

RESEARCH

Open Access



# Ageing-associated changes in DNA methylation in X and Y chromosomes

Laura Kananen<sup>1,2,3,4\*</sup> and Saara Marttila<sup>4,5\*</sup> 

## Abstract

**Background:** Ageing displays clear sexual dimorphism, evident in both morbidity and mortality. Ageing is also associated with changes in DNA methylation, but very little focus has been on the sex chromosomes, potential biological contributors to the observed sexual dimorphism. Here, we sought to identify DNA methylation changes associated with ageing in the Y and X chromosomes, by utilizing datasets available in data repositories, comprising in total of 1240 males and 1191 females, aged 14–92 years.

**Results:** In total, we identified 46 age-associated CpG sites in the male Y, 1327 age-associated CpG sites in the male X, and 325 age-associated CpG sites in the female X. The X chromosomal age-associated CpGs showed significant overlap between females and males, with 122 CpGs identified as age-associated in both sexes. Age-associated X chromosomal CpGs in both sexes were enriched in CpG islands and depleted from gene bodies and showed no strong trend towards hypermethylation nor hypomethylation. In contrast, the Y chromosomal age-associated CpGs were enriched in gene bodies, and showed a clear trend towards hypermethylation with age.

**Conclusions:** Significant overlap in X chromosomal age-associated CpGs identified in males and females and their shared features suggest that despite the uneven chromosomal dosage, differences in ageing-associated DNA methylation changes in the X chromosome are unlikely to be a major contributor of sex dimorphism in ageing. While age-associated CpGs showed good replication across datasets in the present study, only a limited set of previously reported age-associated CpGs were replicated. One contributor to the limited overlap are differences in the age range of individuals included in each data set. Further study is needed to identify biologically significant age-associated CpGs in the sex chromosomes.

**Keywords:** Ageing, DNA methylation, Sex chromosomes, X chromosome, Y chromosome

## Background

Ageing displays clear sexual dimorphism. Females have a longevity advantage, which can be observed across nationalities and in historical data, where and when reliable demographic information is available [1, 2]. Despite the longer lifespan, females have poorer health in later life, the so-called mortality–morbidity paradox [3, 4].

Epidemiology, pathophysiology, and symptoms of ageing-associated diseases, such as cardiometabolic diseases, cancer, and neurodegenerative diseases, can also differ between males and females [3]. The mechanisms underlying this sexual dimorphism remain poorly understood, but contributing factors include both biological as well as behavioral and societal [1, 3, 5].

Sex chromosomes are one of the biological factors contributing to sexual dimorphism in humans and other mammals [6]. The Y chromosome is small (57 million base pairs, bp) and relatively gene poor, with 566 genes, majority of which are pseudogenes. For comparison, chromosomes 22 and 19 (51 and 59 million bp,

\*Correspondence: [laura.kananen@tuni.fi](mailto:laura.kananen@tuni.fi); [saara.marttila@tuni.fi](mailto:saara.marttila@tuni.fi)

<sup>1</sup> Faculty of Social Sciences (Health Sciences), Tampere University, Tampere, Finland

<sup>5</sup> Department of Clinical Chemistry, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Full list of author information is available at the end of the article



respectively) contain 1376 and 2914 genes, respectively [7, 8]. The X chromosome is 156 million bp long and contains 2408 genes [7, 8]. Dosage compensation between XX females and XY males is achieved via X chromosome inactivation (XCI) in females, where one of the X chromosomes is packaged as heterochromatin and is nearly fully methylated. XCI occurs at random at an early developmental stage, leading to mosaic expression from both maternal and paternal X. In addition, a proportion of X chromosomal genes escape XCI [6].

Ageing-associated changes in DNA methylation have been widely studied (for example [9–12], see also references in “Discussion”), but in the great majority of studies, focus has been on the autosomes and sex chromosomes have been excluded. Here, we sought to identify DNA methylation changes associated with ageing in the Y and X chromosomes, by utilizing five datasets, all of whole blood, available in data repositories, comprising in total of 1240 males and 1191 females, aged 14–92 years.

**Results**

**Y chromosome**

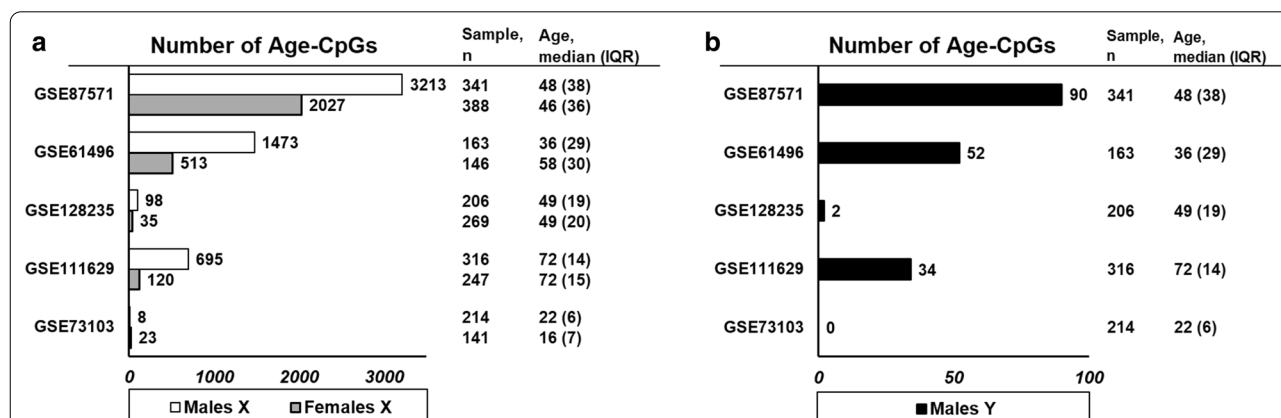
Of the five studied datasets, we identified age-associated CpG sites (hereafter referred as age-CpGs) in the Y chromosome in four datasets. The number of age-CpGs varied from 2 to 90 across datasets (Fig. 1). All Y chromosomal age-CpGs can be found from Additional file 1: Table S1A–D. There was considerable overlap between the datasets, as in all but one pairwise comparison, there was more overlap than can be expected by chance (Table 1). Of the identified CpG sites, 46 were age-associated in two or more datasets with corresponding

**Table 1** Overlap of age-CpGs between data sets used in the study

	GSE61496	GSE73103	GSE87571	GSE111629
<b>Male Y</b>				
GSE73103	na			
GSE87571	7.17e-9 (34)	na		
GSE111629	4.22e-7 (18)	na	4.15e-9 (26)	
GSE128235	0.030 (2)	na	ns (2)	0.013 (2)
<b>Male X</b>				
GSE73103	ns (2)			
GSE87571	3.07e-269 (1099)	ns (5)		
GSE111629	1.04e-164 (406)	ns (1)	3.71e-146 (553)	
GSE128235	5.15e-25 (60)	ns (1)	1.32e-33 (90)	1.09e-24 (44)
<b>Female X</b>				
GSE73103	0.007 (5)			
GSE87571	3.13e-51 (259)	ns (5)		
GSE111629	3.24e-16 (34)	ns (1)	1.50e-30 (84)	
GSE128235	1.29e-06 (11)	0.003 (2)	1.82e-8 (23)	3.21e-10 (9)

In the table, the hypergeometric test *p* value for the overlap and the number of common age-CpGs in parentheses are presented. *na* not available due to no significant age-CpGs in GSE73103, *ns* not statistically significant, hypergeometric test *p* value > 0.05

direction of association (Additional file 1: Table S1E). As there are 416 Y chromosomal probes on the Illumina 450 K array, 11% of probes were identified as age-associated. Two CpGs were identified in four datasets (Table 2) and 14 CpGs in three datasets.



**Fig. 1** Characteristics of the study populations used and number of age-CpGs identified. Age distributions in the datasets are also visualized as histograms in Additional file 5: Fig. S1. Notably, the number of age-CpGs identified in each dataset was associated with the age range in the dataset. The data sets are shown in ranked order according the age interquartile range (IQR). Number of identified age-CpGs in each data set in males and females and in X (A) and Y chromosome (B) is visualized as bars. The number was higher in the X chromosome, as compared to the Y chromosome, reflecting the number of probes available for each chromosome on the Illumina 450 K array (11648 X chromosomal probes, 416 Y chromosomal probes). *n* = number of individuals in the data set

**Table 2** Age-CpGs identified in four or more datasets in one sex (\*only cg25140188 identified in all five datasets studied). In boldface, age-CpGs identified in both females and males (in two or more datasets for the other sex)

	ID	Change with age	Gene (hg19)	Median $\beta$ , males	Median $\beta$ , females	CpG island
Y	cg27214488	Hyper	<i>NLGN4Y</i>	0.12		Shore
	cg20401549	Hypo	<i>PRKY</i>	0.07		Island
Male X	<b>cg12737514</b>	<b>Hyper</b>	<b><i>BCOR</i></b>	0.27	0.18	Island
	cg16140952	Hyper	<i>CHRD1</i>	0.14		Shore
	cg21480420	Hyper	<i>FAM70A</i>	0.27		Shore
	cg06711837	Hyper	<i>FGF13</i>	0.15		Shore
	cg08783090	Hyper	<i>GABRE</i>	0.19		Island
	cg01222180	Hyper	<i>GPR50</i>	0.28		Island
	cg21672057	Hyper	<i>KIAA2022</i>	0.27		Island
	cg01289637	Hyper	<i>LRCH2</i>	0.09		Shore
	cg24823082	Hyper	<i>MID1IP1</i>	0.08		Shelf
	<b>cg26481961</b>	<b>Hyper</b>	<b><i>MXRA5</i></b>	0.19	0.17	Island
	cg00575851	Hyper	<i>NCRNA00087</i>	0.15		Shore
	cg12848223	Hyper	<i>NRK</i>	0.30		Shore
	cg14951132	Hyper	<i>NRK</i>	0.28		Island
	cg09935073	Hyper	<i>PAK3</i>	0.21		Open sea
	cg13775533	Hyper	<i>PCDH19</i>	0.37		Island
	cg05134041	Hyper	<i>POU3F4</i>	0.30		Shore
	cg09610569	Hyper	<i>SLITRK2</i>	0.11		Island
	cg10288278	Hyper	<i>SLITRK4</i>	0.30		Island
	<b>cg01538344</b>	<b>Hyper</b>	<b><i>TAF7L</i></b>	0.46	0.49	Island
	cg27326620	Hyper	<i>ZC4H2</i>	0.43		Open sea
	cg00525383	Hyper	<i>ZIC3</i>	0.23		Island
	cg03671371	Hyper		0.18		Island
	<b>cg03718079</b>	<b>Hyper</b>		0.18	0.35	Island
	<b>cg17156570</b>	<b>Hyper</b>		0.38	0.51	Shore
	<b>cg03660876</b>	<b>Hypo</b>	<b><i>ARHGAP6</i></b>	0.74	0.80	Open sea
	<b>cg18422972</b>	<b>Hypo</b>	<b><i>BCOR</i></b>	0.52	0.32	Island
	cg06758848	Hypo	<i>FLNA</i>	0.33		Island
	cg09761247	Hypo	<i>IDS</i>	0.19		Shore
	<b>cg00168417</b>	<b>Hypo</b>	<b><i>KIAA1210</i></b>	0.73	0.69	Shore
	cg03675615	Hypo	<i>PAGE4</i>	0.91		Open sea
	<b>cg25863147</b>	<b>Hypo</b>	<b><i>PDK3</i></b>	0.65	0.63	Shore
	<b>cg09641151</b>	<b>Hypo</b>	<b><i>PLXNA3</i></b>	0.67	0.44	Shore
	<b>cg04532200</b>	<b>Hypo</b>	<b><i>PLXNB3</i></b>	0.62	0.64	Shore
	cg19406003	Hypo	<i>PLXNB3</i>	0.42		Shelf
	cg18478786	Hypo	<i>SLC25A5</i>	0.11		Island
	cg21019788	Hypo	<i>SUV39H1</i>	0.28		Island
	<b>cg06461462</b>	<b>Hypo</b>	<b><i>TSC22D3</i></b>	0.66	0.75	Shore
	cg21900764	Hypo	<i>UPF3B</i>	0.09		Island
	cg19570406	Hypo	<i>ZFP92</i>	0.71		Shore
	<b>cg25140188*</b>	<b>Hypo</b>		0.44	0.46	Shelf
Female X	<b>cg01538344</b>	<b>Hyper</b>	<b><i>TAF7L</i></b>	0.46	0.49	Island
	<b>cg08704539</b>	<b>Hyper</b>	<b><i>TAF7L</i></b>	0.58	0.55	Island
	cg07811386	Hypo	<i>CA5B</i>		0.35	Open sea
	<b>cg00168417</b>	<b>Hypo</b>	<b><i>KIAA1210</i></b>	0.73	0.69	Shore
	cg05204193	Hypo	<i>MAP3K7IP3</i>		0.49	Shelf
	<b>cg04532200</b>	<b>Hypo</b>	<b><i>PLXNB3</i></b>	0.62	0.64	Shore
	cg02991082	Hypo	<i>TSC22D3</i>		0.10	Shore
	<b>cg06461462</b>	<b>Hypo</b>	<b><i>TSC22D3</i></b>	0.66	0.75	Shore
	cg17938879	Hypo			0.37	Open sea

**Table 2** (continued)

Median methylation level across all datasets is given for both males and females

Majority, 76%, of the Y chromosomal age-CpGs identified in two or more datasets were hypermethylated. Their chromosomal distribution in relation to CpG islands was not different to what would be expected by chance. Regarding gene regions, we detected an enrichment of age-CpG sites in gene bodies and a depletion in intergenic regions (Additional file 1: Table S1F). Location of age-CpGs across the Y chromosome is presented in Additional file 5: Fig. S3. The 46 CpG sites were located in 18 genes (Additional file 1: Table S1E), but these genes were not enriched in any GO term categories, when Y chromosomal genes that were represented by one or more probes in the Illumina HumanMethylation450 BeadChip were used as background.

#### X chromosome in males

We detected age-CpGs in the X chromosome in males in all five data sets studied, and the number of age-CpGs varied from 8 to 3213 across datasets (Fig. 1). All age-CpGs can be found from Additional file 2: Table S2A-E. There was considerable overlap between the age-associated CpGs in each dataset, as there was more overlap than can be expected by chance in the majority pairwise comparisons, with the exception of GSE73103, where we identified only 8 age-CpGs (Table 1). Of the identified CpGs, 1327 were identified as age-associated in two or more datasets with corresponding direction of association (Additional file 2: Table S2F). As there are 11 233 X chromosomal probes on the Illumina 450 K array, 12% of probes were identified as age-associated. Of these, one CpG (cg25140188) was identified in all five datasets and 39 CpGs were identified in four datasets (Table 2).

The identified age-CpGs were both hyper- (47%) and hypomethylated (53%) with age. The age-CpGs were enriched in CpG islands and depleted from open sea and island shelves. Regarding gene regions, the age-CpGs were enriched in 1st Exon and 5'UTR but depleted from gene body, 3'UTR and from intragenic regions (Additional file 2: Table S2G). Location of age-CpGs across the X chromosome is presented in Additional file 5: Fig. S3. The 1327 age-CpGs were located in 491 genes (Additional file 2: Table 2F) and these genes were enriched in 47 GO terms, when X chromosomal genes that were represented by one or more probes in the Illumina HumanMethylation450 BeadChip were used a background (Additional file 2: Table 2H). For the most part, the identified GO terms were associated with cellular component organization and signal transduction.

#### X chromosome in females

We detected age-CpGs in the X chromosome in females in all five datasets studied, and the number of age-CpGs varied from 23 to 2027 across datasets (Fig. 1). All age-CpGs can be found from Additional file 3: Table S3A-E. There was considerable overlap between the age-associated CpGs in each dataset, as there was more overlap than can be expected by chance in the majority pairwise comparisons, with the exception of GSE73103 (Table 1). Of the identified CpGs, 325 were identified in two or more datasets with corresponding direction of association (Additional file 3: Table S3F). As there are 11 233 X chromosomal probes on the Illumina 450 K array, 3% of probes were identified as age-associated. Of these, 9 CpGs were identified in four datasets (Table 2) and 29 in three datasets.

The identified age-CpGs were both hyper- (54%) and hypomethylated (46%) with age. As compared to males, the proportion of hypermethylated age-CpGs was higher and this difference was statistically significant ( $p < 0.05$ ). The age-CpGs were enriched in CpG islands and depleted in open sea and CpG island shelves. In relation to genes, the age-CpGs were enriched in TSS200 and depleted in gene body (Additional file 3: Table S3G). Location of age-CpGs across the X chromosome is presented in Additional file 5: Fig. S3. The 325 age-CpGs were located in 199 genes (Additional file 3: Table S3F), and these were enriched in two GO component terms, when X chromosomal genes that were represented by one or more probes in the Illumina HumanMethylation450 BeadChip were used as background (GO0044464, *cell part*, FDR  $q$  value = 0.003; GO0016020, *membrane*, FDR  $q$  value = 0.003).

#### Overlap between males and females in the X chromosome

Of the 1327 and 325 X chromosomal age-CpGs identified in males and females, respectively, 122 were identified in both sexes with corresponding direction of association and additional 5 were identified in both sexes but with opposite direction of association (Additional file 4: Table S4A). Overlap of CpGs with comparable direction of association was more than would be expected by chance (hypergeometric test,  $p$  value =  $1.1 \times 10^{-35}$ ). Four CpGs (cg00168417, cg01538344, cg04532200, and cg06461462) were identified in four datasets in both females and males (Table 2).

Of the age-CpGs identified in both sexes, 57% were hypermethylated with age. The age-CpGs common for males and females were depleted from open sea as well

as from gene bodies (Additional file 4: Table S4B). The 122 age-CpGs were located in 84 genes (Additional file 4: Table S4A), that were enriched in one GO process term (GO0071840: *cellular component organization or biogenesis*, FDR  $q$  value = 0.049), when X chromosomal genes that were represented by one or more probes in the Illumina HumanMethylation450 BeadChip were used as background.

## Discussion

To gain insight on age-associated changes in DNA methylation in sex chromosomes, we utilized whole blood methylation data available in data repositories and report results from five individual datasets, comprising of ~2400 individuals aged 14–92 years.

The age-CpGs identified in the five datasets showed significant overlap across the datasets. In X in both sexes, and in Y in males, the number of age-CpGs was dependent on the age range of the sample; the wider the range, the more age-CpGs were detected (Fig. 1). In the X chromosome, we identified more significant age-CpGs in males as compared to females. Despite the difference in number of age-CpGs and the uneven chromosome dosage, the X chromosomal age-CpGs showed significant overlap between females and males and displayed similar characteristics in terms of genomic locations and direction of change. These age-CpGs were enriched in CpG islands and towards the 5' end of genes but depleted from gene bodies. For the X chromosome, we did not observe a strong trend towards hyper- or hypomethylation. The proportion of hypermethylated age-CpGs was higher in females as compared to males, but this difference can be affected by the different number of age-CpGs identified in females and males (325 and 1327, respectively). Age-CpGs in the Y chromosome, however, were predominantly hypermethylated with age and contrary to the X chromosomal age-CpGs, were enriched in gene bodies.

Ageing is associated with a global hypomethylation of the genome [13], and in the autosomes, majority of studies on age-associated changes in DNA methylation have identified more hypomethylated age-associated sites as compared to hypermethylated [10, 14–21]. Notably, some studies do report more hypermethylated age-associated sites [22–25]. For the X chromosome, Li et al. (2020) [26] reported a trend towards hypermethylation for female-specific age-associated CpGs and inconsistent results for male-specific CpGs. In our own results, we also identified more hypermethylated age-CpGs in females, but the trend was not as strong as reported by Li et al. [26]. For age-associated X chromosomal probes identified in both sexes, they reported a strong trend towards hypermethylation (78–94% of identified CpGs hypermethylated), similar to the more modest trend of hypermethylation we

observed for X chromosomal age-CpGs observed in both sexes (57%). For the Y chromosome, Lund et al. (2020) [27] reported a strong trend towards hypermethylation with age (>82% hypermethylated), and we observe a similar trend in our analyses (76% hypermethylated). This suggests that the age-associated changes in DNA methylation especially in the Y chromosome are distinct from those in the autosomes. However, as the number of Y chromosomal probes on the Illumina 450 K array is considerably lower as compared to X chromosomal probes (416 versus 11 233 probes, respectively), these results should be interpreted with caution.

GO term enrichment analysis of the sex-chromosomal age-CpGs did not associate these CpGs with functions that are known to be strongly associated with the molecular mechanisms of ageing [28]. The individual age-CpGs identified were located in genes with various functionalities. Of the individual age-CpGs, cg25140188 was identified as age-associated in males in all five datasets studied, in females in two of the datasets studied, and previously by McCartney et al. (2020) [29]. However, this CpG is not annotated with a gene. Four X chromosomal CpG sites were identified as age-associated in four datasets in both females and males, and each of these were located within a protein coding gene. Two of these, *KIAA1210* (cg13498184) and *TAF7L* (TATA-box-binding protein associated factor 7 like) (cg01538344), are mainly expressed in testes and have been implicated to have a role in spermatogenesis [30, 31]. *TAF7L* is also classified as a cancer-testis antigen (CTAg), belonging to a group of genes typically expressed only in the germline but also in malignant tumours [32]. Two additional CpG sites located in *TAF7L* were also identified as age-CpGs in at least two datasets in both females and males included in this study (Additional file 4: Table S4A). *PLXNB3* (plexin B3) (cg04532200) is highly expressed in the brain and plays a role in axon guidance [33]. *TSC22D3* (TSC22 domain family member 3) (cg06461462) regulates T-cell activation and has immunosuppressive and anti-inflammatory effects [34]. Dysregulation of the immune system, the so-called immunosenescence, is an important feature of the ageing phenotype [35].

The two Y chromosomal CpGs that were identified in four of the five studied datasets are located in *PRKY* (cg20401549) and *NLGN4Y* (cg27214488). One additional CpG site located in *PRKY* and seven additional CpG sites located in *NLGN4Y* were also identified as age-CpGs in at least two datasets included in this study (Additional file 1: Table S1E). *PRKY* is a protein kinase pseudogene, and there is a similar gene in the pseudoautosomal region of the X chromosome. Interestingly, three a-CpGs identified in both females and males were located in *NLGN4X* (neuroligin 4 X-linked) (Additional file 4: Table S4A).

*NLGN4X* and *NLGN4Y* share 97% sequence identity, but a difference of only one amino acid leads to deficit in *NLGN4Y* function, as compared to *NLGN4X* [36]. Neuroligins are associated with neuronal development and synaptic transmission, and they have been linked with autism spectrum disorders and intellectual disability [37]. In addition, *NLGN4Y* has been suggested to be associated with male homosexuality [38] and prostate cancer [39]. A CpG (cg01707559) identified previously by Lund et al. [27] and replicated in the present study is located in *TBLIY* (transducin beta like 1 Y-linked), a gene that is highly expressed in the prostate and shares great similarity with *TBLIX* in the X chromosome. We also identified two additional a-CpGs located in *TBLIY* in the present study (Additional file 1: Table S1E).

In most previous studies on age-associated DNA methylation changes, sex chromosomes have been excluded from the analyses [9, 10, 14–18, 23, 24, 40–44]. Other studies, while not explicitly stating whether or not sex chromosomes have been included in the analyses, report no significant findings in X or Y chromosomes related to age [11, 12, 19, 21, 22, 25, 45]. However, while majority of studies have not explored this, it should be noted that also autosomes can display sexually dimorphic age-associated changes in DNA methylation [46].

Only a handful of studies, that have not specifically focused on sex chromosomes, including only women, report significant age-associated findings in the X chromosome [20, 47, 48]. Each of these studies report five or less age-associated CpGs in the X chromosome, none of which are replicated between these three studies. Of these, one CpG site (cg27250462) reported by Teschendorff et al. [20] is replicated in a study focusing on age-associated methylation changes in the X chromosome (Li et al. 2020) [26], and one CpG site (cg13277716) reported by Xu & Taylor (2014) [48] is replicated in both females and males in three of the datasets analyzed in this study (GSE111629, GSE61496, GSE87571).

Some studies on age-associated DNA methylation changes have included X chromosomal probes in the analyses, but processed and analyzed female and male samples together [29, 49, 50]. Of these, Jansen et al. (2019) [50] report no significant age associations in the X chromosome, whereas Kim et al. (2014) [49] report an ageing-signature that is highly enriched to the X chromosome. As the statistical analyses were not sex-stratified, the findings might be biased due to differential methylation profile in female and male X chromosomes. In the study by McCartney et al. (2020) [29], sex and sex\*age interaction were included in the analysis, and among their high-confidence age-associated CpGs, they report 5 sex-independent and 6 sex-dependent X chromosomal sites. Of the sex-independent CpGs, one (cg25140188)

was identified as age-CpG in both females and males in this study, and in males in all of the five datasets studied. Of the sex-dependent CpGs reported by McCartney et al. (2020) [29], one CpG (cg20202246) was identified as age-CpG in both females and males and another CpG (cg08814148) only in males.

Only a few studies specifically focusing on age-associated DNA methylation changes in X or Y chromosomes have been published. Cotton et al. (2015) [51] reported no age-associated CpGs in the X chromosome using a small dataset (n = 111). More recently, Li et al. (2020) [26] have studied ageing-associated changes in DNA methylation in the X chromosome in two discovery datasets, and a third replication dataset. They identified 27 CpGs in females and 19 CpGs in males as age-associated in all three datasets in their study. In our study, in females, none of the 27 CpGs reported by Li et al. [26] were identified as age-CpGs, whereas in males, 4 CpGs out of the 19 reported by Li et al. were identified as age-CpGs with similar direction of change with age (Table 3).

One study only has previously focused specifically on ageing-associated DNA methylation changes in the Y chromosome (Lund et al. 2019) [27]. In total, they utilized four datasets, and report 7 CpGs that were consistently hypermethylated with age across all data sets. Of these, one (cg01707559), located in *TBLIY*, is replicated in our study in three datasets (GSE111629, GSE61496, GSE87571).

In genome-wide DNA methylation studies on autosomes, at the level of individual CpGs, the results have not shown great replication from study to study. Only a small number of CpG sites, for example those located in *ELOVL2*, *FHL2*, and *EDARADD*, have been replicated in majority of studies [10, 20, 25, 40, 42, 45, 48, 52]. This pattern of small accordance between studies is similar for the sex chromosomes, as a very limited number of CpGs are identified as age associated in more than one study. Presumably, both sample and dataset characteristics as well as data processing and analysis methods contribute to this discrepancy.

Specifically, age distribution within and between datasets likely influences the replication of age-CpGs. In our

**Table 3** Age-CpGs in male X chromosome identified in Li et al. (2020) [26] and in the present study, all hypermethylated with age

ID	Data sets in this study	Gene
cg05070565	GSE61496, GSE87571	<i>WAS</i>
cg15222604	GSE111629, GSE61496, GSE87571	<i>HCFC1</i>
cg18086582	GSE61496, GSE87571	<i>WDR13</i>
cg21120047	GSE61496, GSE87571	<i>XPNPEP2</i>

analysis, the number of age-CpGs was dependent on the age distribution of the dataset, with higher the IQR for age, the higher the power was to detect significant age-CpGs (Fig. 1). Previously, variation in DNA methylation has been shown to increase with age for both autosomes [42, 53], and the Y chromosome [27]. Modest variation in DNA methylation at younger adult ages, and low IQR for age, may explain why we did not identify many changes in dataset GSE73103. Furthermore, the study by Li et al. (2020) [26] consisted of older individuals aged > 55 years, which may in part explain the low accordance with our results of X chromosomal age-CpGs.

To decipher to what extent data processing methods could explain the limited overlap between our results and those reported by Li et al. (2020) [26], we repeated our analysis in a data set with the highest number of age-CpGs (GSE87571) as Li et al. [26] had done. While both age-association analyses were performed for X chromosomal probes only, Li et al. [26] excluded autosomal probes before normalization, whereas in our study, the autosomal probes were excluded from the analysis only after normalization. However, this did not change our results for dataset GSE87571, and the overlap with Li et al. remained limited, further suggesting that dataset and sample characteristics might explain the observed discrepancies. Further details of this additional analysis can be found from Additional file 5.

The sex chromosomes can also undergo more large-scale changes with age, which could affect the observed DNA methylation pattern. Both loss of Y (LOY) and loss of X (LOX) have been associated with ageing [54, 55]. Especially, LOX could affect the observed DNA methylation pattern, as the fully methylated, inactive X is more often affected by LOX [55]. In the present study, we cannot exclude that some of the age-associated changes observed are actually due to LOX. Skewing in the proportion of inactivated X has also been suggested to be associated with ageing [56–58]. Several papers catalogue human X chromosomal genes escaping from XCI [51, 59, 60]. The X chromosomal age-CpGs in females in this study mapped to 199 protein coding genes. Of these, 48 were classified as escaping XCI in these three studies [51, 59, 60], so in total, 25% of genes harboring age-CpGs in this study have been implicated to escape XCI. Total percentage of X chromosomal genes escaping XCI is considered to be approximately one third of all X chromosomal genes [51, 60], thus suggesting that genes escaping XCI are not overrepresented among those harboring age-associated DNA methylation changes. However, Li et al. (2020) [26] reported a much lower proportion of XCI escape, approximately 6% of all genes in the X chromosome (648 CpGs located in 133 genes). Of the 232 age-CpGs in females identified in this study, 55 (17%) were

described as escapees by Li et al. (2020) [26], suggesting an overrepresentation of escapees among the genes harboring age-CpGs. Taken together, age-associated DNA methylation changes in the X chromosome occur in both XCI escaping genes and those that remain inactivated.

A limitation of our study is that as we used data from a data repository, we had only limited phenotypic information that could be adjusted for in the analysis. Adjusting for confounding factors such as disease status or lifestyle factors could increase the sensitivity of the analysis to identify age-CpGs. We did adjust for cell-type proportions, as these can be estimated from the methylation data itself, and they have been shown to explain majority of variation in DNA methylation [61].

## Conclusion

Taken together, we report here ageing-associated changes in DNA methylation in both X and Y chromosomes. Our results show the similarities in age-associated DNA methylation changes in females and males and suggest that they share features with DNA methylation changes observed in the autosomes. Significant overlap in X chromosomal age-CpGs identified in males and females and their shared features suggest that despite the uneven chromosomal dosage, differences in ageing-associated DNA methylation changes in the X chromosome are unlikely to be a major contributor of sex dimorphism in ageing. Age-associated DNA methylation changes in the Y chromosome appear to differ from those observed in the autosomes, but due to technical reasons, the results should be interpreted with caution.

Very little has been published regarding the age-associated DNA methylation changes in the sex chromosomes and the published results show very limited overlap. Future study is needed to pinpoint analysis methods that yield biologically the most significant results. In addition, functionality of the observed DNA methylation changes in the sex chromosomes and more broadly throughout the genome requires more study. Analysis strategies, including for example pathway and enrichment analysis, that allow the combination of results from autosomes and sex chromosomes could possibly produce the most accurate picture of these changes. In addition, and especially for the Y chromosome, DNA methylation analysis with, for example, a bisulphite-sequencing method with better coverage could help to differentiate true biological phenomenon from technical bias.

## Materials and methods

### Datasets

Methylation data were downloaded from GEO NCBI database [62] using the following inclusion criteria: number of individuals in the dataset was >200, information

on age and sex was available, sample type was whole blood, DNA methylation was determined using Illumina HumanMethylation450 BeadChip, and raw data were available as idat-files. Data sets included were GSE61496 [63], GSE73103 [64], GSE87571 [15], GSE111629 [65], and GSE128235 [66]. Datasets are further described in Fig. 1 and in Additional file 5: Fig. S1.

GSE61496 consists of monozygotic twin pairs from Danish Twin Registry, who are discordant for birth weight. GSE73103 consists of healthy individuals recruited in Uppsala, Sweden. GSE87571 is a sample of The Northern Sweden Population Health Study in Sweden. GSE111629 is a sample of a population-based case-control study of Parkinson's disease in California, USA. GSE128235 is comprised of depressed and control subjects who were recruited at the Max Planck Institute of Psychiatry, Germany. Detailed information can be found under accession numbers at GEO NCBI database [62].

### Processing of methylation data

For each dataset, male and female samples were pre-processed separately using R software and minfi therein. Visual inspection of the data (Additional file 5: Fig. S2) shows the different distribution of methylation values for female and male X chromosome, consistent with XCI in females, demonstrating the rationale of analyzing male and female samples separately. Correct assignment of sex in the phenotypic data for each dataset was verified by the median methylation value of the X chromosome.

From each dataset, multimapping and cross-reactive probes, probes with detection p value median > 0.05 and probes with minor allele frequency of 1% or higher [67] were excluded. Annotation was hg19. The datasets were background corrected and quantile normalized in minfi, and dye bias was corrected with BMIQ in wateRmelon. Methylation profiles after preprocessing are shown in Additional file 5: Fig. S2.

Individual samples were excluded based on low methylated and unmethylated signals with a cut-off value for the average level at 10.5. With this criterion, we excluded 60 samples from GSE128235 (23 male samples, 37 female samples) and one male sample from both GSE61496 and GSE111629. Individual methylation level values were excluded when detection p value was more than 0.05. After preprocessing, the number of probes remaining for statistical testing across datasets was 9505–9513 for female X, 9717–9807 for male X, and 296–298 for male Y.

Methylation sites associated with age in X and Y chromosomes were identified within each data set using linear multivariate modeling. Methylation level,  $\beta$ -value, was set as the dependent variable and the model was adjusted for blood cell composition. The

blood cell composition was estimated using function EstimateCellCounts. Both twins in GSE61496 were included and considered as singletons in the analysis. Threshold for statistical significance within each data set was Benjamini–Hochberg-adjusted  $p$  value < 0.05.

### Characterizing, visualizing, and comparing datasets

Overlap between datasets in this study and other studies were visualized using webtool Venn [68]. Chromosomal ideograms were produced using PhenoGram software [69]. Statistical significance of overlap was analyzed with hypergeometric test. Enrichment of GO terms was analyzed with GOrilla [70, 71]. Only X and Y chromosomal genes that were represented by one or more probes in the Illumina HumanMethylation450 BeadChip were used as background in the analyses.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13072-021-00407-6>.

**Additional file 1:** Y chromosomal age-CpGs. Contains lists of Y chromosomal age-CpGs identified in each individual dataset (A–D), list of age-CpGs identified in multiple datasets (E) and location of identified Y chromosomal age-CpGs in relation to CpG islands and gene regions (F).

**Additional file 2:** X chromosomal age-CpGs identified in males. Contains lists of X chromosomal age-CpGs identified in males in each individual dataset (A–E), list of age-CpGs identified in multiple datasets (F), location of identified X chromosomal age-CpGs in relation to CpG islands and gene regions (G) and enriched GO terms for X chromosomal age-CpGs in males (H).

**Additional file 3:** X chromosomal age-CpGs identified in females. Contains lists of X chromosomal age-CpGs identified in females in each individual dataset (A–E), list of age-CpGs identified in multiple datasets (F) and location of identified X chromosomal age-CpGs in relation to CpG islands and gene regions (G).

**Additional file 4:** X chromosomal age-CpGs identified in both females and males. Contains list of X chromosomal age-CpGs identified in both females and males (A) and location of those age-CpGs in relation to CpG islands and gene regions (B).

**Additional file 5:** Figures and Supplementary analysis.

### Acknowledgements

The authors wish to acknowledge all the researchers who contributed towards the collection of datasets utilized in this study. The work was conducted in the framework of the Centre of Excellence in Research on Ageing and Care (CoEAgeCare).

### Authors' contributions

LK and SM designed the study, conducted the analyses, and wrote the manuscript. All authors read and approved the final manuscript.

### Funding

L.K. has been financially supported by the Emil Aaltonen Foundation and the Yrjö Jahnsson Foundation (Grant 20197181). S.M. has been financially supported by the Paulo Foundation, the Juho Vainio Foundation, and Yrjö Jahnsson Foundation. This study was also supported from the Academy of Finland through its funding to the Centre of Excellence in Research on Ageing and Care (CoEAgeCare, grants number 326567 and 336670).



### Availability of data and materials

The datasets supporting the conclusions of this article are available in the Gene Expression Omnibus (GEO, [www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) repository, under accession numbers GSE61496, GSE73103, GSE87571, GSE111629, and GSE128235.

### Declarations

#### Ethics approval and consent to participate

Datasets used in this study were retrieved from Gene Expression Omnibus (GEO, [www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) repository. All datasets are in compliance with the Declaration of Helsinki and have been approved by local ethical committees, and details can be found from the original publications, Tan et al. 2014 [63] (GSE61496), Voisin et al. 2015 [64] (GSE73103), Johansson et al. 2013 [15] (GSE87571), Chuang et al. 2017 [65] (GSE111629), and Zannas et al. 2019 [66] (GSE128235).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Faculty of Social Sciences (Health Sciences), Tampere University, Tampere, Finland. <sup>2</sup>Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. <sup>3</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. <sup>4</sup>Gerontology Research Center, Tampere University, Tampere, Finland. <sup>5</sup>Department of Clinical Chemistry, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.

Received: 26 April 2021 Accepted: 23 June 2021

Published online: 02 July 2021

### References

- Austad S, Fischer K. Sex differences in lifespan. *Cell Metab.* 2016;23(6):1022–33.
- WHO [https://www.who.int/data/gho/data/indicators/indicator-details/GHO/life-expectancy-at-birth-\(years\)](https://www.who.int/data/gho/data/indicators/indicator-details/GHO/life-expectancy-at-birth-(years)) Accessed 24.11.2020.
- Ostan R, Monti D, Guerresi P, Bussoletto M, Franceschi C, Baggio G. Gender, aging and longevity in humans: an update of an intriguing/neglected scenario paving the way to a gender-specific medicine. *Clin Sci (Lond).* 2016;130(19):1711–25.
- Van Oyen H, Nusselder W, Jagger C, Kolip P, Cambois E, Robine JM. Gender differences in healthy life years within the EU: an exploration of the “health-survival” paradox. *Int J Public Health.* 2013;58(1):143–55.
- Sampathkumar NK, Bravo JI, Chen Y, Danthi PS, Donahue EK, Lai RW, et al. Widespread sex dimorphism in aging and age-related diseases. *Hum Genet.* 2020;139(3):333–56.
- Snell DM, Turner JMA. Sex chromosome effects on male–female differences in mammals. *Curr Biol.* 2018;28(22):R1313–24.
- Human genome build GRCh38.p13
- Piovesan A, Pelleri MC, Antonaros F, Strippoli P, Caracausi M, Vitale L. On the length, weight and GC content of the human genome. *BMC Res Notes.* 2019;12(1):106.
- Bacalini MG, Boattini A, Gentilini D, Giampieri E, Pirazzini C, Giuliani C, et al. A meta-analysis on age-associated changes in blood DNA methylation: results from an original analysis pipeline for Infinium 450k data. *Aging (Albany NY).* 2015;7(2):97–109.
- Marttila S, Kananen L, Häyrynen S, Jylhävä J, Nevalainen T, Hervonen A, et al. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. *BMC Genomics.* 2015;16:179.
- Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res.* 2010;20(4):434–9.
- Tharakan R, Ubaida-Mohien C, Moore AZ, Hernandez D, Tanaka T, Ferrucci L. Blood DNA methylation and aging: a cross-sectional analysis and longitudinal validation in the InCHIANTI study. *J Gerontol A Biol Sci Med Sci.* 2020;75(11):2051–5.
- Xiao FH, Wang HT, Kong QP. Dynamic DNA methylation during aging: A “Prophet” of age-related outcomes. *Front Genet.* 2019;10:107.
- Fernández AF, Bayón GF, Urduñigo RG, Torano EG, García MG, Carella A, et al. H3K4me1 marks DNA regions hypomethylated during aging in human stem and differentiated cells. *Genome Res.* 2015;25(1):27–40.
- Johansson A, Enroth S, Gyllenstein U. Continuous aging of the human DNA methylome throughout the human lifespan. *PLoS ONE.* 2013;8(6):e67378.
- Kananen L, Marttila S, Nevalainen T, Jylhävä J, Mononen N, Kähönen M, et al. Aging-associated DNA methylation changes in middle-aged individuals: the Young Finns study. *BMC Genomics.* 2016;17:103.
- Li S, Christiansen L, Christensen K, Kruse TA, Redmond P, Marioni RE, et al. Identification, replication and characterization of epigenetic remodelling in the aging genome: a cross population analysis. *Sci Rep.* 2017;7(1):8183.
- McClay JL, Aberg KA, Clark SL, Nerella S, Kumar G, Xie LY, et al. A methylome-wide study of aging using massively parallel sequencing of the methyl-CpG-enriched genomic fraction from blood in over 700 subjects. *Hum Mol Genet.* 2014;23(5):1175–85.
- Reynolds LM, Taylor JR, Ding J, Lohman K, Johnson C, Siscovick D, et al. Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nat Commun.* 2014;5:5366.
- Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Res.* 2010;20(4):440–6.
- Yuan T, Jiao Y, de Jong S, Ophoff RA, Beck S, Teschendorff AE. An integrative multi-scale analysis of the dynamic DNA methylation landscape in aging. *PLoS Genet.* 2015;11(2):e1004996.
- Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* 2012;8(4):e1002629.
- Gervin K, Andreassen BK, Hjorthaug HS, Carlsen KCL, Carlsen KH, Undlien DE, et al. Intra-individual changes in DNA methylation not mediated by cell-type composition are correlated with aging during childhood. *Clin Epigenetics.* 2016;8:110.
- Gopalan S, Carja O, Fagny M, Patin E, Myrck JW, McEwen LM, et al. Trends in DNA methylation with age replicate across diverse human populations. *Genetics.* 2017;206(3):1659–74.
- Steeenga WT, Boekschoten MV, Lute C, Hooiveld GJ, de Groot PJ, Morris TJ, et al. Genome-wide age-related changes in DNA methylation and gene expression in human PBMCs. *Age (Dordr).* 2014;36(3):9648.
- Li S, Lund JB, Christensen K, Baumbach J, Mengel-From J, Kruse T, et al. Exploratory analysis of age and sex dependent DNA methylation patterns on the X-chromosome in whole blood samples. *Genome Med.* 2020;12(1):39.
- Lund JB, Li S, Christensen K, Mengel-From J, Soerensen M, Marioni RE, et al. Age-dependent DNA methylation patterns on the Y chromosome in elderly males. *Aging Cell.* 2020;19(2):e12907.
- López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell.* 2013;153(6):1194–217.
- 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.
- Iwamori T, Iwamori N, Matsumoto M, Ono E, Matzuk MM. Identification of KIAA1210 as a novel X-chromosome-linked protein that localizes to the acrosome and associates with the ectoplasmic specialization in testes. *Biol Reprod.* 2017;96(2):469–77.
- Akinloye O, Gromoll J, Callies C, Nieschlag E, Simoni M. Mutation analysis of the X-chromosome linked, testis-specific TAF7L gene in spermatogenic failure. *Andrologia.* 2007;39(5):190–5.
- CT database, <http://www.cta.incc.br/>, accessed 10.6.2021
- Rujescu D, Meisenzahl EM, Krejcova S, Giegling I, Zetzsche T, Reiser M, et al. Plexin B3 is genetically associated with verbal performance and white matter volume in human brain. *Mol Psychiatry.* 2007;12(2):190–4.
- Cannarile L, Delfino DV, Adorisio S, Riccardi C, Ayroldi E. Implicating the role of GILZ in glucocorticoid modulation of T-cell activation. *Front Immunol.* 2019;10:1823.

35. Nikolich-Zugich J. The twilight of immunity: emerging concepts in aging of the immune system. *Nat Immunol.* 2018;19(1):10–9.
36. Nguyen TA, Wu K, Pandey S, Lehr AW, Li Y, Bembem MA, et al. A cluster of autism-associated variants on X-Linked NLGN4X functionally resemble NLGN4Y. *Neuron.* 2020;106(5):759–768.e7.
37. Nguyen TA, Lehr AW, Roche KW. Neuroligins and neurodevelopmental disorders: X linked genetics. *Front Synaptic Neurosci.* 2020;12:33.
38. Bogaert AF, Skorska MN, Wang C, Gabrie J, MacNeil AJ, Hoffarth MR, et al. Male homosexuality and maternal immune responsiveness to the Y-linked protein NLGN4Y. *Proc Natl Acad Sci U S A.* 2018;115(2):302–6.
39. Gong Y, Wang L, Chippada-Venkata U, Dai X, Oh WK, Zhu J. Constructing Bayesian networks by integrating gene expression and copy number data identifies NLGN4Y as a novel regulator of prostate cancer progression. *Oncotarget.* 2016;7(42):68688–707.
40. Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell.* 2012;11(6):1132–4.
41. Gentilini D, Garagnani P, Pisoni S, Bacalini MG, Calzari L, Mari D, et al. Stochastic epigenetic mutations (DNA methylation) increase exponentially in human aging and correlate with X chromosome inactivation skewing in females. *Aging (Albany NY).* 2015;7(8):568–78.
42. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, et al. Distinct DNA methylomes of newborns and centenarians. *Proc Natl Acad Sci U S A.* 2012;109(26):10522–7.
43. Tserel L, Kolde R, Limbach M, Tretyakov K, Kasela S, Kisand K, et al. Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. *Sci Rep.* 2015;5:13107.
44. Zaghlood SB, Al-Shafai M, Al Muftah WA, Kumar P, Falchi M, Suhre K. Association of DNA methylation with age, gender, and smoking in an Arab population. *Clin Epigenetics.* 2015;7:6.
45. Florath I, Butterbach K, Müller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Hum Mol Genet.* 2014;23(5):1186–201.
46. Yusipov I, Bacalini MG, Kalyakulina A, Krivosov M, Pirazzini C, Gensous N, et al. Age-related DNA methylation changes are sex-specific: a comprehensive assessment. *Aging (Albany NY).* 2020;12(23):24057–80.
47. Benton MC, Sutherland HG, Macartney-Coxson D, Haupt LM, Lea RA, Griffiths LR. Methylome-wide association study of whole blood DNA in the Norfolk Island isolate identifies robust loci associated with age. *Aging (Albany NY).* 2017;9(3):753–68.
48. Xu Z, Taylor JA. Genome-wide age-related DNA methylation changes in blood and other tissues relate to histone modification, expression and cancer. *Carcinogenesis.* 2014;35(2):356–64.
49. Kim J, Kim K, Kim H, Yoon G, Lee K. Characterization of age signatures of DNA methylation in normal and cancer tissues from multiple studies. *BMC Genomics.* 2014;15(1):997.
50. Jansen RJ, Tong L, Argos M, Jasmine F, Rakibuz-Zaman M, Sarwar G, et al. The effect of age on DNA methylation in whole blood among Bangladeshi men and women. *BMC Genomics.* 2019;20(1):704.
51. Cotton AM, Price EM, Jones MJ, Balaton BP, Kobor MS, Brown CJ. Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Hum Mol Genet.* 2015;24(6):1528–39.
52. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, Friend S, Ideker T, Zhang K. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* 2013;49(2):359–67.
53. Talens RP, Christensen K, Putter H, Willemsen G, Christiansen L, Kremer D, et al. Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell.* 2012;11(4):694–703.
54. Forsberg LA. Loss of chromosome Y (LOY) in blood cells is associated with increased risk for disease and mortality in aging men. *Hum Genet.* 2017;136(5):657–63.
55. Machiela MJ, Chanock SJ. The ageing genome, clonal mosaicism and chronic disease. *Curr Opin Genet Dev.* 2017;42:8–13.
56. Gentilini D, Castaldi D, Mari D, Monti D, Franceschi C, Di Blasio AM, et al. Age-dependent skewing of X chromosome inactivation appears delayed in centenarians' offspring. Is there a role for allelic imbalance in healthy aging and longevity? *Aging Cell.* 2012;11(2):277–83.
57. Knudsen GP, Pedersen J, Klingenberg O, Lygren I, Ørstavik KH. Increased skewing of X chromosome inactivation with age in both blood and buccal cells. *Cytogenet Genome Res.* 2007;116(1–2):24–8.
58. Mossner M, Nolte F, Hütter G, Reins J, Klaumünzer M, Nowak V, et al. Skewed X-inactivation patterns in ageing healthy and myelodysplastic haematopoiesis determined by a pyrosequencing based transcriptional clonality assay. *J Med Genet.* 2013;50(2):108–17.
59. Tukiainen T, Villani AC, Yen A, Rivas MA, Marshall JL, Satija R, et al. Landscape of X chromosome inactivation across human tissues. *Nature.* 2017;550(7675):244–8.
60. Carrel L, Willard HF. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature.* 2005;434:400–4.
61. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol.* 2014;15:R31.
62. GEO NCBI database [www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)
63. Tan Q, Frost M, Heijmans BT, von Bornemann HJ, Christiansen K, et al. Epigenetic signature of birth weight discordance in adult twins. *BMC Genomics.* 2014;15:1062.
64. Voisin S, Almén MS, Zheleznyakova GY, Lundberg L, Zarei S, Castillo S, et al. Many obesity-associated SNPs strongly associate with DNA methylation changes at proximal promoters and enhancers. *Genome Med.* 2015;7:103.
65. Chuang YH, Paul KC, Bronstein JM, Bordelon Y, Horvath S, Ritz B. Parkinson's disease is associated with DNA methylation levels in human blood and saliva. *Genome Med.* 2017;9(1):76.
66. Zannas AS, Jia M, Hafner K, Baumert J, Wiechmann T, Pape JC, et al. Epigenetic upregulation of FKBP5 by aging and stress contributes to NF- $\kappa$ B-driven inflammation and cardiovascular risk. *Proc Natl Acad Sci U S A.* 2019;116(23):11370–9.
67. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8(2):203–9.
68. <http://bioinformatics.psb.ugent.be/webtools/Venn/>
69. Wolfe D, Dudek S, Ritchie MD, Pendergrass SA. Visualizing genomic information across chromosomes with PhenoGram. *BioData Min.* 2013;6(1):18.
70. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: A tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinform.* 2009;10:48.
71. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of dna sequences. *PLoS Comput Biol.* 2007;3(3):39.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.