amplified chr2:164,589,883-164,590,418 and chr12:11,001,978-11,002,636 (GRCh37/hg19 genome assembly) respectively.

Enrichment analyses and Transcriptomine search

Enrichment of genomic regions, imprinted genes and sex-hormone related genes was calculated using Fisher exact test, as implemented in the *fisher.test* function in the *stats* R package. Enrichment of Reactome annotations was calculated using the *methylgometh* function implemented in the *methylGSA* R package, using default settings. Transcriptomine database available at NURSA website (www.nursa.org/nursa/transcriptomine) was searched for the 5 genes *FIGN*, *PRR4*, *SOGA3*, *PEX10* and *BAG3*, selecting the following pathways: "AR & Androgens", "ERs and estrogens", "ERR subfamily".

Results

Identification of sex-specific age-associated differentially methylated positions (saDMPs)

As a first step, we aimed at identifying those probes undergoing hyper- or hypo-methylation during aging (normative changes) and displaying differences between males and females in these changes (sex-specific age-associated differentially methylated positions, saDMPs). To this end, we applied a "polygon" approach, described in the Materials and Methods section and represented in Figure 1A. The number of CpGs selected according to the polygon approach in each dataset and their intersections are shown in Figure 1B, while the complete list of probes is reported in Supplementary File 1.

The number of probes satisfying the selection criteria was drastically lower in GSE40279 dataset compared to the other datasets. This difference can be attributed to the fact that GSE40279 has the smallest size, or to the fact that only not normalized beta values were available in GEO for this dataset. We identified 7 saDMPs common to all the 4 datasets (Table 1), and 29 when GSE40279 was excluded (Supplementary File 1). Figure 1C reports DNA methylation values according to age and sex in each dataset for each of the 7 saDMPs. Six out 7 of the saDMPs that we identified in our analysis have been reported to have sex-dependent methylation (independently from age) in previous reports [20–22], 3 of them also when newborns were considered [22] (Supplementary File 1). The 7 probes common to all datasets map in the 5 genes *FIGN*, *PRR4*, *C6orf174/SOGA3*, *PEX10* and *BAG3*. We searched these 5 genes in the Transcriptomine database, which allows to evaluate genes for their regulation by nuclear receptors, including receptors for androgens and estrogens. We found different evidences, summarized in Supplementary File 2, suggesting that the target genes can be regulated by estrogens and/or androgens in *in vitro* cell cultures, in part depending on the type of molecule, its concentration and duration of the exposure.

Age-associated changes in DNA methylation values can reflect modifications in the relative proportions of blood cell counts. To take into account this potentially confounding effect, we estimated blood cell counts from methylation data and we repeated the polygon analysis on the residuals of methylation values regressed for blood cells counts. Two CpG sites, mapping in *FIGN* and *PRR4* genes, were identified in all the 4 datasets (Supplementary Figure 2A). Both probes were the same identified in the analysis performed on beta values. Again, GSE40279 returned the smallest number of selected CpG probes, and removing it from the analysis led to the identification of 26 probes common to GSE55763, GSE87571 and EPIC datasets (Supplementary Figure 2B; Supplementary File 3). Seventeen of these probes were common to those identified above in the analysis performed on beta values.

Then, we aimed at investigating the possible functional meaning of differential regulation of DNA methylation in males and females during aging. As the list of 7 saDMPs common to all the 4 datasets was too short to perform this analysis, we focused only on GSE87571 dataset, as it had the most homogeneous distribution of ages. Using the above-described criteria, we identified 354 saDMPs in GSE87571 dataset (Supplementary File 1).

First of all, we explored whether these 354 saDMPs were enriched in specific genomic regions (Supplementary Figure 3A). We found that, when the analysis was performed on beta values, the selected probes were significantly enriched for CpG islands, north shores and north shelves (p-value 1.31e-8, 2.7e-3 and 2.7e-3 respectively). A similar trend was observed when the analysis was performed on the residuals, with the selected probes (n=306; Supplementary File 3) significantly enriched in CpG islands (p-value 4.90e-14; Supplementary Figure 3 B).

Next, we tested the enrichment of the 354 saDMPs in imprinted genes and in sex hormone-related genes (Supplementary File 4). The 354 probes mapped in 170 genes, 5 of which are reported as imprinted in the Geneimprint database; this enrichment was significant according to Fisher exact test (p-value 0.047; odds ratio: 2.64). We also checked for an enrichment in sex hormone-related genes, as suggested by [21], but no enrichment was found (Supplementary File 4). No enrichment for imprinted genes or for sex hormone-related genes was found when we considered the list of 306 saDMPs deriving from the analysis on residuals.

Finally, we performed pathway analysis using Reactome database. The list of saDMPs was enriched in genes involved in dermatan sulfate biosynthesis, in the response to metal ions and in biosynthesis of lipid mediators (Supplementary File 4). Some of these pathways (dermatan sulfate biosynthesis, biosynthesis of lipid mediators) were found also when pathway analysis was performed on the list of saDMPs calculated correcting for blood cell counts (Supplementary File 5).

Validation of saDMPs

The above-identified saDMPs were experimentally validated using the EpiTYPER assay, a high throughput approach for target DNA methylation analysis. Target regions were chosen within FIGN and PRR4 in order to include the cg01620164 and cg23256579 probes respectively. We analyzed whole blood from 198 males from 15 to 98 years old and 221 females from 23 to 98 years old.

The *FIGN* target region included 13 CpG sites; of these, 7 were measurable by the assay, grouped in 5 CpG units. CpG unit 3.4.5 included the microarray probe cg01620164. We found that this group of CpG sites showed a sex-specific DNA hypomethylation trajectory comparable to the one resulting from the microarray (Supplementary Figure 4); also the adjacent CpG sites showed a similar profile (Supplementary Figure 4), in particular CpG unit 9 (Figure 2A). This result suggests that in blood the CpG sites in this locus of at least 250bp are commonly regulated during aging according to the sex of the individual.

The *PRR4* target region included 5 CpG sites, all assessable by EpiTYPER and all corresponding to an Infinium450k probe; CpG units 3 and 4, corresponding to the Infinium450k probes cg23256579 and cg27615582, had the same mass and returned the same methylation value in the EpiTYPER assay. While CpG units 1 and 2 did not show age-dependent changes nor sex specificity (Supplementary Figure 5), CpG units 3 and 4 showed sex-dependent trajectories with aging (Figure 2B). However, while in the GSE87571 methylation values of cg23256579 and cg27615582 tended to diverge between the two sexes during aging, in EpiTYPER results, an opposite trend was found. This discrepancy could be due to the fact that the EpiTYPER assay is not able to distinguish the two CpG sites. Although less evident, also CpG 5 showed sex-related differences in age-associated methylation changes (Supplementary Figure 5).

As a further step, we used the EpiTYPER assay to evaluate the two validated loci in samples from additional cohorts available in our laboratory: persons affected by Down syndrome, that we previously demonstrated to have an acceleration in epigenetic age [40]; and centenarians and their offspring, as a model of successful aging experiencing a deceleration in epigenetic age [14]. Interestingly, we found sex-specific patterns of *FIGN* and *PRR4* methylation also in these models. Compared to aged controls (>80 years old), centenarian males displayed a highly variable DNA methylation profiles for FIGN amplicon, with about half of the subjects showing a female-like DNA methylation level (Figure 2A and Supplementary Figure 6); the differences in variance between control and centenarians' males (but not females) reached statistical significance for CpG unit 9 (F-test p-value: 0.02). No specific trends were found for PRR4 amplicon in centenarians' cohort (Supplementary Figure 7). Centenarians' offspring showed DNA methylation patterns comparable to age-matched controls for both the amplicons (Supplementary Figures 6 and 7). Persons affected by Down Syndrome showed DNA methylation profiles similar to age-matched controls in FIGN locus (Supplementary Figure 6). On the contrary, females affected by Down syndrome showed lower values of CpG unit 3 in PRR4 amplicon compared to sex- and age-matched healthy controls (p-value: 6.2*10-5; anova corrected for age), while no significant differences were found between males affected by Down syndrome compared to sex- and age-matched controls (Figure 2B and Supplementary Figure 7). The results of the statistical analyses performed on the centenarians', centenarians' offspring and Down syndrome cohorts are summarized in Supplementary File 6.

Identification of sex-specific age-associated variably methylated positions (saVMPs)

An increase in inter-individual DNA methylation variability has been described during aging, but a possible sex-specific effect has not been investigated. To have a general view of the sex-dependent trends in age-

related increase in DNA methylation variability, we plotted the density distributions of standard deviation values, calculated in the GSE87571 dataset in 3 age-ranges (14-39 years; 40-59 years; 60-94 years) considering the whole cohort (Supplementary Figure 8A) or separating males and females (Supplementary Figure 8B). A clear increase in standard deviation across the 3 age ranges was evident when considering the entire cohort. No clear differences between males and females were evident in the first 2 age ranges, while we found a trend towards higher variability in males in the oldest group.

To identify probes having sex-specific differences in age-dependent variability (sex-specific age-associated variably methylated positions, saVMPs), we applied the approach described in Materials and Methods and reported in Figure 3A. We identified 65 saVMPs showing different beta values variability in males and females at different ages, in a reproducible way among the 4 datasets. All 65 saVMPs showed increasing variance with age for both sexes in all datasets. Some examples are represented in Figure 3B and the full list is reported in Supplementary File 7. Only 2 of the saVMPs have sex differences according to previous reports [20–22]. As above, we repeated the same analysis on residuals of methylation values regressed for blood cells counts. We identified 99 saVMPs, all showing increasing variance with age for both sexes in all datasets (Supplementary File 8). Notably, 33 probes were identified in both beta values and residuals analysis, suggesting that sex-specific age-related variability of these probes is reproducible and not affected by changes in blood cell counts.

As for saDMPs, we focused on GSE87571 dataset and we used Fisher exact test to explore whether the saVMPs were enriched in specific genomic regions (Supplementary Figure 9). We found that, when the analysis was performed on beta values, the selected probes (n=1682; Supplementary File 6) were significantly enriched for CpG islands, south and north shelves (p-values 1.42e-72, 5.67e-8 and 4.92e-6 respectively). A similar trend was observed when the analysis was performed on the residuals, with the selected probes (n=1562; Supplementary File 7) significantly enriched in CpG islands, north and south shelves (p-values 7.96e-108, 8.50e-11 and 5.54e-9 respectively). The list of 1682 saDMPs was highly enriched in imprinted genes (p-value 1.7e-10; odds ratio: 3.68), and a similar enrichment was found when we considered the list of 1562 saVMPs calculated on residuals (p-value 9.2e-8; odds ratio: 3.22; Supplementary File 8). No enrichment in sex-related genes was found using both the lists (Supplementary File 8).

Finally, when we performed pathway analysis using Reactome database, we found only one pathway (Regulation of gene expression in early pancreatic precursor cells) enriched in saVMPs calculated on beta values (Supplementary File 9).

Epimutations and entropy analysis

Number of epimutations in the 4 datasets is shown in Figure 4A. Results demonstrated a non-sex-specific increase in the number of epimutations with age. The use of the second criterion of the “polygon” approach, described in Materials and Methods section, to epimutations data confirmed that there was no sex-specificity in the increase in epimutation numbers during aging.

The dependence of Shannon entropy on age for the 4 datasets is shown in Figure 4B. For all the datasets, entropy showed non-sex-specific a increase with age. Again, we applied the second criterion of “polygon” approach to entropy data, confirmeing that there was no statistically significant difference between sexes in Shannon entropy age-dependent increase.

Discussion

In the present study we exploited 4 large EWAS studies on human whole blood to investigate sex-specific changes in DNA methylation during aging, considering 4 different layers of epigenetic remodelling that potentially can, either individually or in combination, contribute to sex-specific traits of human aging: normative changes (age-associated hyper- or hypomethylation events), changes in variability, increase in

epimutations and increase in entropy. We focused on these 4 different aspects of age-associated remodelling of DNA methylation profiles as they have been largely described in their association with human aging, but they have been rarely (in the case of normative changes) or never (for the other 3 aspects) investigated for their sex-specificity.

Our results suggest that males and females do not differ for the age-associated increase in epimutations and in entropy, while we found a short list of CpG sites showing sex-specific changes in DNA methylation values (sex-specific age-associated differentially methylated positions, saDMPs) and variability (sex-specific age-associated variably methylated positions, saVMPs) during aging. Importantly, these short lists result from the intersection of analyses performed independently in 4 distinct datasets and include therefore robust, highly reproducible candidates of sex-specificity in age-associated epigenetic changes. Furthermore, the simultaneous analysis of different cohorts, recruited in different geographic area (United States, Sweden, Italy, United Kingdom) and including different ethnicities (Caucasian and Hispanic), contributes to disentangle the effects of sex from those of potentially confounding factors like genetic background and socio-cultural aspects contributing to gender definition. At the same time, it must be considered that the analyzed datasets differ in terms of size, age-range and data pre-processing procedures. It is therefore likely that our rigorous selection excluded additional CpG sites displaying a sex-specificity in their age-associated methylation trends, but not evident in all the datasets due to the above-mentioned differences between them. For this reason, for more descriptive analyses like region and pathways enrichment we focused on the GSE87571 dataset, which includes a large number of subjects well distributed in a wide age-range.

The 7 saDMPs map in 6 genes with different functions. *FIGN* gene encodes for Fidgetin protein, an ATP-dependent microtubule severing enzyme which catalyzes internal breaks in microtubules and is involved in different cellular processes, including cell division and neurogenesis [41]. The gene is poorly expressed in whole blood, while it is expressed at higher levels in arteries and in female reproductive organs. *PRR4* gene encodes for Proline-rich protein 4, a poorly characterized protein expressed in tear fluid. The gene has very low expression in most human tissues, including whole blood. Interestingly, whole blood methylation of cg23256579 probe has been previously reported to be associated with lupus nephritis in women with systemic lupus erythematosus [42]. *PEX10* encodes for a protein localized to the peroxisomal membrane and involved in import of peroxisomal matrix proteins. The gene is highly expressed in testis, and a meta-analysis suggests that *PEX10* polymorphisms are associated with male infertility, especially with non-obstructive azoospermia susceptibility. *SOGA3* is a member of SOGA (suppressor of glucose from autophagy) proteins, a poorly characterized family that inhibit glucose production, while BAG3 is part of a family of anti-apoptotic proteins acting as co-chaperones of Hsp70 [43]. Recent studies indicate that BAG3 plays a central role in selective macroautophagy pathway, which mediates degradation of aggregation-prone proteins that can accumulate upon cellular stress or aging [43,44]. As member of the protein quality control machinery, whose activity undergoes substantial changes during aging [45,46], *BAG3* has been implicated in several age-associated diseases, including cancer and neurodegenerative diseases [47]. *FIGN* and *PRR4* were found also when our analytical approach for the identification of saDMPs was corrected for blood cells counts, considering the intersection of all 4 the datasets; in addition, *PEX10* and *SOGA3* were found when correcting for blood cell counts but excluding the GSE40279 dataset. These results suggest that the differences that we identified in saDMPs (and in the associated genes) tend to be robust to the changes in blood cell counts that can be related to differences in the immune system between the two sexes during aging [48,49]. In addition, we found a large overlap between the probes in the short saDMPs list (and those in the GSE87571 saDMPs list) and the probes reported as differentially methylated between males and females, regardless the effect of aging, in previous studies [20–22]. It is worth to note that the 5 genes showed different trends in their sex-specific methylation changes during aging. For *FIGN*, *PRR4* and *BAG3* differences between males and females are divergent during aging: they arise early, at 20-30 years or before menopause, and tend to be more marked at older ages. On the contrary, *SOGA3* shows convergent patterns,

in which the differences between males and females tend to reduce during aging. Finally, differences in *PEX10* methylation between males and females tend do remain constant during aging.

It is difficult to speculate the mechanisms leading to differential methylation patterns of these genes in males and females aging. When considering autosomal differences of DNA methylation between men and women (correcting for age), Singmann et al. found an enrichment in CpG island shores and in imprinted genes, but they did not find an enrichment in sex hormone-related genes. Similarly, when considering only GSE87571 dataset, we found an enrichment in CpG island shores and, in addition, in CpG islands. We also found a mild enrichment in imprinted genes, and noteworthy one out the 5 genes in our short list is imprinted (*PEX10*). Similarly to Singmann et al., we did not found any enrichment in sex-hormone related genes, as only 1 out 170 genes in our list (*POMC*, which encodes for proopiomelanocortin) belonged to this category. On the other side, the Transcriptomine search suggests that sex hormones regulate the expression of the genes in which saDMPs map in *in vitro* cell cultures. Future studies should assess if sex hormones can directly regulate the methylation of the genes that we identified, as previously demonstrated for other genomic loci [50].

Another question is if and how the saDMPs that we identified contribute to the sex gap in health span and longevity. The data on Down syndrome (a model of premature/accelerated aging [51]) and on centenarians (a model of successful aging [14]) are intriguing, as both showed a sex-specific alteration in the observed sex-specific trends of *FIGN* and *PRR4* epigenetic patterns. In particular, a subset of centenarian males showed a "feminization" of *FIGN* methylation values, while females with Down syndrome showed a "masculinization" of *PRR4* methylation values. Further studies should deepen these results and identify other changes in saDMPs that are associated to age-related diseases or longevity.

Finally, we focused on the differences in age-associated variance in methylation patterns between the two sexes. First of all, we confirmed that an increase in DNA methylation variance occurs with aging [5] and we further reported a trend towards higher epigenetic variance in males compared to females at older ages. This result mimics what observed for gene expression in the hippocampus of male and female mice at different ages [52], thus suggesting that the loss of epigenetic and transcriptional control that occurs during aging is more marked in males than in females. Accordingly, higher genomic instability has been suggested in adult males compared to females [53]. A more specific analysis of saVMPs returned a list of 65 loci, 33 of which were replicated also after correction for blood cell counts. Interestingly, the list of saVMPs calculated on the GSE87571 dataset was not enriched in particular pathways, but was highly enriched in imprinted genes.

In conclusion, we provided a comprehensive description of sex-differences in DNA methylation changes with aging in whole blood. Future studies should investigate the tissue-specificity of these patterns and their relationship with gene expression differences between males and females, in order to identify possible molecular triggers of sex gap in aging and longevity.

Acknowledgments

We acknowledge support by the grant of the Ministry of Education and Science of the Russian Federation Agreement No. 075-15-2019-871.

References

[1] M.G. Bacalini, P. D'Aquila, E. Marasco, C. Nardini, A. Montesanto, C. Franceschi, G. Passarino, P. Garagnani, D. Bellizzi, The methylation of nuclear and mitochondrial DNA in ageing phenotypes and longevity, Mech. Ageing Dev. 165 (2017) 156–161. https://doi.org/10.1016/j.mad.2017.01.006.

[2] F. Ciccarone, S. Tagliatesta, P. Caiafa, M. Zampieri, DNA methylation dynamics in aging: how far are we from understanding the mechanisms?, Mech. Ageing Dev. (2017). https://doi.org/10.1016/j.mad.2017.12.002.

[3] V.V. Ashapkin, L.I. Kutueva, B.F. Vanyushin, Epigenetic Clock: Just a Convenient Marker or an Active Driver of Aging?, Adv. Exp. Med. Biol. 1178 (2019) 175–206. https://doi.org/10.1007/978-3-030-25650-0_10.

[4] R. Feil, M.F. Fraga, Epigenetics and the environment: emerging patterns and implications, Nat. Rev. Genet. 13 (2012) 97–109. https://doi.org/10.1038/nrg3142.

[5] R.C. Slieker, M. van Iterson, R. Luijk, M. Beekman, D.V. Zhernakova, M.H. Moed, H. Mei, M. van Galen, P. Deelen, M.J. Bonder, A. Zhernakova, A.G. Uitterlinden, E.F. Tigchelaar, C.D.A. Stehouwer, C.G. Schalkwijk, C.J.H. van der Kallen, A.

Hofman, D. van Heemst, E.J. de Geus, J. van Dongen, J. Deelen, L.H. van den Berg, J. van Meurs, R. Jansen, P.A.C. 't Hoen, L. Franke, C. Wijmenga, J.H. Veldink, M.A. Swertz, M.M.J. van Greevenbroek, C.M. van Duijn, D.I. Boomsma, P.E. Slagboom, B.T. Heijmans, Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms, Genome Biol. 17 (2016) 191. https://doi.org/10.1186/s13059-016-1053-6.

[6] D. Gentilini, P. Garagnani, S. Pisoni, M.G. Bacalini, L. Calzari, D. Mari, G. Vitale, C. Franceschi, A.M. Di Blasio, Stochastic epigenetic mutations (DNA methylation) increase exponentially in human aging and correlate with X chromosome inactivation skewing in females, Aging. 7 (2015) 568–578. https://doi.org/10.18632/aging.100792.

[7] M.J. Jones, S.J. Goodman, M.S. Kobor, DNA methylation and healthy human aging, Aging Cell. 14 (2015) 924–932. https://doi.org/10.1111/acel.12349.

[8] A. Unnikrishnan, W.M. Freeman, J. Jackson, J.D. Wren, H. Porter, A. Richardson, The role of DNA methylation in epigenetics of aging, Pharmacol. Ther. 195 (2019) 172–185. https://doi.org/10.1016/j.pharmthera.2018.11.001.

[9] T. Zhu, S.C. Zheng, D.S. Paul, S. Horvath, A.E. Teschendorff, Cell and tissue type independent age-associated DNA methylation changes are not rare but common, Aging. 10 (2018) 3541–3557. https://doi.org/10.18632/aging.101666.

[10] R.C. Slieker, C.L. Relton, T.R. Gaunt, P.E. Slagboom, B.T. Heijmans, Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception, Epigenetics Chromatin. 11 (2018) 25. https://doi.org/10.1186/s13072-018-0191-3.

[11] C.G. Bell, R. Lowe, P.D. Adams, A.A. Baccarelli, S. Beck, J.T. Bell, B.C. Christensen, V.N. Gladyshev, B.T. Heijmans, S. Horvath, T. Ideker, J.-P.J. Issa, K.T. Kelsey, R.E. Marioni, W. Reik, C.L. Relton, L.C. Schalkwyk, A.E. Teschendorff, W. Wagner, K. Zhang, V.K. Rakyan, DNA methylation aging clocks: challenges and recommendations, Genome Biol. 20 (2019) 249. https://doi.org/10.1186/s13059-019-1824-y.

[12] P.D. Fransquet, J. Wrigglesworth, R.L. Woods, M.E. Ernst, J. Ryan, The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis, Clin. Epigenetics. 11 (2019) 62. https://doi.org/10.1186/s13148-019-0656-7.

[13] S. Horvath, K. Raj, DNA methylation-based biomarkers and the epigenetic clock theory of ageing, Nat. Rev. Genet. 19 (2018) 371–384. https://doi.org/10.1038/s41576-018-0004-3.

[14] S. Horvath, C. Pirazzini, M.G. Bacalini, D. Gentilini, A.M. Di Blasio, M. Delledonne, D. Mari, B. Arosio, D. Monti, G. Passarino, F. De Rango, P. D'Aquila, C. Giuliani, E. Marasco, S. Collino, P. Descombes, P. Garagnani, C. Franceschi, Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring, Aging. 7 (2015) 1159–1170. https://doi.org/10.18632/aging.100861.

[15] G.M. Fahy, R.T. Brooke, J.P. Watson, Z. Good, S.S. Vasanawala, H. Maecker, M.D. Leipold, D.T.S. Lin, M.S. Kobor, S. Horvath, Reversal of epigenetic aging and immunosenescent trends in humans, Aging Cell. 18 (2019) e13028. https://doi.org/10.1111/acel.13028.

[16] WHO | Life expectancy, WHO. (n.d.). http://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends_text/en/ (accessed January 13, 2020).

[17] S.N. Austad, K.E. Fischer, Sex Differences in Lifespan, Cell Metab. 23 (2016) 1022–1033. https://doi.org/10.1016/j.cmet.2016.05.019.

[18] R. Ostan, D. Monti, P. Gueresi, M. Bussolotto, C. Franceschi, G. Baggio, Gender, aging and longevity in humans: an update of an intriguing/neglected scenario paving the way to a gender-specific medicine, Clin. Sci. Lond. Engl. 1979. 130 (2016) 1711–1725. https://doi.org/10.1042/CS20160004.

[19] Y. Zeng, C. Nie, J. Min, H. Chen, X. Liu, R. Ye, Z. Chen, C. Bai, E. Xie, Z. Yin, Y. Lv, J. Lu, J. Li, T. Ni, L. Bolund, K.C. Land, A. Yashin, A.M. O'Rand, L. Sun, Z. Yang, W. Tao, A. Gurinovich, C. Franceschi, J. Xie, J. Gu, Y. Hou, X. Liu, X. Xu, J.-M. Robine, J. Deelen, P. Sebastiani, E. Slagboom, T. Perls, E. Hauser, W. Gottschalk, Q. Tan, K. Christensen, X. Shi, M. Lutz, X.-L. Tian, H. Yang, J. Vaupel, Sex Differences in Genetic Associations With Longevity, JAMA Netw. Open. 1 (2018) e181670. https://doi.org/10.1001/jamanetworkopen.2018.1670.

[20] M. Inoshita, S. Numata, A. Tajima, M. Kinoshita, H. Umehara, H. Yamamori, R. Hashimoto, I. Imoto, T. Ohmori, Sex differences of leukocytes DNA methylation adjusted for estimated cellular proportions, Biol. Sex Differ. 6 (2015) 11. https://doi.org/10.1186/s13293-015-0029-7.

[21] P. Singmann, D. Shem-Tov, S. Wahl, H. Grallert, G. Fiorito, S.-Y. Shin, K. Schramm, P. Wolf, S. Kunze, Y. Baran, S. Guarrera, P. Vineis, V. Krogh, S. Panico, R. Tumino, A. Kretschmer, C. Gieger, A. Peters, H. Prokisch, C.L. Relton, G. Matullo, T. Illig, M. Waldenberger, E. Halperin, Characterization of whole-genome autosomal differences of DNA methylation between men and women, Epigenetics Chromatin. 8 (2015) 43. https://doi.org/10.1186/s13072-015-0035-3.

[22] P. Yousefi, K. Huen, V. Davé, L. Barcellos, B. Eskenazi, N. Holland, Sex differences in DNA methylation assessed by 450 K BeadChip in newborns, BMC Genomics. 16 (2015) 911. https://doi.org/10.1186/s12864-015-2034-y.

[23] S. Horvath, M. Gurven, M.E. Levine, B.C. Trumble, H. Kaplan, H. Allayee, B.R. Ritz, B. Chen, A.T. Lu, T.M. Rickabaugh, B.D. Jamieson, D. Sun, S. Li, W. Chen, L. Quintana-Murci, M. Fagny, M.S. Kobor, P.S. Tsao, A.P. Reiner, K.L. Edlefsen, D. Absher, T.L. Assimes, An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease, Genome Biol. 17 (2016) 171. https://doi.org/10.1186/s13059-016-1030-0.

[24] S.M. Tajuddin, D.G. Hernandez, B.H. Chen, N. Noren Hooten, N.A. Mode, M.A. Nalls, A.B. Singleton, N. Ejiogu, K.N. Chitrala, A.B. Zonderman, M.K. Evans, Novel age-associated DNA methylation changes and epigenetic age acceleration in middle-aged African Americans and whites, Clin. Epigenetics. 11 (2019) 119. https://doi.org/10.1186/s13148-019-0722-1.

[25] D.R. Masser, N. Hadad, H.L. Porter, C.A. Mangold, A. Unnikrishnan, M.M. Ford, C.B. Giles, C. Georgescu, M.G. Dozmorov, J.D. Wren, A. Richardson, D.R. Stanford, W.M. Freeman, Sexually divergent DNA methylation patterns with hippocampal aging, Aging Cell. 16 (2017) 1342–1352. https://doi.org/10.1111/acel.12681.

[26] S.L. Klein, K.L. Flanagan, Sex differences in immune responses, Nat. Rev. Immunol. 16 (2016) 626–638. https://doi.org/10.1038/nri.2016.90.

[27] H. Lotter, M. Altfeld, Sex differences in immunity, Semin. Immunopathol. 41 (2019) 133–135. https://doi.org/10.1007/s00281-018-00728-x.
[28] M.R. Gubbels Bupp, T. Potluri, A.L. Fink, S.L. Klein, The Confluence of Sex Hormones and Aging on Immunity, Front. Immunol. 9 (2018) 1269. https://doi.org/10.3389/fimmu.2018.01269.
[29] S. Horvath, DNA methylation age of human tissues and cell types, Genome Biol. 14 (2013) R115. https://doi.org/10.1186/gb-2013-14-10-r115.
[30] G. Hannum, J. Guinney, L. Zhao, L. Zhang, G. Hughes, S. Sadda, B. Klotzle, M. Bibikova, J.-B. Fan, Y. Gao, R. Deconde, M. Chen, I. Rajapakse, S. Friend, T. Ideker, K. Zhang, Genome-wide methylation profiles reveal quantitative views of human aging rates, Mol. Cell. 49 (2013) 359–367. https://doi.org/10.1016/j.molcel.2012.10.016.
[31] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K.A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: archive for functional genomics data sets--update, Nucleic Acids Res. 41 (2013) D991-995. https://doi.org/10.1093/nar/gks1193.
[32] A. Johansson, S. Enroth, U. Gyllensten, Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan, PloS One. 8 (2013) e67378. https://doi.org/10.1371/journal.pone.0067378.
[33] B. Lehne, A.W. Drong, M. Loh, W. Zhang, W.R. Scott, S.-T. Tan, U. Afzal, J. Scott, M.-R. Jarvelin, P. Elliott, M.I. McCarthy, J.S. Kooner, J.C. Chambers, A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies, Genome Biol. 16 (2015) 37. https://doi.org/10.1186/s13059-015-0600-x.
[34] D. Palli, F. Berrino, P. Vineis, R. Tumino, S. Panico, G. Masala, C. Saieva, S. Salvini, M. Ceroti, V. Pala, S. Sieri, G. Frasca, M.C. Giurdanella, C. Sacerdote, L. Fiorini, E. Celentano, R. Galasso, A. Decarli, V. Krogh, EPIC-Italy, A molecular epidemiology project on diet and cancer: the EPIC-Italy Prospective Study. Design and baseline characteristics of participants, Tumori. 89 (2003) 586–593.
[35] G. Campanella, S. Polidoro, C. Di Gaetano, G. Fiorito, S. Guarrera, V. Krogh, D. Palli, S. Panico, C. Sacerdote, R. Tumino, P. Elliott, G. Matullo, M. Chadeau-Hyam, P. Vineis, Epigenetic signatures of internal migration in Italy, Int. J. Epidemiol. 44 (2015) 1442–1449. https://doi.org/10.1093/ije/dyu198.
[36] W. Zhou, P.W. Laird, H. Shen, Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes, Nucleic Acids Res. 45 (2017) e22. https://doi.org/10.1093/nar/gkw967.
[37] Home | DNA Methylation Age Calculator, (n.d.). https://dnamage.genetics.ucla.edu/home (accessed January 13, 2020).
[38] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise, in: Proc. Second Int. Conf. Knowl. Discov. Data Min., AAAI Press, Portland, Oregon, 1996: pp. 226–231.
[39] M.J. Jones, S.A. Islam, R.D. Edgar, M.S. Kobor, Adjusting for Cell Type Composition in DNA Methylation Data Using a Regression-Based Approach, Methods Mol. Biol. Clifton NJ. 1589 (2017) 99–106. https://doi.org/10.1007/7651_2015_262.
[40] S. Horvath, P. Garagnani, M.G. Bacalini, C. Pirazzini, S. Salvioli, D. Gentilini, A.M. Di Blasio, C. Giuliani, S. Tung, H.V. Vinters, C. Franceschi, Accelerated epigenetic aging in Down syndrome, Aging Cell. 14 (2015) 491–495. https://doi.org/10.1111/acel.12325.
[41] F.J. McNally, A. Roll-Mecak, Microtubule-severing enzymes: From cellular functions to molecular mechanism, J. Cell Biol. 217 (2018) 4057–4069. https://doi.org/10.1083/jcb.201612104.
[42] A. Mok, O. Solomon, R.R. Nayak, P. Coit, H.L. Quach, J. Nititham, A.H. Sawalha, L.F. Barcellos, L.A. Criswell, S.A. Chung, Genome-wide profiling identifies associations between lupus nephritis and differential methylation of genes regulating tissue hypoxia and type 1 interferon responses, Lupus Sci. Med. 3 (2016) e000183. https://doi.org/10.1136/lupus-2016-000183.
[43] C. Behl, Breaking BAG: The Co-Chaperone BAG3 in Health and Disease, Trends Pharmacol. Sci. 37 (2016) 672–688. https://doi.org/10.1016/j.tips.2016.04.007.
[44] E. Stürner, C. Behl, The Role of the Multifunctional BAG3 Protein in Cellular Protein Quality Control and in Disease, Front. Mol. Neurosci. 10 (2017) 177. https://doi.org/10.3389/fnmol.2017.00177.
[45] R. Higuchi-Sanabria, P.A. Frankino, J.W. Paul, S.U. Tronnes, A. Dillin, A Futile Battle? Protein Quality Control and the Stress of Aging, Dev. Cell. 44 (2018) 139–163. https://doi.org/10.1016/j.devcel.2017.12.020.
[46] N. Chondrogianni, K. Voutetakis, M. Kapetanou, V. Delitsikou, N. Papaevgeniou, M. Sakellari, M. Lefaki, K. Filippopoulou, E.S. Gonos, Proteasome activation: An innovative promising approach for delaying aging and retarding age-related diseases, Ageing Res. Rev. 23 (2015) 37–55. https://doi.org/10.1016/j.arr.2014.12.003.
[47] M. Duggan, B. Torkzaban, T.M. Ahooyi, K. Khalili, J. Gordon, Age-related neurodegenerative diseases, J. Cell. Physiol. 235 (2020) 3131–3141. https://doi.org/10.1002/jcp.29248.
[48] E. Bongen, H. Lucian, A. Khatri, G.K. Fragiadakis, Z.B. Bjornson, G.P. Nolan, P.J. Utz, P. Khatri, Sex Differences in the Blood Transcriptome Identify Robust Changes in Immune Cell Proportions with Aging and Influenza Infection, Cell Rep. 29 (2019) 1961-1973.e4. https://doi.org/10.1016/j.celrep.2019.10.019.
[49] S.T. Gal-Oz, B. Maier, H. Yoshida, K. Seddu, N. Elbaz, C. Czysz, O. Zuk, B.E. Stranger, H. Ner-Gaon, T. Shay, ImmGen report: sexual dimorphism in the immune system transcriptome, Nat. Commun. 10 (2019) 4295. https://doi.org/10.1038/s41467-019-12348-6.
[50] Y.-M. Yang, D. Sun, S. Kandhi, G. Froogh, J. Zhuge, W. Huang, B.D. Hammock, A. Huang, Estrogen-dependent epigenetic regulation of soluble epoxide hydrolase via DNA methylation, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 613–618. https://doi.org/10.1073/pnas.1716016115.
[51] W.B. Zigman, Atypical aging in Down syndrome, Dev. Disabil. Res. Rev. 18 (2013) 51–67. https://doi.org/10.1002/ddrr.1128.

[52] C.A. Mangold, B. Wronowski, M. Du, D.R. Masser, N. Hadad, G.V. Bixler, R.M. Brucklacher, M.M. Ford, W.E. Sonntag, W.M. Freeman, Sexually divergent induction of microglial-associated neuroinflammation with hippocampal aging, J. Neuroinflammation. 14 (2017) 141. https://doi.org/10.1186/s12974-017-0920-8.

[53] K.E. Fischer, N.C. Riddle, Sex Differences in Aging: Genomic Instability, J. Gerontol. A. Biol. Sci. Med. Sci. 73 (2018) 166–174. https://doi.org/10.1093/gerona/glx105.

Tables

Table 1. saDMPs common to all the 4 Infinium450k datasets analyzed

ID_REF	CHR	MAPINFO	REFGENE NAME	REFGENE GROUP	RELATION TO CPG ISLAND
cg01620164	2	164590272	FIGN	Body	N_Shelf
cg23256579	12	11002403	PPR4	TSS1500;Body	
cg27615582	12	11002411	PPR4	TSS1500; Body	
cg14079463	6	127796989	C6orf174	Body	Island
cg04580344	6	127797022	C6orf174	Body	Island
cg23928726	1	2344998	PEX10	TSS1500	N_Shore
cg17076667	10	121418050	BAG3	Body	

Supplementary Table 1. Characteristics of the Infinium450k datasets investigated in the present study

	GSE40279	GSE87571	EPIC	GSE55763
Number of subjects	656	729	1803	2670
Number of females	338	388	1114	860
Number of males	318	341	689	1810
Age range	19 – 101	14 – 94	34 – 74	35 – 75

Supplementary Table 2. Characteristics of the samples analysed by the EpiTYPER assay

	Healthy controls n=419	Centenarians n=49	Centenarians' offspring n=48	Down syndrome n=49
Males	n=198 25-98 years	n=15 100-105 years	n=15 58-84 years	n=25 22-63 years
Females	n=221 23-98 years	n=34 100-112 years	n=33 55-89 years	n=19 19-66 years

Figure legends

Figure 1. Identification of sex-specific age-associated differentially methylated positions (saDMPs). (A) The polygon approach used to identify saDMPs. The scatter plot reports methylation, expressed as beta values, respect to age in females (red) and males (blue). Allowed regions for females and males are highlighted in red and blue respectively, while the intersection between the allowed regions is highlighted in yellow. (B) Venn diagram of the number of saDMPs identified according to the polygon approach in the 4 datasets. (C) Scatter plots of saDMPs identified in all the 4 Infinium450k analyzed datasets. x axis represents the age of subjects, y axis the methylation levels expressed as beta values.

Figure 2 Validation of FIGN and PRR4 locus by EpiTYPER. (A) Methylation of CpG unit 9 in FIGN amplicon vs age. (B) Methylation of CpG unit 3 in PRR4 amplicon vs age. For each CpG unit , DNA methylation in controls (general population), centenarians, centenarian's offspring and persons with Down

syndrome is reported vs the age of the subjects. Males are in blue, females are in red. Linear regression between DNA methylation and age was calculated separately for males and females in control subjects and was reported in each plot.

Figure 3. Identification of sex-specific age-associated variably methylated positions (saVMPs) (A) The approach used to indentify saVMPs. The scatter plot reports methylation, expressed as beta values, respect to age in females (red) and males (blue). The best 5% and 95% percentiles fittings for females and males are shown as red and blue lines, respectively. (B) Some examples of saVMPs common to the 4 datasets GSE40279, GSE87571, EPIC, GSE55763. Lines represent best type of 5% and 95% percentiles fitting. X axis corresponds to age of subjects, Y axis to methylation levels.

Figure 4. (A) Number of epimutations (log scale) in dependence on age in females (red) and males (blue). (B) Shannon entropy for 4 considered datasets: GSE40279, GSE87571, EPIC, GSE55763.

Supplementary Figure 1. Histograms of the number of females (red) and males (blue) according to age in GSE40279, GSE87571, EPIC and GSE55763 datasets.

Supplementary Figure 2. Identification of sex-specific age-associated differentially methylated positions (saDMPs) when correcting for blood cell counts. (A) *Venn diagram of the number of saDMPs identified in the 4 datasets when the polygon approach was applied to residuals obtained after regressing out blood cell counts. (B) Scatter plots of saDMPs common to all the 4 Infinium450k datasets analyzed, when the polygon approach was applied to residuals. x axis represents the age of subjects, y axis the methylation levels expressed as beta values.*

Supplementary Figure 3. Enrichment (odds ratio) of genomic localizations for saDMPs calculated from beta values (A) or residuals (B).

Supplementary Figure 4. Validation of PRR4 locus by EpiTYPER. For each of the CpG units returned by the EpiTYPER assay, DNA methylation in controls (general population), centenarians, centenarian's offspring and persons with Down syndrome is reported vs the age of the subjects. Males are in blue, females are in red. Linear regression between DNA methylation and age was calculated separately for males and females in control subjects and was reported in each plot.

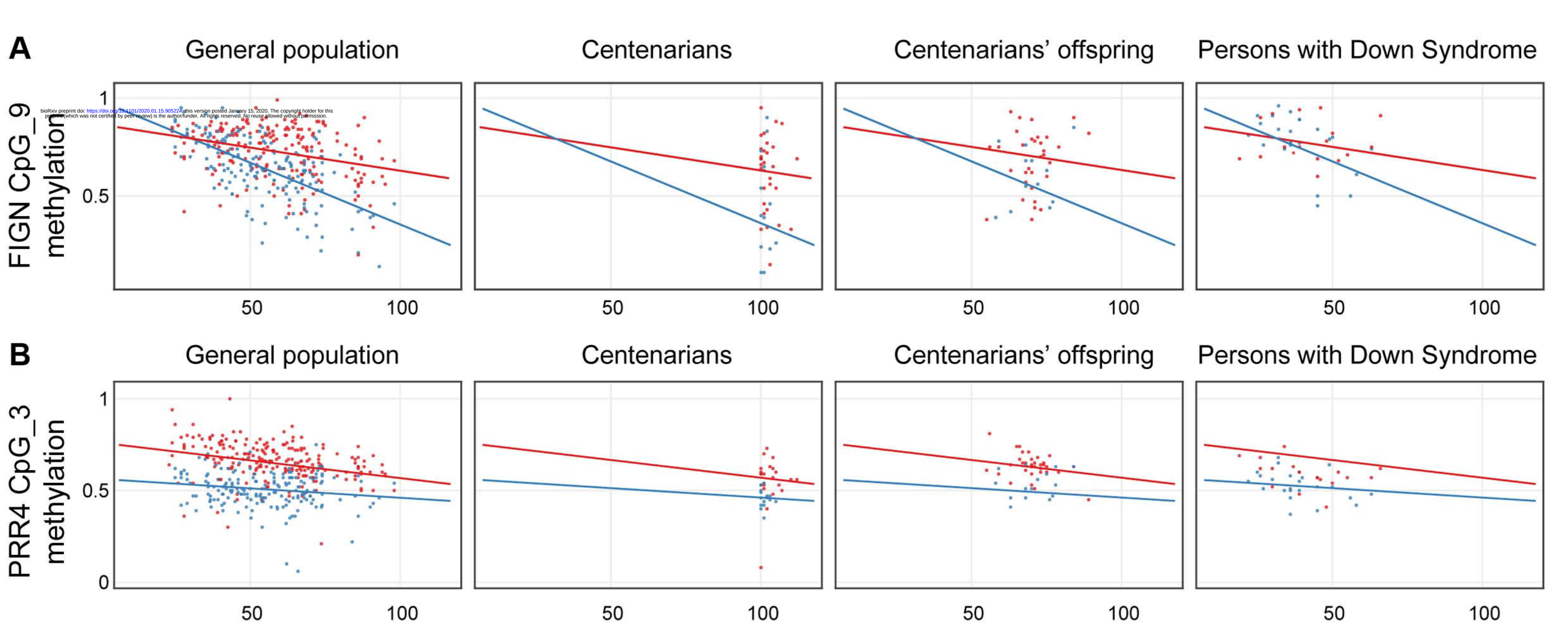
Supplementary Figure 5. Validation of PRR4 locus by EpiTYPER. For each of the 5 CpG units returned by the EpiTYPER assay, DNA methylation in controls (general population), centenarians, centenarian's offspring and persons with Down syndrome is reported vs the age of the subjects. Males are in blue, females are in red. Linear regression between DNA methylation and age was calculated separately for males and females in control subjects and was reported in each plot.

Supplementary Figure 6. Boxplots of DNA methylation for each CpG unit in FIGN amplicon in centenarians, centenarians' offspring and Down syndrome cohorts. Left panels: for each CpG unit in FIGN locus, boxplots of DNA methylation in centenarian males and females, compared to control (>80, < 100 years) males and females. Middle panels: for each CpG unit in FIGN locus, boxplots of DNA methylation in centenarians' offspring males and females, compared to age-matched control (>54, < 90 years) males and females. Right panels: for each CpG unit in FIGN locus, boxplots of DNA methylation in Down syndrome males and females, compared to age-matched control (>18, < 67 years) males and females.

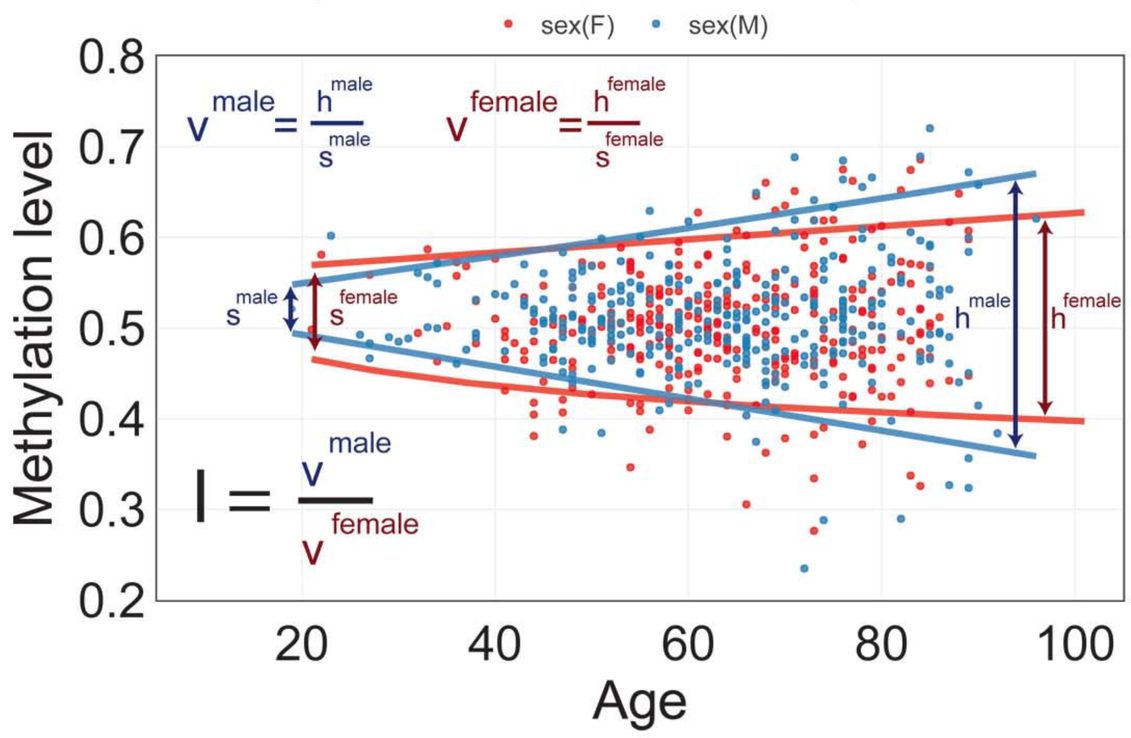
Supplementary Figure 7. Boxplots of DNA methylation for each CpG unit in PRR4 amplicon in centenarians, centenarians' offspring and Down syndrome cohorts. Left panels: for each CpG unit in PRR4 locus, boxplots of DNA methylation in centenarian males and females, compared to control (>80, < 100 years) males and females. Middle panels: for each CpG unit in PRR4 locus, boxplots of DNA methylation in centenarians' offspring males and females, compared to age-matched control (>54, < 90 years) males and females. Right panels: for each CpG unit in PRR4 locus, boxplots of DNA methylation in Down syndrome males and females, compared to age-matched control (>18, < 67 years) males and females.

Supplementary Figure 8. Density distribution of standard deviation values calculated in the GSE87571 dataset for 3 age classes, considering males and females together (A) or separated (B).

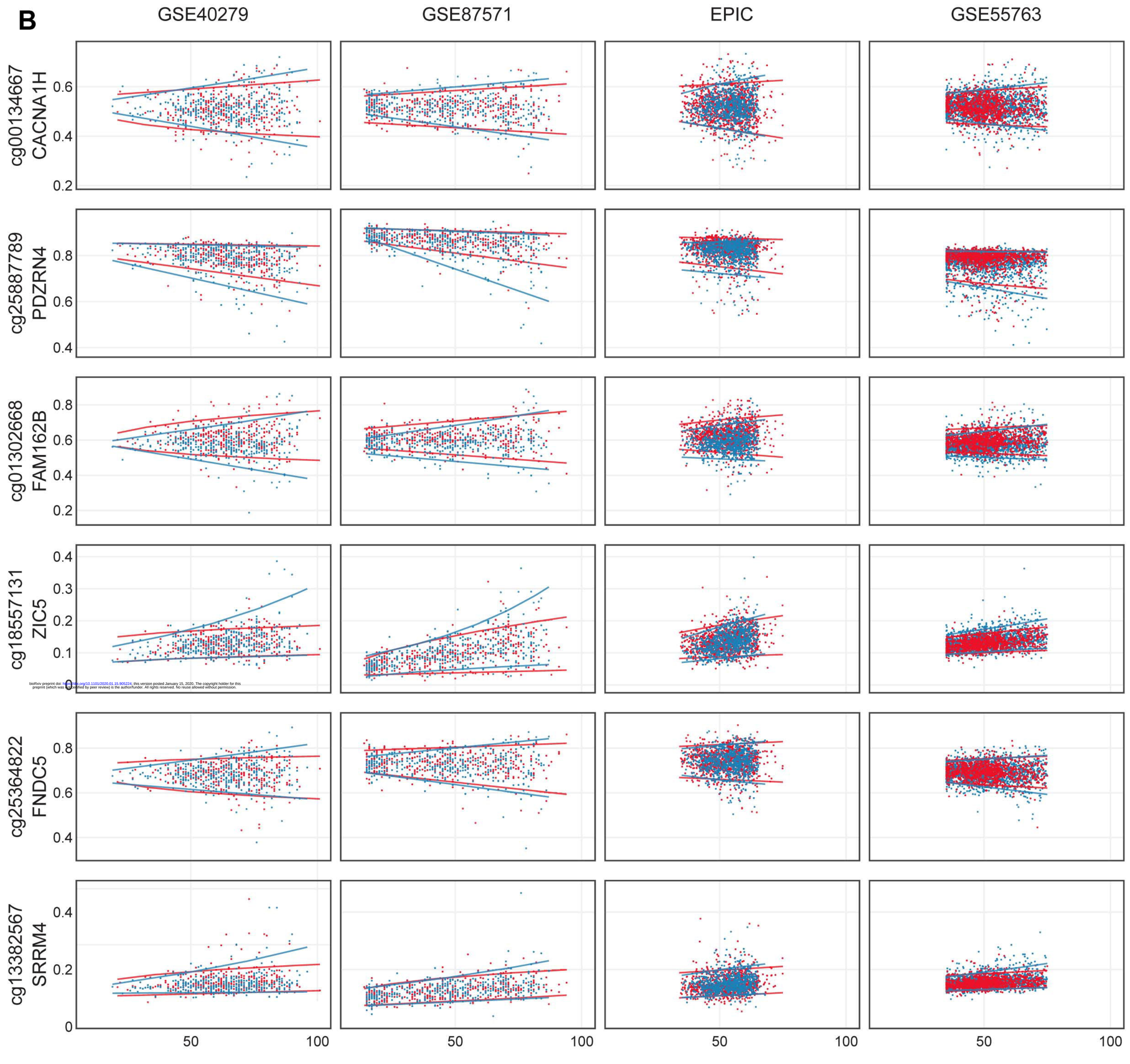
Supplementary Figure 9. Enrichment (odds ratio) of genomic localizations for saVMPs calculated from beta values (A) or residuals (B).

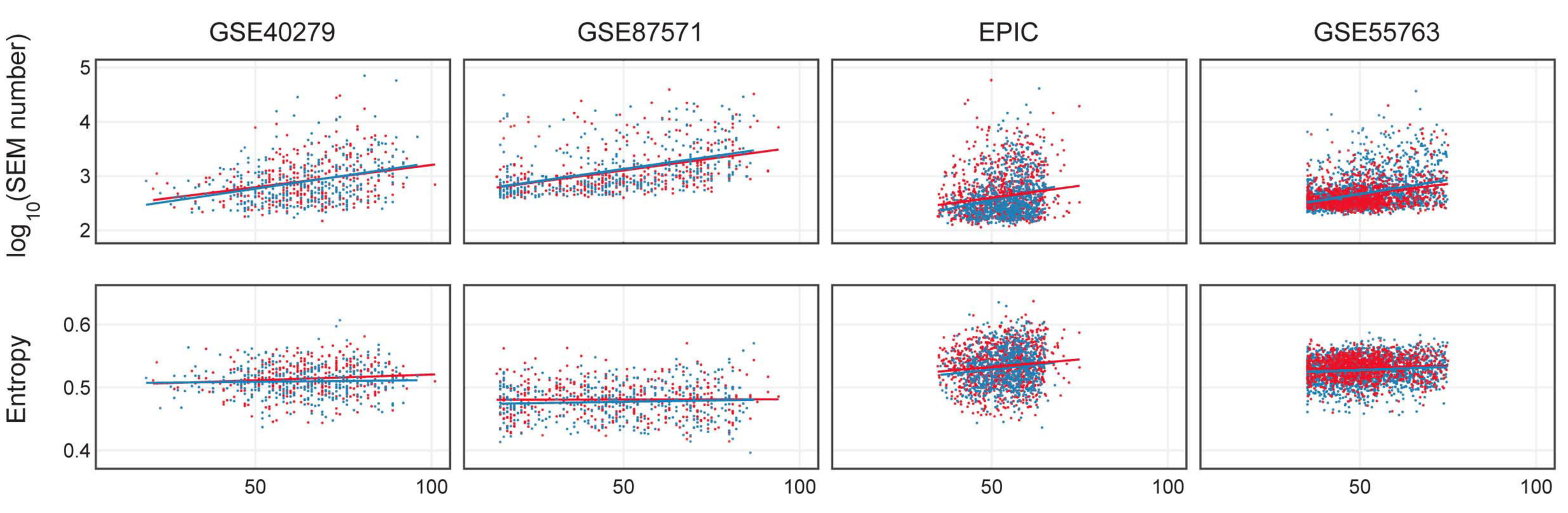


A



B




GSE40279
GSE87571
EPIC
GSE55763
log10(SEM number)
5
4
3
2
Entropy
0.6
0.5
0.4
50
100