Sex-chromosome-dependent aging in female heterogametic methylomes

Authors:

Marianthi Tangili [1] * †, Joanna Sudyka [1,2] †, Fabricio Furni[1], Per J. Palsbøll[1,3], Simon Verhulst* [1]

Affiliations:

[1] Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands

[2] Institute of Environmental Sciences, Jagiellonian University, Kraków, Poland

[3]Center for Coastal Studies, Provincetown, Massachusetts, USA

† These authors contributed equally to this work.

* Corresponding authors: Marianthi Tangili, Email: tangilim@gmail.com and Simon Verhulst, Email: s.verhulst@rug.nl

Abstract:

DNA methylation-based indices predict mortality and provide insights into aging mechanisms. Even though sex chromosomes are thought to play an important role in sex-specific aging, they have been largely ignored in epigenetic aging studies. We identified age-related CpG (AR-CpG) sites in the full methylomes of two avian species (zebra finch and jackdaw), including for the first time the avian sex chromosomes (Z and the female-specific, haploid W). In both species, we find a high concentration of AR-CpG sites on the W chromosome, with DNA methylation levels predominantly decreasing with age. Our results highlight distinct age-related changes in sex chromosome DNA methylation compared to the rest of the genome, suggesting that this feature of sex chromosomes may contribute to sex-dependent aging.

Main

The underlying causes of aging, the decline in biological functioning with time, remain an enigma. Phenotypic and molecular biomarkers that serve as proxies of biological age (*1*) are instrumental in unraveling aging mechanisms. In this context, indices based on age-dependent epigenetic modifications have shown an unrivaled capacity to predict chronological and biological age (*2*). These indices are mostly based on age-related changes in DNA methylation (DNAm) at CpG sites; sites in the genome where a 5' cytosine is followed by a guanine (*3*). Indices of chronological age based upon the degree of DNAm in certain CpG loci ('epigenetic clocks') have opened novel, pivotal windows into the complex landscape of age-associated phenotypes. Numerous epigenetic clocks have been developed in recent years for humans (*4–6*), model (*7*) and non-model animals (*8*, *9*), and even plants (*10*). Because DNAm contributes to the regulation of gene expression (*3*) and is known to change with age (*4*), studying DNAm in relation to aging across a broad range of species holds the potential to identify fundamental sites, sequences, and processes related to aging conserved across diverse taxa. Casting the net widely will enable identifying aging pathways shared among species, or classes, that likely are fundamental to aging.(*11*).

Sex differences in age-related pathology and lifespan are ubiquitous in the animal kingdom, with the homogametic sex in female heterogametic species living on average 7.1% longer (*12*), suggesting a link between sex chromosomes and aging. However, epigenetic aging studies routinely ignore sex chromosomes, with a few notable exceptions only in humans (*13*, *14*). Employing a novel, robust pipeline we identified age-related CpG (AR-CpG) sites where the degree of DNAm correlates with age among blood samples from captive zebra finches (*Taeniopygia guttata*) and wild jackdaws (*Corvus monedula*). We employed longitudinal data, i.e. multiple samples collected from an individual during its lifespan, to identify AR-CpG sites across the whole genome. We included, for the first time, the avian sex chromosomes (Z and the female-specific haploid W) and tested for deviations of the observed distribution of AR-CpG sites across chromosomes from null expectations. We hypothesized that an important role of sex chromosomes in aging will be revealed in AR-CpG sites being more numerous on sex chromosomes.

We provide details of the pipeline we developed to identify AR-CpG sites from longitudinal methylomes in the supplementary material (Text 1 and Fig. S1, S2). In both species, 0.006% of CpG sites were identified to show age-related changes in DNAm (zebra finch: 1,085/18,906,252 CpG sites; jackdaw: 1,299/21,835,349; Fig. S1). We did not detect AR-CpG sites in the mitochondrial genome, but AR-CpG sites were present on most nuclear chromosomes (Fig. 1). AR-CpG sites were more numerous than expected by chance on the W chromosome in both species (Fig.2; zebra finch: 17% of all AR-CpG sites, Table S2, $p<0.0001$; jackdaw: 14% of all AR-CpG sites, Table S3, $p<0.0001$, Fig.2). The AR-CpG site density on the Z chromosome was elevated in the zebra finch (13% of all age-related sites; $\chi^2= 2622$, df=36, $p<0.001$, Fig.2) and depressed in the jackdaw (3% of all AR-CpG sites; $\chi^2=2243$, df=35, $p<0.001$, Fig.2). The density of AR-CpG sites also varied among the autosomal chromosomes. In the zebra finch, AR-CpG sites were overrepresented on six chromosomes, and underrepresented on 22 chromosomes (Table S2). The corresponding numbers in the jackdaw were six and nine respectively (Table S3). In the zebra finch, DNAm level increased with age on 56% of AR-CpG sites (n= 604; Fig.

1C), which contrasts with the jackdaw, where DNAm increased with age on 31% of AR-CpG sites (n = 407; Fig. 1D).

AR-CpG sites were predominantly located in intergenic regions in both species, followed by introns, promoters, and exons in the zebra finch and introns, exons, and promoters in the jackdaw (Fig. 3A, D). The observed distribution deviated significantly from null expectations (zebra finch: χ^2= 1040.3, df=3, p<0.001, jackdaw: χ^2=321.49, df=3, p<0.001), also when considered separately for each annotation category (Table S4). AR-CpG sites were absent from exons in the W chromosome in both species (Fig. 3B, F, Fig. S3).

Functional consequences of DNAm mostly arise when a promoter is methylated, suppressing gene expression, making promoter DNAm of special interest. On the sex chromosomes promoters, DNAm increased with age on 12 sites and decreased on 21 sites (Fig, 3D). In the jackdaw, all AR-CpG sites in promoters increased with age (n=4; Fig. 3H). In the zebra finch autosomal chromosomes, 60 AR-CpG sites were detected in promoter regions, with DNAm increasing with age on 29 sites, and decreasing with age on 31 sites (Fig. 3D). In the jackdaw autosomal chromosomes, 46 AR-CpG sites were found in promoters among which 15 and 31 increased and decreased in DNAm with age, respectively (Fig3H). For a small number of AR-CpG sites in promoters the specific genes could be identified (Table S5). The GO enrichment analysis of the genes whose promoters contained AR-CpG sites were not involved in shared biological processes or molecular functions for either species. In both species, DNAm decreased with age on the majority of intergenic W chromosome AR-CpG sites (Fig.S3). Details of the direction of change in the degree of DNAm at AR-CpG sites per annotation category across all chromosomes in each species are depicted in Fig. S3.

The number of AR-CpG sites identified by our pipeline depends on the stringency of the selection criteria, raising the question whether this biases our findings. We therefore repeated the analysis with less stringent criteria, resulting in a six-fold increase in AR-CpG sites (Fig.S4,S5). However, the overall trends regarding the distribution of AR-CpG sites across the chromosomes remained the same in both species (Fig. S4, S5), leading us to assume that our results were robust with respect to pipeline criteria.

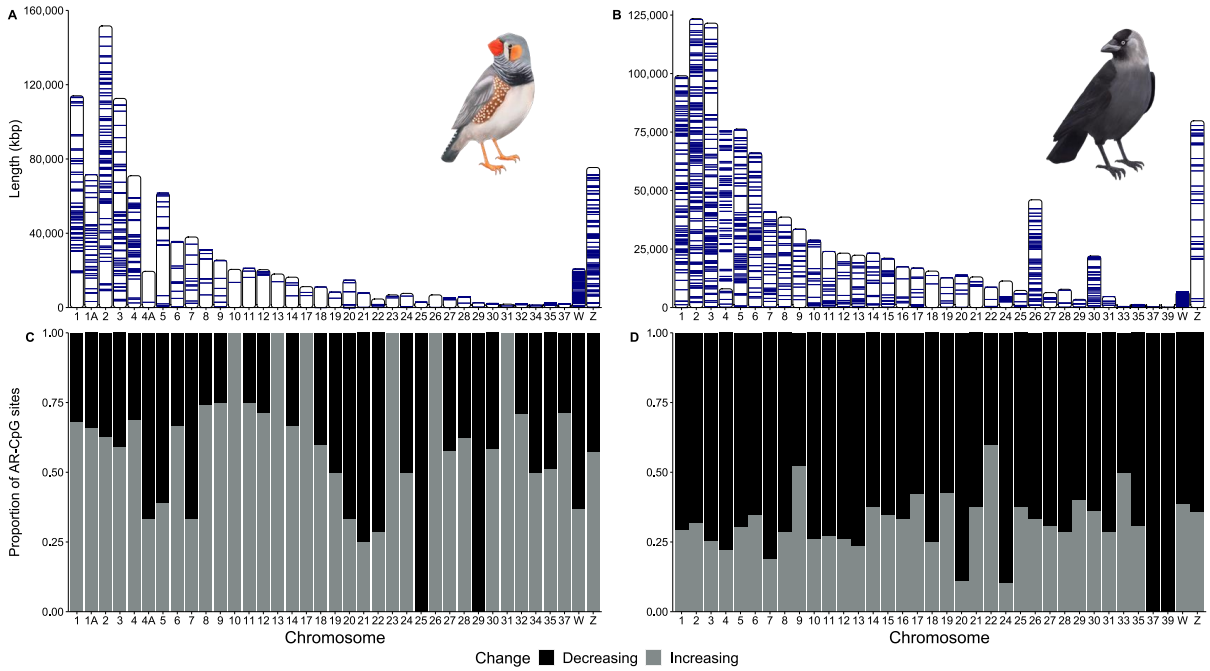


Figure 1. Location of AR-CpG sites across chromosomes and direction of change in DNAm. *Position of each AR-CpG site per chromosome for (**A**) the zebra finch and (**B**) the jackdaw. Proportion of AR-CpG sites on each chromosome whose level of DNAm increased (grey) or decreased (black) with age in the zebra finch (**C**) and the jackdaw (**D**). No AR-CpG sites were detected on zebra finch chromosomes 15, 16, 33, 36 and jackdaw chromosomes 23, 32, 34, 36, 38. The actual AR-CpG site positions for both species can be found in* (*15*).

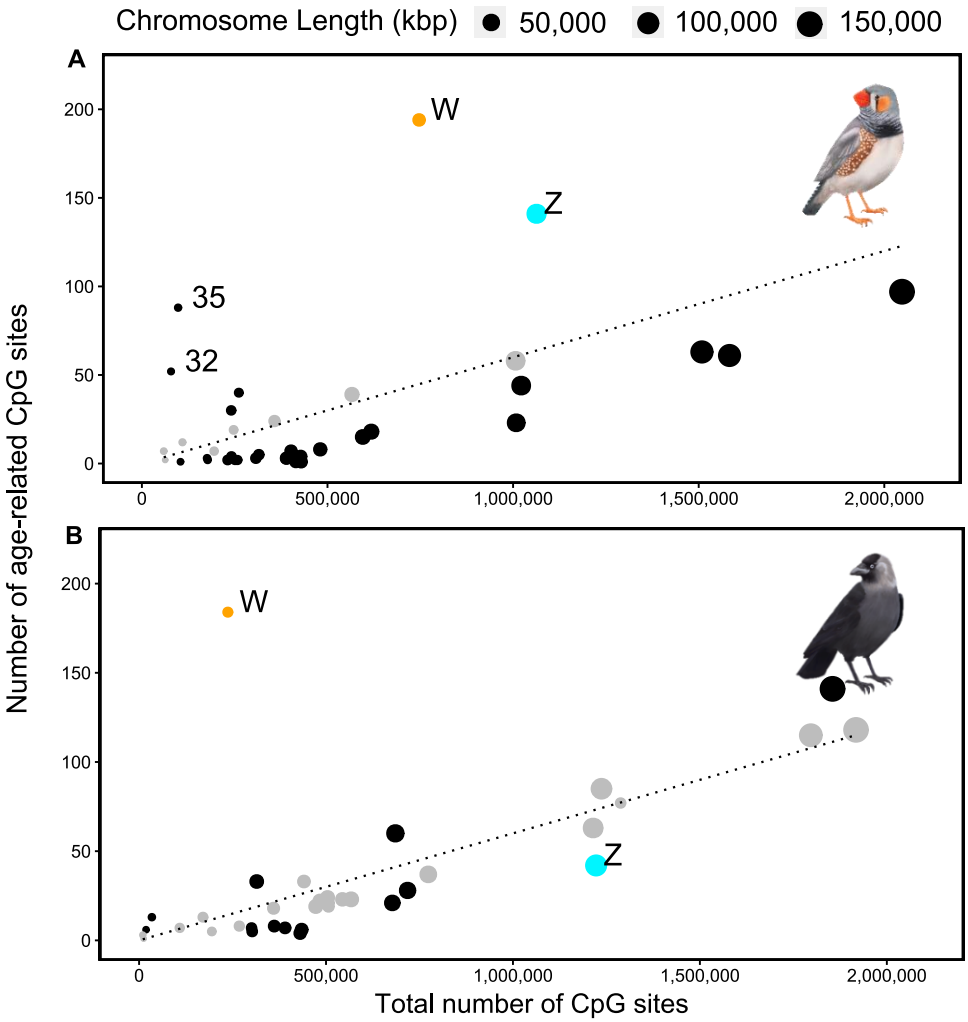


Figure 2. Number of AR-CpG sites in relation to total CpG sites per chromosome. *(**A**) Zebra finch and (**B**) jackdaw. The diameter of each data point is proportional to chromosome length in kilobase pairs (kbp). Black data points denote autosomal chromosomes with an AR-CpG density that deviated significantly from the null expectations (dotted line).*

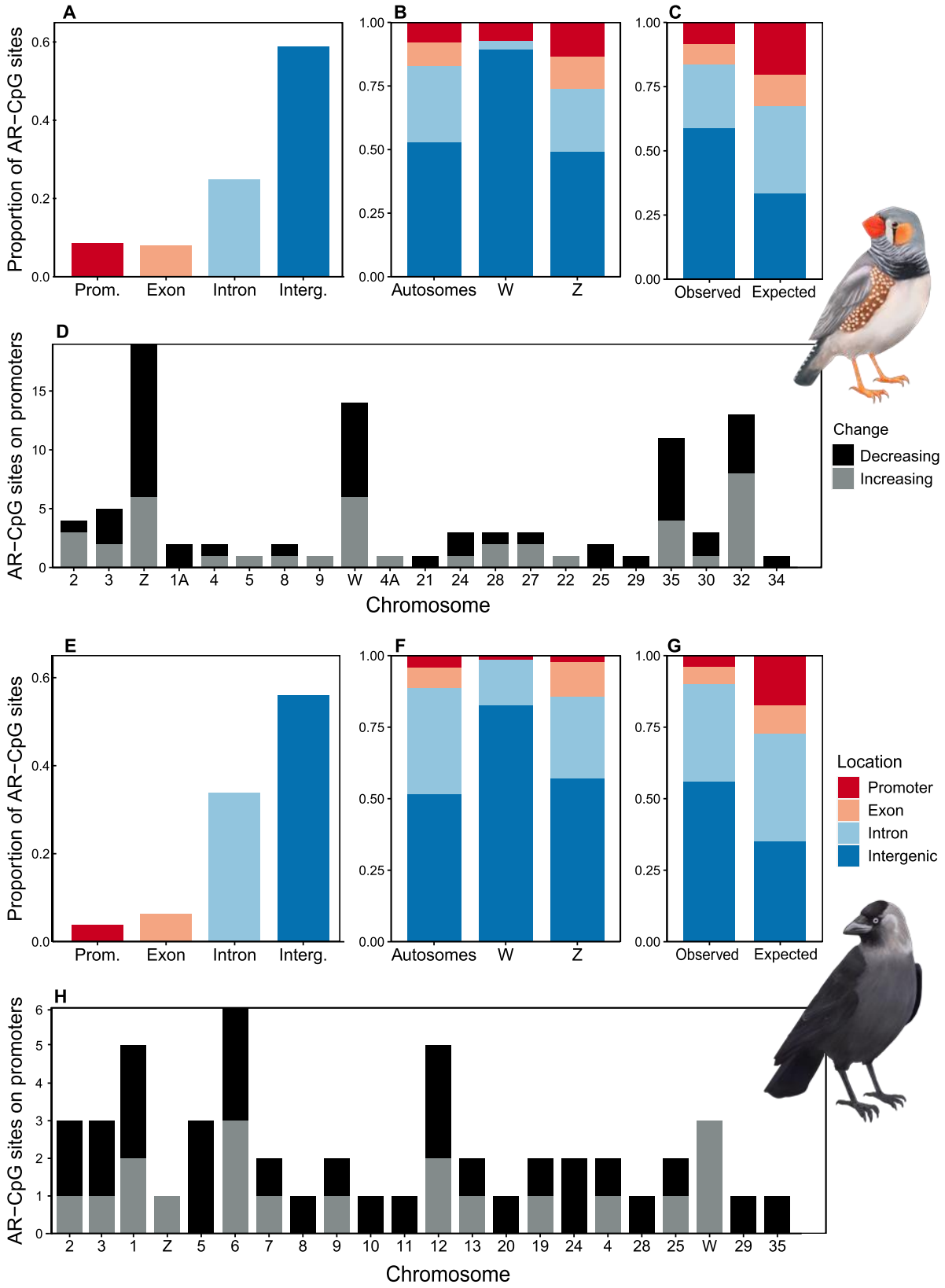

A
B
C
Proportion of AR−CpG sites
0.0
0.2
0.4
0.6
Prom.
Exon
Intron
Interg.
0.00
0.25
0.50
0.75
1.00
Autosomes
W
Z
Observed
Expected
D
AR−CpG sites on promoters
0
5
10
15
2
3
Z
1A
4
5
8
9
W
4A
21
24
28
27
22
25
29
35
30
32
34
Chromosome
Change
Decreasing
Increasing
E
F
G
Location
Promoter
Exon
Intron
Intergenic
H
0
1
2
3
4
5
6
2
3
1
Z
5
6
7
8
9
10
11
12
13
20
19
24
4
28
25
W
29
35
Chromosome

Figure 3. Functional annotation of AR-CpG sites. *Proportion of identified AR-CpG sites in each annotation category (***A & E***), by chromosome type (***B & F***) and observed vs. expected in the whole genome (***C & G***) for the zebra finch and jackdaw respectively. Count of AR-CpG sites increasing or decreasing in the degree of DNAm in promoters per chromosome (ordered by decreasing length) for the zebra finch and jackdaw, respectively (***D & H***).*

Our finding of a high density of AR-CpG sites on the W chromosome in two avian species highlights the potential of sex chromosome DNAm to drive sex-specific aging (*13*, *14*). A unique strength of our study was the longitudinal character of the data, combined with whole-genome methylome data, which provided us with a high-resolution view of the entire methylome. Consequently, we detected previously unrecognized sex-chromosome-specific patterns of epigenetic aging, unlike earlier attempts that were restricted to targeting specific genomic regions (*16*) or relied on cross-sectional analyses (*17*). The similarity of our findings between jackdaws and zebra finches, such as the comparable overall number of AR-CpG sites, their increase and decrease in DNAm over time as well as their distribution within each annotation category on the W chromosome suggest our findings may apply generally to passerine birds.

Whether DNAm increases (hypermethylation) or decreases (hypomethylation) with age typically varies among CpG sites (*18*). On the W chromosome in both species, DNAm decreased with age in the majority of AR-CpG sites. The avian W chromosome contains a high density of transposable elements (*19*), whose expression has negative fitness consequences (*20*), which may explain why DNAm is substantially higher on W compared to the autosomal chromosomes (*21*, also in our study species, data not shown). Decreasing DNAm with age likely leads to declining suppression of transposable element expression, and thereby increased genomic instability and deregulation of gene expression, both drivers of senescence (*23*, *24*). Across all chromosomes, but on W chromosomes in particular, AR-CpG sites were disproportionally located in intergenic regions. Besides transposons, intergenic regions also harbor microRNAs (*18*) that modulate longevity and the aging processes (*27*) while playing a significant role in cell proliferation, cell death, and hematopoiesis (*28*). Dysregulation of microRNA gene expression can lead to malignancy (*29*) and disrupted immune cell homeostasis (*30*). Decreased levels of DNAm with age in intergenic regions of the W chromosome suggests loss of the ability to inhibit detrimental genetic elements such as deleterious mutations leading to genomic instability, which is one of the hallmarks of aging (*24*). Our findings therefore indicate that epigenetic aging of the W chromosome via DNAm contributes to the observed shorter lifespans of female birds compared to males (*25*).

The density of AR-CpG sites was higher than expected by chance on the zebra finch Z chromosome, whereas there were slightly fewer AR-CpG sites than expected on the jackdaw Z chromosome. This difference matches the sex difference in lifespan in the two species: in zebra finches, females have a shorter lifespan (*31*), whereas there is no difference in lifespan between the sexes in jackdaws (*32*). The density of AR-CpG sites deviated significantly from the null expectations on most autosomes, but most deviations were small. The notable exceptions were chromosomes 32 and 35 in the zebra finch, where the density of AR-CpG sites was higher than expected. Whether variation in the AR-CpG density among chromosomes (other than the W chromosome) reflects the frequency of transposable elements or other functional components among chromosomes remains to be investigated.

Our finding that AR-CpG sites on the W chromosome mostly decreased in DNAm with age and that most such sites were located in intergenic regions contrasts research on the human haploid, heterogametic Y chromosome where the majority of AR-CpG sites were found to become hypermethylated with age and were located in genes (*13*). On the other hand, a single early study found the mammalian X chromosome to harbor relatively few AR-CpG sites (*33*), as we found for the jackdaw Z chromosome. The avian sex chromosomes ZW are not homologous to the mammalian XY (*34*) and it is therefore unsurprising that DNAm patterns differ from male heterogametic (XY) species. Our findings suggest that DNAm on sex chromosomes likely plays different roles in male and female heterogametic systems. Furthermore, our results underscore the notion that some of the key mechanisms driving sex-specific differences in aging might be fundamental and thus conserved among closely related species, but not across taxa. Consequently, efforts should be directed towards deciphering the role of sex chromosomes in aging across a broader range of species as well as understanding the mechanisms underlying this process in species with various sex determination systems.

Ethics approval

All methods and experiments detailed in this manuscript were performed under the approval of the Central Committee for Animal Experiments (Centrale Commissie Dierproeven) of the Netherlands, under licenses AVD1050020174344 and AVD1050020184967.

Acknowledgments

We thank the Center for Information Technology at the University of Groningen for support and access to the Peregrine and Hábrók high-performance computing clusters. We express our gratitude to Ellis Mulder for her help in the laboratory, Berber Maarsingh for illustrations, and Raphaël Scherrer for assistance with the statistical analyses. We are also indebted to the animal caretakers at the University of Groningen as well as numerous students whose invaluable help made this project possible.

Funding: MT was supported by an Adaptive Life Scholarship awarded by the University of Groningen. JS received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 101025890. Contributions by FB and PJP were supported by the European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no. 813383, and the University of Groningen.

Author contributions: MT, JS, and SV conceived and designed the study. MT and JS performed the bioinformatics and data analysis under the supervision of SV and PJP. FF performed the genomic annotation. MT and SV wrote the manuscript with input from all authors. All authors approved the final version of the manuscript. **Competing interests:** The authors declare no competing interests**.**

Data and materials availability

Raw reads for whole methylomes of both species are available at the National Center for Biotechnology Information under BioProject ID PRJNA1108628. The bioinformatics code and R code used in this study can be found at https://datadryad.org/stash/share/6L2TYwRWmgzi4c35v1pf987XyECT1wIdJBoqcwwEXeA.

Supplementary Materials

Materials and Methods

Sample origin

Zebra finch blood samples were collected during a long-term experiment where birds were housed in outdoor aviaries (320 × 150 × 210 cm) each containing single sex groups of 18–24 adults (*31*). Twenty longitudinal samples were collected from ten known-age individuals (six males and four females) sampled twice during their lifetime at an average interval of 1,470 days between the two sampling points. Jackdaw blood samples were collected from individuals of a free-ranging population breeding in nest-boxes south of Groningen, the Netherlands (53.1708°N, 6.6064°E). We analyzed 22 longitudinal blood samples collected during 2007 – 2021 from 11 known age adults (five males and six females) with an average sampling interval of 2,429 days. For both species, the precise age of sampled individuals was known because individuals were followed since birth (Table S1). Samples were taken between June 2008 and December 2014 and stored in -80°C in EDTA buffer.

Whole genome bisulfite sequencing

We extracted total-cell DNA using innuPREP DNA Mini Kit (Analytik Jena GmbH+Co) from 3µl (nucleated) red blood cells according to the manufacturer's protocol. Whole genome sequencing was performed by The Hospital for Sick Children (Toronto, Canada) where paired-end Illumina next-generation sequencing (150bp) was carried out on either an Illumina HiSeqX™ (12 zebra finch samples) or an Illumina NovaSeq™ sequencer (eight zebra finch samples and 22 jackdaw samples). Libraries were prepared using the Swift Biosciences Inc. Accel NGS Methyl Seq kit (part no. 30024 and 30096) and the DNA was bisulfite converted using the EZ-96 DNA Methylation-Gold kit from (Zymo Research Inc., part no. D5005) as per the manufacturer's protocol and subsequently subjected to whole-genome amplification.

Bioinformatic processing

Sequences were trimmed using Trim Galore! v. 0.6.10 (*35*) in paired-end mode. Visual quality controls of the data were carried out before and after trimming using FastQC v. 011.9 (*36*) and MultiQC v. 11.14 (*37*). Because the Swift Biosciences Inc. Accel NGS Methyl Seq kit was used for library preparation, the first ~10 bp showed extreme biases in sequence composition and M-bias, so after checking the M-bias plots, the first 10 bps were further trimmed from each sequence.

Alignments were performed using Bismark v. 0.14.433 (*38*) using the Bowtie 2 v. 2.4.5 alignment algorithm (*39*) for both in silico bisulfite conversion of the reference genomes and alignments. For zebra finch, trimmed reads were aligned against the in silico bisulfite converted zebra finch (*Taeniopygia guttata*) reference genome (GCA_003957565.4, *46*) and the average mapping efficiency was 64.5% (SD: 3.22). Because the jackdaw reference genome did not contain an assembled W chromosome, the jackdaw bisulfite sequencing data was aligned to an *in silico* bisulfite converted Hawaiian crow genome (*Corvus hawaiiensis*, GCA_020740725.1, *46*, the closest relative species with an assembled W chromosome) the average mapping efficiency was 64.7% (SD: 0.98).

Functional annotation of AR-CpG sites

The GFT-formatted annotation files for the zebra finch (GCF_003957565.2) and the Hawaiian crow (GCF_020740725.1) were converted into BED12 format using the University of California, Santa Cruz (UCSC) utilities *gtfToGenePred* and *genePredToBed* (available at https://hgdownload.soe.ucsc.edu/downloads.html#utilities_downloads). AR-CpG sites were annotated using the tool *annotateWithGeneParts* from the R 4.1.2 package *genomation* 1.4.1 (*41*). This tool hierarchically classifies the sites into pre-defined functional regions, i.e., promoter, exon, intron, or intergenic, hereon referred to as annotation categories. The predefined functional regions were based on the annotation information present in the BED12 files accessed with the *genomation* tool *readTranscriptFeatures*. Annotations were performed as a hierarchical assignment (promoter > exon). Subsequently, a customized R script was employed to integrate the annotation results of AR-CpG sites with their respective annotation category information.

Gene ontology enrichment analysis

The gene ontology (GO) enrichment analysis was conducted using the PANTHER knowledgebase (version 18.0, *48*) and performed a Fisher's exact test and false discovery rate calculation with the *Gallus gallus* genome as it was the most closely related species to our study species available in the database.

Statistical assessment

We conducted χ^2 tests using the R package *janitor v*.2.2.0 (*43*) to compare the observed and expected number of AR-CpG sites in each species and applied exact binomial tests using the R package *stats* (*44*) to compare the observed and expected number of AR-CpG for each chromosome. The expected number of AR-CpG sites for each chromosome was inferred as the product of the observed number of CpG sites on each chromosome and 0.006; the average frequency of AR-CpG sites for the whole genome in both species. We applied a Benjamini-Hochberg correction to control for the false discovery rate to account for multiple hypotheses testing (*45*).

We assessed the observed and expected number of AR-CpG sites per annotation category using a χ^2 test. The proportion of expected AR-CpG sites per annotation category was estimated as the proportion of all CpG sites captured by our analysis located in each annotation category. For all χ^2 tests, we used an exact binomial test as a *post hoc* test to assess observed vs. expected occurrences. All statistical tests were conducted using the CRAN R (4.1.2).

Table S1. Chronological age at collection and sex of all the individuals in our dataset.

Species	ID	Age young (days)	Age old (days)	Sex
Zebra finches	2269	295	2636	male
	2288	289	2644	female
	2257	298	884	male
	2277	292	631	male
	2247	468	2651	female
	1923	853	2124	male
	2172	732	2002	female
	1942	847	2484	male
	2354	351	2367	male
	2423	213	912	female
Jackdaws	480	749	5496	female
	775	743	3666	male
	888	748	2584	female
	923	753	2948	female
	973	742	3670	male
	977	1119	3666	male
	1146	1458	4022	female
	1350	1099	2939	male
	1502	745	2935	female
	1605	747	2228	female
	1762	745	2210	male

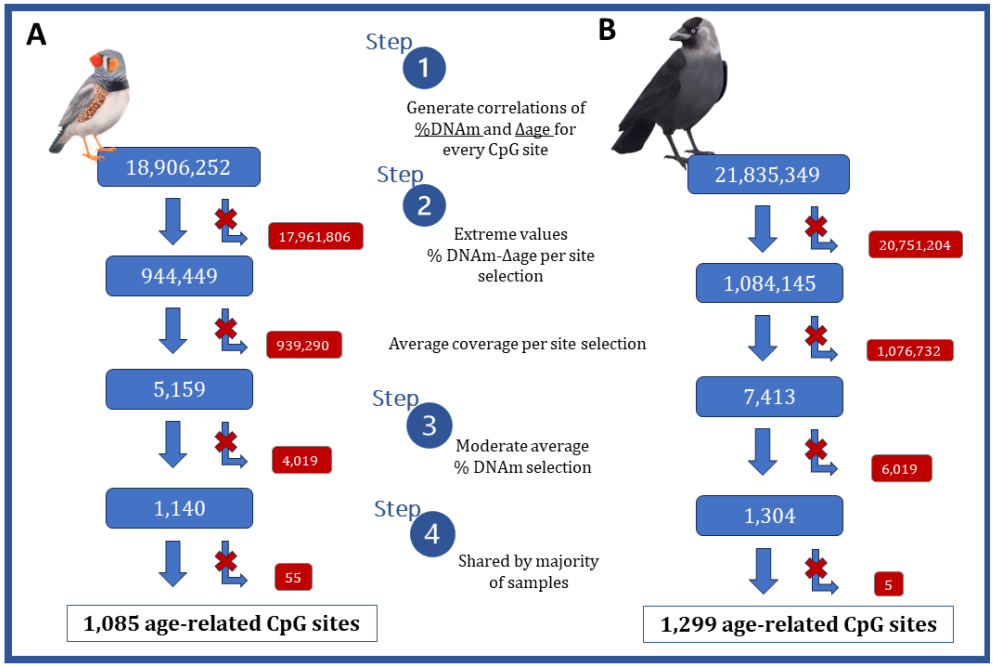


*Figure S1. Number of CpG sites remaining (blue panels) and lost (red panels) after each step in the pipeline for the zebra finch (**A**) and jackdaw (**B**).*

Supplementary Text 1

Pipeline for identification of AR-CpG sites (Fig.S1)

Step 1. For each sequenced CpG site we generated a correlation coefficient between DNAm at the focal sampling point and delta age (the difference between an individual's chronological age at the time of sampling and its average age calculated for both samplings, van de Pol & Wright, 2009) to account for between and within subject effects.

Step 2. We then correlated the values from Step 1 with the average sequencing coverage per CpG site over all samples. To include the sites with the highest correlations between DNAm at the focal sampling point and delta age (negative or positive as we want to include methylations increases and decreases with age across all ranges of coverage) we selected the top 5% sites for every 20 units of coverage. We assumed a threshold of minimum coverage to ensure high quality of reads for the selected sites. This step depends on the quality of sequencing data and can be tailored to the coverage distribution in a particular dataset. It can be manipulated to produce the final outcome (after step 4) of 1000 candidate sites (±20%) per species. Although there is no set minimum coverage for identification of age-related sites, sites with higher coverage hold the power to produce more accurate data. Therefore, the threshold coverage values in our datasets were >=20 for zebra finches and >=25 for jackdaws.

Step 3. We then selected CpG sites with average DNAm across all samples between 40% and 60% that are likely to have a higher absolute change in DNAm over time.

Step 4. Finally, we selected only the sites that were shared by the majority of the samples to increase the site detection probability across the whole population. In our datasets that was >=15 samples for autosomes (75% of the samples in the zebra finch, 68% in the jackdaw) and >=6 samples for the W chromosome for the zebra finch (75% of female samples) and >=8 for the jackdaw (67% of female samples), as it is present only in females in both species.

Since mtDNA was not found to contain any AR-CpG sites when included in the initial analyses, it was excluded from subsequent analyses.

This newly-developed pipeline to identify AR-CpG sites across the genome can be applied to (i) an array of different species, (ii) on DNAm data of varying qualities (in terms of sequencing coverage) and (iii) for which a limited number of individuals sampled longitudinally is sufficient. The robustness of our pipeline was confirmed by the high correlation between the number of AR-CpG sites generated with different cutoff parameters. This robustness stems from numerous factors. Firstly, we analyzed whole genome DNAm data, meaning that we had the opportunity to investigate the change in DNAm of CpG sites throughout the entirety of all of the chromosomes, including the sex chromosomes. Moreover, as previously stated, the longitudinal nature of our data makes it ideal to identify within-individual changes in DNAm with age. Another strength of our method is that we take extra steps to ensure that we not only identify CpG sites whose DNAm changes with age, but CpG sites that have extreme changes (in the same direction) in DNAm that are shared by the majority of the samples in the population. Additionally, we use the metric of DNAm correlated with Δage (the difference between an individual's chronological age at the time of sampling and its average age calculated for both samplings) instead of chronological age, meaning that individuals with larger timespans between sampling points have a higher influence on the identification of AR-CpG sites. A larger time interval between samplings exerts more pronounced changes in DNAm patterns over time, thereby enhancing the

detectability of AR-CpG sites. Lastly, we observe comparable yet not identical patterns of AR-CpG site location in the genomes of our two study species which is to be expected as they belong to the same class but are not very closely related phylogenetically. Because of all the above, we maintain confidence in the strength and robustness of our results.

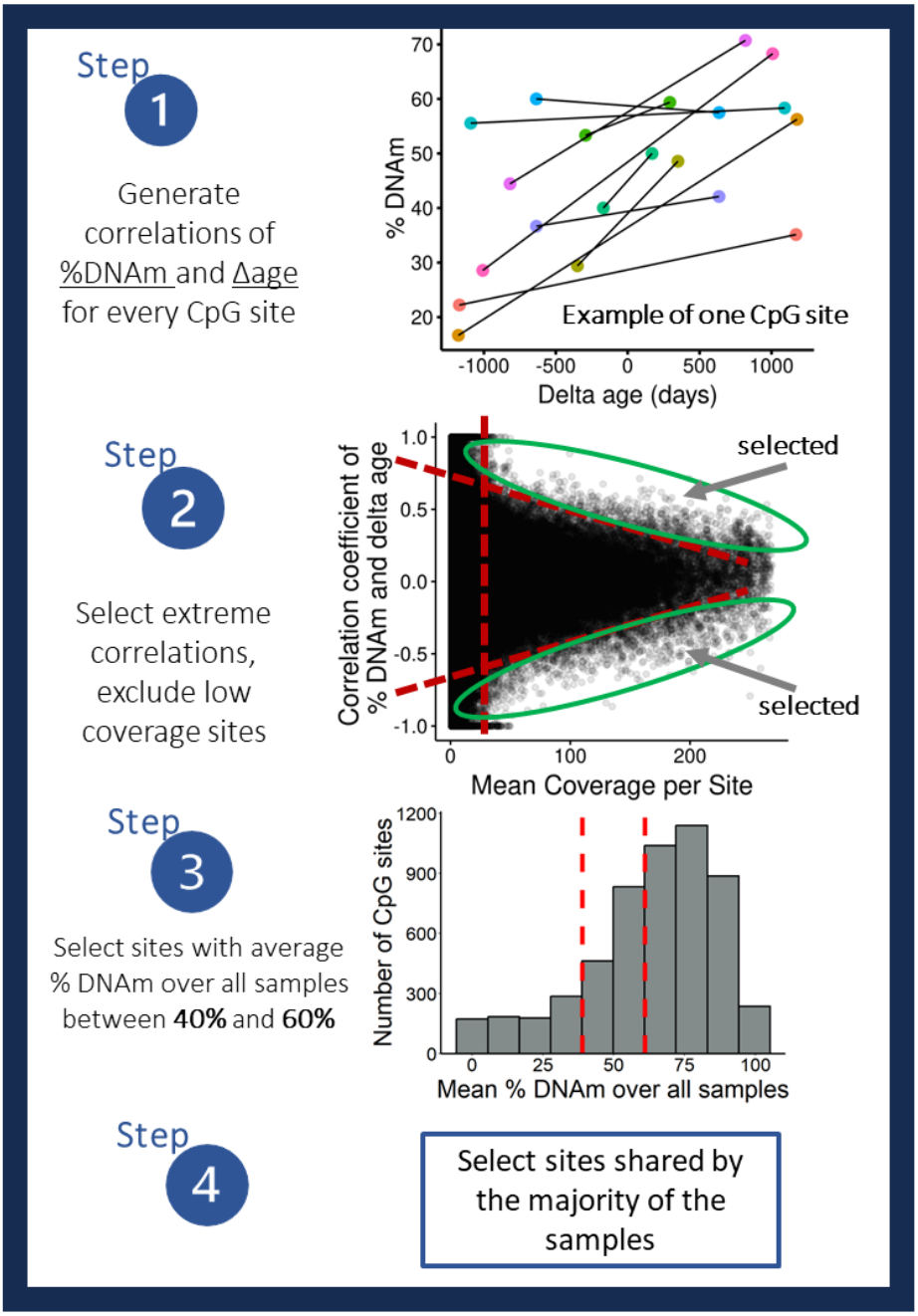
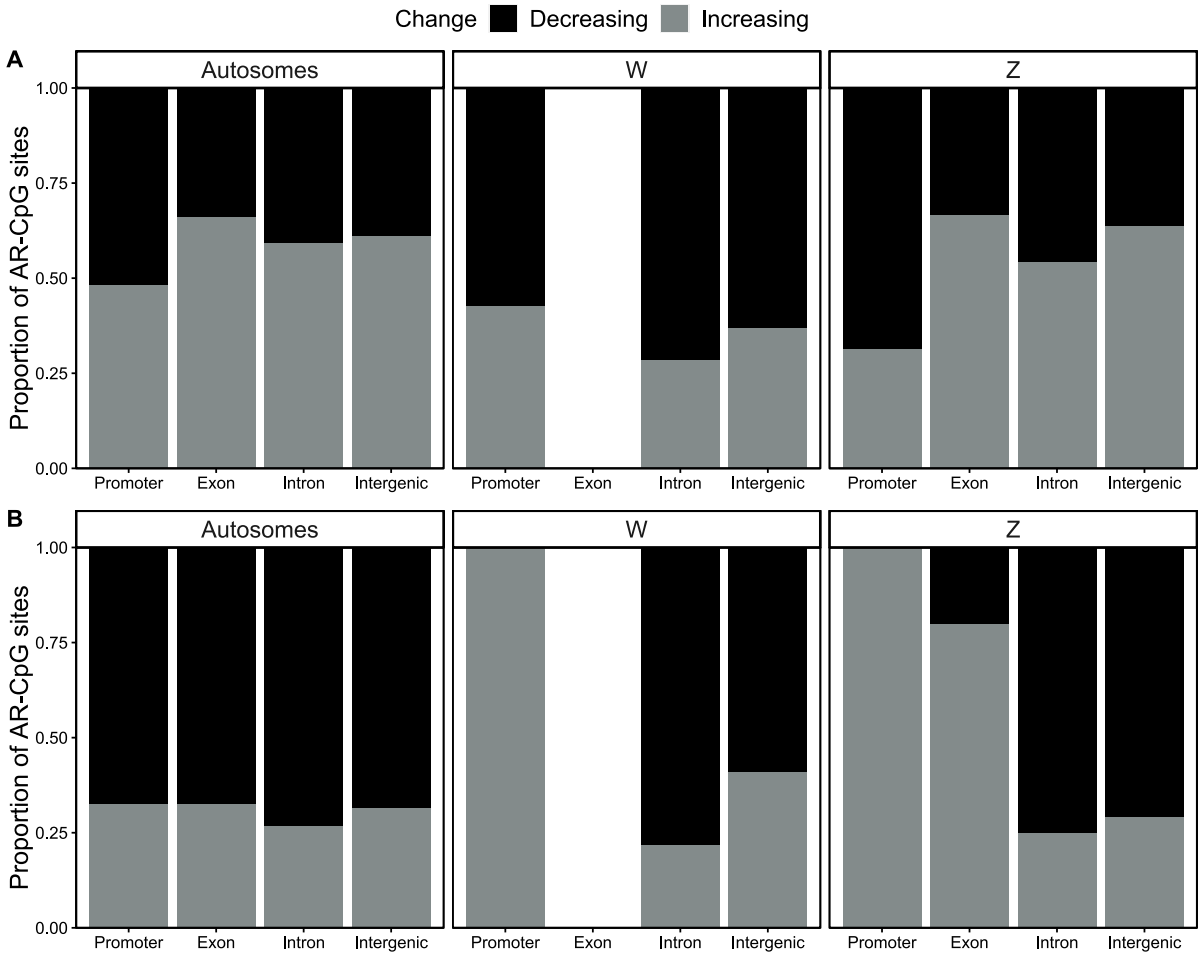


Figure S2. Graphical representation of the pipeline for identification of AR-CpG sites from longitudinal, whole genome bisulfite sequencing (WGBS) data.



415 *Figure S3. Proportion of AR-CpG sites increasing and decreasing in DNAm with age in each annotation category in the zebra finch (**A**) and the jackdaw (**B**).*

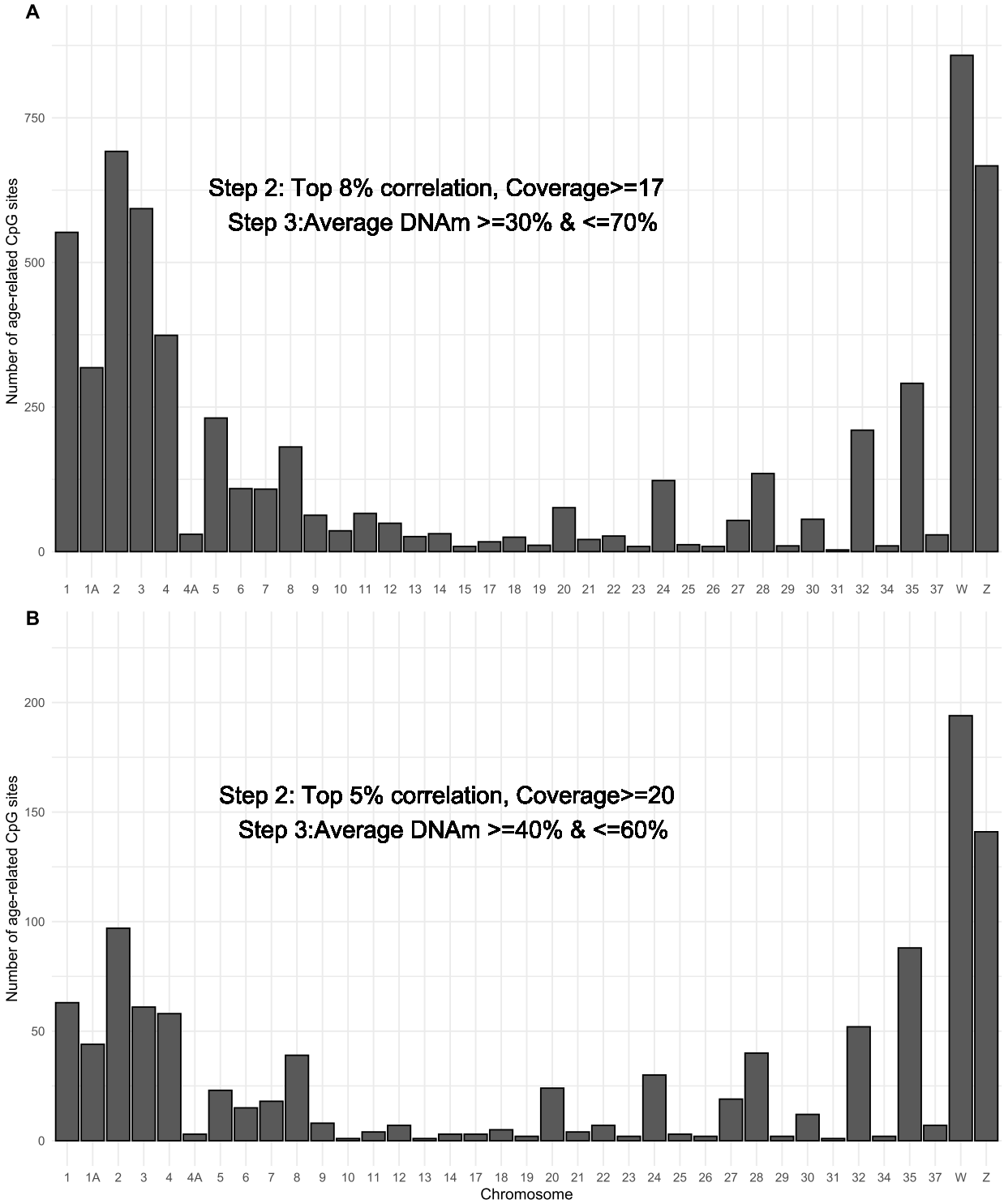


Figure S4. Comparison of "lenient" (A) and "standard" (B) pipelines for the zebra finch. The lenient pipeline selects the top 8% of the correlations between % DNAm and Δage per 20 coverage, selects positions with average coverage of >=17 (Step 2) and sites with average % DNAm between 30% and 70% (Step 3). As a result the total number of sites finally selected in Step 4 increases to 6556 in the zebra finch. Note that the numbers of sites per chromosomes are highly correlated between the two pipelines ($r_{(35)}$ =0.93, p<0.001) pointing to the high

repeatability and robustness of the pipeline in response to threshold modifications making it applicable to the whole ranges of methylation data.

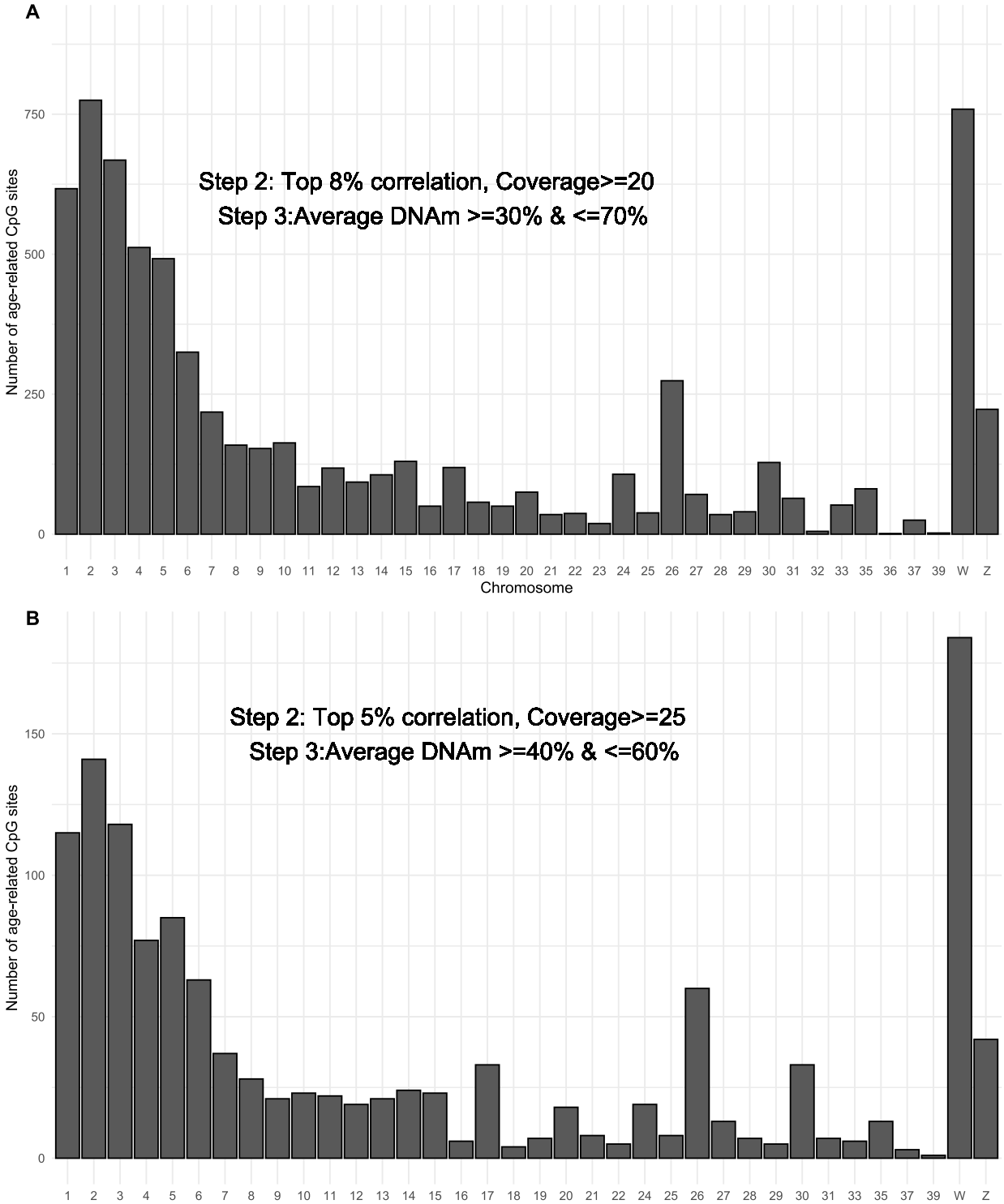


Figure S5. Comparison of "lenient" (A) and "standard" (B) pipelines for the jackdaw. The lenient pipeline selects the top 8% of the correlations between % DNAm and Δage per 20 coverage, selects positions with average coverage of >=25 (Step 2) and sites with average % DNAm between 30% and 70% (Step 3). As a result the total number of sites finally selected in Step 4 increases to 6961 in the jackdaw. Note that the numbers of sites per chromosomes are highly correlated between the two pipelines ($r_{(34)}$ =0.98, p<0.001) pointing to the high

repeatability and robustness of the pipeline in response to threshold modifications making it applicable to the whole ranges of methylation data.

Table S2. Proportions of observed and expected distribution of age-related CpG sites in each chromosome and results of two-sided exact binomial tests of observed vs. expected proportions of AR-CpG sites in each chromosome in the zebra finch. Significant differences (p<0.05) marked in bold.

Chromosome	AR-CpG sites	Proportion observed AR-CpG sites	Proportion expected AR-CpG sites	P-adj.
1	**63**	**0.058**	**0.082**	**0.006**
1A	**44**	**0.041**	**0.055**	**0.044**
2	**97**	**0.089**	**0.111**	**0.032**
3	**61**	**0.056**	**0.086**	**0.001**
4	58	0.053	0.055	0.947
4A	**3**	**0.003**	**0.021**	**<0.0001**
5	**23**	**0.021**	**0.055**	**<0.0001**
6	**15**	**0.014**	**0.032**	**<0.0001**
7	**18**	**0.017**	**0.033**	**0.002**
8	39	0.036	0.031	0.317
9	**8**	**0.007**	**0.026**	**<0.0001**
10	**1**	**0.001**	**0.023**	**<0.0001**
11	**4**	**0.004**	**0.023**	**<0.0001**
12	**7**	**0.006**	**0.022**	**<0.0001**
13	**1**	**0.001**	**0.022**	**<0.0001**
14	**3**	**0.003**	**0.021**	**<0.0001**
17	**3**	**0.003**	**0.017**	**<0.0001**
18	**5**	**0.005**	**0.017**	**0.001**
19	**2**	**0.002**	**0.013**	**0.001**
20	24	0.022	0.019	0.536
21	**4**	**0.004**	**0.013**	**0.005**
22	7	0.006	0.011	0.261
23	**2**	**0.002**	**0.014**	**<0.0001**
24	**30**	**0.028**	**0.013**	**<0.0001**
25	**3**	**0.003**	**0.010**	**0.025**
26	**2**	**0.002**	**0.014**	**0.000**
27	19	0.018	0.013	0.261
28	**40**	**0.037**	**0.014**	**<0.0001**
29	**2**	**0.002**	**0.010**	**0.006**

30	12	0.011	0.006	0.053
31	**1**	**0.001**	**0.006**	**0.048**
32	**52**	**0.048**	**0.004**	**<0.0001**
34	2	0.002	0.003	0.614
35	**88**	**0.081**	**0.005**	**<0.0001**
37	7	0.006	0.003	0.111
W	**194**	**0.179**	**0.040**	**<0.0001**
Z	**141**	**0.130**	**0.058**	**<0.0001**

Table S3. Proportions of observed and expected distribution of age-related CpG sites in each chromosome and results of two-sided exact binomial tests of observed vs. expected proportions of AR-CpG sites in each chromosome in the jackdaw. Significant differences (p<0.05) marked in bold.

Chromosome	**AR-CpG sites**	**Proportion observed AR-CpG sites**	**Proportion expected AR-CpG sites**	**P-adj.**
1	115	0.089	0.084	0.599
2	**141**	**0.109**	**0.087**	**0.017**
3	118	0.091	0.090	0.909
4	77	0.059	0.060	0.953
5	85	0.065	0.058	0.309
6	63	0.048	0.057	0.288
7	37	0.028	0.036	0.236
8	**28**	**0.022**	**0.034**	**0.033**
9	**21**	**0.016**	**0.032**	**0.003**
10	23	0.018	0.027	0.084
11	22	0.017	0.023	0.275
12	19	0.015	0.022	0.124
13	21	0.016	0.024	0.128
14	24	0.018	0.024	0.337
15	23	0.018	0.025	0.127
16	**6**	**0.005**	**0.020**	**<0.0001**
17	33	0.025	0.021	0.309
18	**4**	**0.003**	**0.020**	**<0.0001**
19	**7**	**0.005**	**0.018**	**0.001**
20	18	0.014	0.017	0.588
21	**8**	**0.006**	**0.017**	**0.004**
22	**5**	**0.004**	**0.014**	**0.002**
24	19	0.015	0.024	0.060
25	8	0.006	0.013	0.067
26	**60**	**0.046**	**0.032**	**0.016**
27	13	0.010	0.008	0.519
28	**7**	**0.005**	**0.014**	**0.013**

29	5	0.004	0.009	0.077
30	**33**	**0.025**	**0.015**	**0.012**
31	7	0.005	0.005	0.893
33	**6**	**0.005**	**0.001**	**0.004**
35	**13**	**0.010**	**0.002**	**<0.0001**
37	3	0.002	0.000	0.051
39	1	0.001	0.001	0.588
W	**184**	**0.142**	**0.011**	**<0.0001**
Z	**42**	**0.032**	**0.057**	**<0.0001**

*Table S4. Proportions of observed and expected distribution of age-related CpG sites in each annotation category and results of two-sided exact binomial tests of observed vs. expected proportions of AR-CpG sites in each annotation category for the zebra finch (**A**) and the jackdaw (**B**). Expected proportions were calculated from the location of all CpG sites captured by our analysis in each annotation category.*

	A			***B***		
Category	*Observed proportions*	*Expected proportions*	*p-value (two-sided)*	*Observed proportions*	*Expected proportions*	*p-value (two-sided)*
Promoters	0.086	0.124	**<0.001**	0.039	0.174	**<0.001**
Exons	0.080	0.335	**<0.001**	0.063	0.101	**<0.01**
Introns	0.248	0.339	**<0.001**	0.339	0.376	**0.04**
Intergenic regions	0.587	0.203	**<0.001**	0.560	0.351	**<0.001**

*Table S5. Identified genes in whose promoters we located AR-CpG sites and whether DNAm increased or decreased with time at that AR-CpG site in the zebra finch (**A**) and jackdaw (**B**). The 88 and 43 remaining AR-CpG sites in promoters for the zebra finch and jackdaw respectively were located in genes with unknown function.*

	Gene symbol	***Gene name***	***Accession number***	***Increasing/ decreasing in DNAm with age***	***Chromosome***
A	KRT18	keratin 18	XM_030259242.3	Decreasing	29
	MGAT2	alpha-1,6-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase	XM_030275310.3	Increasing	Z
	MIR2973	microRNA precursor	NR_048982.1	Decreasing	Z
	NTMT2	N-terminal Xaa-Pro-Lys N-methyltransferase 2	XM_041717884.1	Increasing	8
	RNF223	ring finger protein 223	XM_041720425.1	Decreasing	21
B	CLASP2	cytoplasmic linker associated protein 2	XM_048298364.1	Increasing	1
	FGF7	fibroblast growth factor 7	XM_048318267.1	Decreasing	13
	MAB21L3	mab-21 like 3	XM_048293924.1	Decreasing	2
	SLC16A7	solute carrier family 16 member 7	XM_048300392.1	Increasing	4
	TEKT1	tektin 1	XM_048324488.1	Decreasing	20
	TSPAN13	tetraspanin 13	XM_048304222.1	Decreasing	1
	TTC39A	tetratricopeptide repeat domain 39A	XM_048313046.1	Increasing	9

References

35. F. Krueger, F. James, P. Ewels, E. Afyounian, M. Weinstein, B. Schuster-Boeckler, G. Hulselmans, sclamons, FelixKrueger/TrimGalore: v0.6.10. Zenodo (2023). https://doi.org/10.5281/zenodo.7598955.

36. S. Andrews, FastQC: a quality control tool for high throughput sequence data. (2010). https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

37. P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

38. F. Krueger, S. R. Andrews, Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).

39. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

40. A. Rhie, S. A. McCarthy, O. Fedrigo, J. Damas, G. Formenti, S. Koren, M. Uliano-Silva, W. Chow, A. Fungtammasan, J. Kim, C. Lee, B. J. Ko, M. Chaisson, G. L. Gedman, L. J. Cantin, F. Thibaud-Nissen, L. Haggerty, I. Bista, M. Smith, B. Haase, J. Mountcastle, S. Winkler, S. Paez, J. Howard, S. C. Vernes, T. M. Lama, F. Grutzner, W. C. Warren, C. N. Balakrishnan, D. Burt, J. M. George, M. T. Biegler, D. Iorns, A. Digby, D. Eason, B. Robertson, T. Edwards, M. Wilkinson, G. Turner, A. Meyer, A. F. Kautt, P. Franchini, H. W. Detrich, H. Svardal, M. Wagner, G. J. P. Naylor, M. Pippel, M. Malinsky, M. Mooney, M. Simbirsky, B. T. Hannigan, T. Pesout, M. Houck, A. Misuraca, S. B. Kingan, R. Hall, Z. Kronenberg, I. Sović, C. Dunn, Z. Ning, A. Hastie, J. Lee, S. Selvaraj, R. E. Green, N. H. Putnam, I. Gut, J. Ghurye, E. Garrison, Y. Sims, J. Collins, S. Pelan, J. Torrance, A. Tracey, J. Wood, R. E. Dagnew, D. Guan, S. E. London, D. F. Clayton, C. V Mello, S. R. Friedrich, P. V Lovell, E. Osipova, F. O. Al-Ajli, S. Secomandi, H. Kim, C. Theofanopoulou, M. Hiller, Y. Zhou, R. S. Harris, K. D. Makova, P. Medvedev, J. Hoffman, P. Masterson, K. Clark, F. Martin, K. Howe, P. Flicek, B. P. Walenz, W. Kwak, H. Clawson, M. Diekhans, L. Nassar, B. Paten, R. H. S. Kraus, A. J. Crawford, M. T. P. Gilbert, G. Zhang, B. Venkatesh, R. W. Murphy, K.-P. Koepfli, B. Shapiro, W. E. Johnson, F. Di Palma, T. Marques-Bonet, E. C. Teeling, T. Warnow, J. M. Graves, O. A. Ryder, D. Haussler, S. J. O'Brien, J. Korlach, H. A. Lewin, K. Howe, E. W. Myers, R. Durbin, A. M. Phillippy, E. D. Jarvis, Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746 (2021).

41. A. Akalin, V. Franke, K. Vlahovic, C. E. Mason, D. Schu, Genome analysis genomation : a toolkit to summarize , annotate and visualize genomic intervals. *Bioinformatics* **31**, 1127–1129 (2015).

42. P. D. Thomas, D. Ebert, A. Muruganujan, T. Mushayahama, L.-P. Albou, H. Mi, PANTHER : Making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2021).

43. S. Firke, janitor: Simple Tools for Examining and Cleaning Dirty Data. R package version 2.2.0. (2023). https://cran.r-project.org/package=janitor.

44. R Core Team, R: A language and environment for statistical computing. RStudio 2021.09.0+351 "Ghost Orchid" Release (077589bcad3467ae79f318afe8641a1899a51606,

2021-09-20) for Windows Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) QtWebEngine/5.12.8 Chrome/69.0.3497.128 Safari/537.36 (2023). https://www.r-project.org/.

45. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate : A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).