Larval diapause slows adult epigenetic ageing in an insect model, *Nasonia vitripennis*

Erin E.B. Foley[a], Christian L. Thomas[a], Charalambos P. Kyriacou[a], and Eamonn B. Mallon[a,1]

This manuscript was compiled on May 22, 2025

Epigenetic clocks based on DNA methylation provide robust biomarkers of biological age, yet the mechanistic basis and functional significance of slowing these clocks remain unclear. Progress has been limited by the lack of short-lived, genetically tractable model organisms with functional DNA methylation systems. The jewel wasp, *Nasonia vitripennis*, offers a unique solution. It combines a functional DNA methylation system with a short lifespan and established tools for experimental manipulation. We previously developed an epigenetic clock in *Nasonia*, but whether this clock reflects plastic, environmentally driven ageing processes was unknown.

Here, we test this directly by experimentally inducing larval diapause, a naturally occurring developmental arrest triggered by environmental cues. Diapause extended median adult lifespan by 36% and significantly slowed the rate of epigenetic ageing. Using whole-genome bisulfite sequencing across multiple adult timepoints, we show that while diapaused adults initially emerge epigenetically older, their subsequent epigenetic ageing proceeds 29% more slowly than non-diapaused controls.

Clock CpGs were enriched for gene ontology terms related to conserved nutrient-sensing and developmental pathways, including insulin/IGF signaling and mTOR, supporting the established mechanistic link between development and epigenetic ageing. These findings demonstrate that epigenetic ageing is plastic and can be experimentally modulated by early-life environment, establishing *Nasonia* as a tractable system for dissecting the causal mechanisms of epigenetic ageing.

Senescence plasticity | DNA methylation | Epigenetic clock | Hymenoptera

Understanding the biology of ageing is a major scientific and societal challenge. Epigenetic clocks, biomarkers based on DNA methylation, have emerged as powerful predictors of biological age and healthspan that can outperform chronological age (1, 2). Yet despite their utility, the mechanistic basis of these clocks and the biological significance of slowing epigenetic ageing remain poorly understood (3).

Progress in this area has been hindered by limitations in the current model organisms. While invertebrates like *Drosophila melanogaster* and *Caenorhabditis elegans* are invaluable for ageing research due to their genetic tractability and short lifespans, they lack detectable DNA methylation (4, 5), precluding their use for studying the relevance of DNA methylation in ageing.

The jewel wasp, *Nasonia vitripennis*, overcomes this barrier by combining a short lifespan, a well-annotated genome and a functional DNA methylation system (6, 7). We recently established an epigenetic clock in *Nasonia*, making it the first insect model with a methylation-based biomarker of ageing (8). However, whether this clock is plastic and responsive to environmental changes remains unknown.

Here, we test the plasticity of the *Nasonia* epigenetic clock using diapause, a naturally induced larval developmental arrest triggered by environmental cues. Diapause has been linked to extended adult lifespan and modulates conserved ageing pathways such as insulin/IGF signalling and mTOR (9, 10), along with broad DNA methylation reprogramming (11).

We show that exposure to larval diapause slows adult epigenetic ageing in *Nasonia vitripennis*, providing direct evidence that an invertebrate epigenetic clock is responsive to environmental inputs. This establishes *Nasonia* as a powerful new model for dissecting the mechanisms of plastic epigenetic ageing.

Author affiliations: [a]Department of Genetics, Genomics and Cancer Sciences, University of Leicester, University Road, LE1 7RH, U.K.

EBF, CPK and EBM conceived the study. EBF conducted the experiments. EBF, CLT, and EBM performed the data analysis. EBM wrote the initial draft of the manuscript. All authors contributed to manuscript revision and approved the final version.

EBM and CLT are affiliated with Vitality Labs, a spinout exploring commercial applications of insect epigenetic ageing models. This work was conducted independently of the company. All other authors declare no competing interests.

[1]To whom correspondence should be addressed. E-mail: ebm3@le.ac.uk

DRAFT

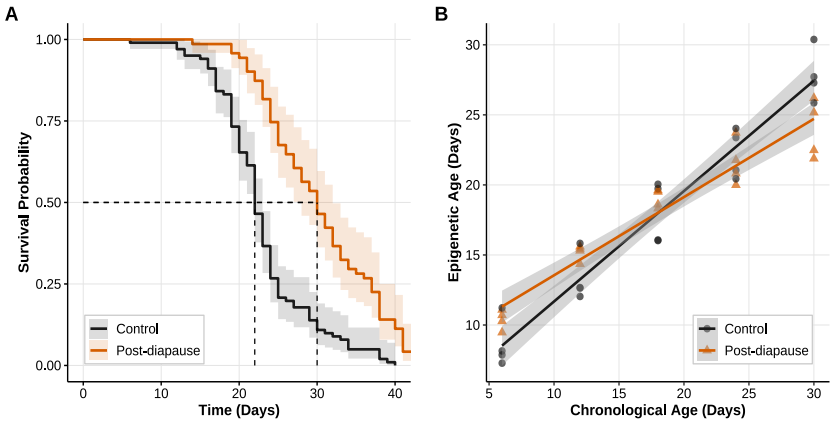

Fig. 1. Diapause effect on lifespan and epigenetic ageing. A) Adults diapaused as larvae live longer than non-diapaused conspecifics. Shaded areas represent 95% confidence intervals. Dotted lines represent median survival. B) Diapause slows adult epigenetic ageing. Shaded areas represent 95% confidence intervals.

Materials and Methods

Diapause was induced by maintaining *Nasonia vitripennis* mothers at 20°C under a 16:8 h light:dark (LD 16:8) cycle. After ten days, their larval offspring are diapaused destined. Fourth instar diapaused larvae were maintained at 4°C in constant darkness (DD) for three months before adult emergence. All experiments were carried out on virgin males.

Post-emergence, diapaused and non-diapaused (control) virgin males were individually housed (25°C, 40% humidity, LD 12:12), fed daily with 20% sucrose solution, and their survival monitored daily.

For whole-genome bisulfite sequencing (WGBS), individuals were sampled at five timepoints post emergence (days 6, 12, 18, 24, 30). Each of the 40 libraries (2 treatments x 5 timepoints x 4 replicates) represented pooled DNA from 10 individuals and included a 1% unmethylated lambda spike-in. Sequencing and bioinformatic processing, including read trimming, alignment, methylation calling, and duplicate removal, followed our standard pipeline (8). Age-related differentially methylated loci (DMLs) were identified (12) and then filtered to retain CpGs strongly correlated (Pearson's $|r| \geq 0.3$, uncorrected $p \leq 0.05$) with chronological age in control samples for epigenetic clock construction (13).

An elastic net regression was trained on age-associated CpG sites to predict chronological age providing an epigenetic clock. Predicted epigenetic age was validated against chronological age using a linear model.

Full methodological details are provided in the Supplementary Information.

Results and Discussion

Diapause treatment extended lifespan (Figure 1A), reducing the hazard of death by approximately 65% compared to controls (Cox Proportional Hazards model; Hazard Ratio [HR] = 0.35, 95% CI: 0.26–0.49; Wald test, $p = 5 \times 10^{-9}$). This was also reflected in the median survival determined by Kaplan-Meier analysis which was 30 days (95% CI: 28–32 days) for the diapause group (n=71), versus 22 days (95% CI: 22–23 days) for controls (n=101).

From 715,987 CpG sites classified as methylated (SI Dataset S1), the generalized linear model (12) identified 7,950 CpGs with significant age-related differential methylation (SI Dataset S2). We further prioritized 289 of these sites that were strongly correlated with chronological age (Pearson's $|r| \geq 0.3$, uncorrected $p \leq 0.05$) as input features for an Elastic Net regression model (SI Dataset S3). The final optimised model (mixing parameter $\alpha = 0.5$; regularization $\lambda = 1.84207$, determined by 10-fold repeated cross-validation [3 repeats] minimizing Root Mean Squared Error [RMSE]) utilized a concise panel of 27 CpGs to estimate epigenetic age (SI Dataset S4).

Our resulting epigenetic clock accurately predicted chronological age in control samples, explaining 91.7% of the variance (cross-validated R^2; RMSE = 2.44 days). Importantly, the clock also performed robustly when applied to diapause samples, accounting for 78.0% of chronological age variance (R^2; RMSE = 3.98 days), demonstrating its potential applicability across distinct physiological conditions.

Curiously, at day 6 post-eclosion, diapaused adults were epigenetically older than age-matched controls by an estimated 2.8 days (diapaused: 11.32 days vs. control: 8.53 days, t = -3.22, d.f. = 36, p = 0.0027, `emmeans` post hoc). One possible explanation is that epigenetic ageing occurs during the diapause period, albeit at a markedly reduced rate compared to adult ageing. Alternatively, the observed overshoot may reflect transient remodelling of DNA methylation during emergence from diapause. Distinguishing between these scenarios will require direct measurement of methylation dynamics during diapause itself.

Despite this initial increase in epigenetic age, diapaused adults subsequently age epigenetically 29% more slowly than controls (Figure 1B; control slope = 0.78812, diapause slope = 0.55828; linear model interaction of day and treatment: t = -3.903, d.f. = 36, p = 0.000399). By day 18, both groups converge on an epigenetic age of approximately 18 days. However, by day 30, diapaused individuals are epigenetically 2.7 days younger than controls (24.71 days vs. 27.44 days; `emmeans` post hoc: t = 3.152, d.f. = 36, p = 0.0033). This striking deceleration mirrors the observed extension in lifespan and 65% reduction in mortality hazard in diapaused individuals. Together, these results suggest that the molecular changes captured by the epigenetic clock are closely aligned with the physiological mechanisms that promote survival post-diapause.

Our findings place diapause-induced epigenetic deceleration within the broader context of early life programming, where environmental cues can reshape long-term molecular and physiological trajectories. The direction and magnitude of changes in the epigenome largely depend on early life environmental conditions (14). This is not unexpected, as the epigenome is highly plastic during early development. Such plasticity may be adaptive, allowing early life environments to reconfigure the epigenome in ways that enhance future fitness,

DRAFT

consistent with the predictive adaptive response hypothesis (15). In this framework, insect larval diapause would serve as a predictive adaptive response: an overwintering strategy that anticipates a more challenging adult environment. Adults that have passed through diapause may be under selection to survive longer, facilitating reproductive success in harsher post-winter conditions. In *Nasonia*, our data suggest that this is mirrored at the molecular level by a long-term slowing of the epigenetic clock.

Epigenetic clocks across diverse species consistently highlight key developmental gene sets as predictors of biological age (16). The CpGs comprising our *Nasonia* epigenetic clock are significantly enriched for gene ontology terms related to conserved developmental and nutrient-sensing pathways, including mTOR and insulin/IGF signaling (SI Dataset S5). These pathways are central regulators of growth, metabolism, and lifespan. Theoretical links between development and ageing have a long history in evolutionary biology (17), and experimental studies, particularly in invertebrates such as *C. elegans*, have shown that developmental alterations, such as entry into the *dauer* stage, can dramatically extend lifespan. Our findings align with recent work in *Drosophila*, where Kang *et al.* (18) demonstrated that delayed development in prothoracicotropic hormone (PTTH)-null mutants extends lifespan and postpones the onset of age-related transcriptional changes. In *Nasonia*, we observe a similar phenomenon, larval diapause, a naturally induced developmental delay, slows the progression of the epigenetic ageing clock. Notably, PTTH suppression has been implicated in diapause induction across diverse insect species (19). Together, these findings point to a conserved endocrine-epigenetic axis through which developmental timing modulates ageing trajectories.

Epigenetic ageing is influenced by inflammation, cell division, metabolic state, and early-life environment (3). With its compact genome, short lifespan, and functional methylation system, *Nasonia* enables experimental dissection of these processes in vivo (6). Our findings demonstrate that epigenetic ageing in this model is not only measurable, but developmentally modifiable. This positions *Nasonia* to address a fundamental translational question: can targeted reductions in epigenetic age improve long-term health and resilience?

ACKNOWLEDGMENTS. EBF was supported by a BBSRC MIBTP DTP studentships. CT was funded by a Leverhulme Trust award RPG-2020-363. EM was funded by a BBSRC Pioneer Award APP3335. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising

1. K Seale, S Horvath, A Teschendorff, N Eynon, S Voisin, Making Sense of the Ageing Methylome. *Nat. Rev. Genet.* **23**, 585–605 (2022) Publisher: Nature Publishing Group.
2. L Drew, Turning back time with epigenetic clocks. *Nature* **601**, S20–S22 (2022) Bandiera_abtest: a Cg_type: Outlook Number: 7893 Publisher: Nature Publishing Group Subject_term: Ageing, Society, Epigenetics.
3. CG Bell, et al., DNA methylation aging clocks: challenges and recommendations. *Genome Biol.* **20**, 249 (2019).
4. F Lyko, R Maleszka, Insects as Innovative Models for Functional Studies of DNA Methylation. *Trends genetics* **27**, 127–131 (2011).
5. CW Hu, JL Chen, YW Hsu, CC Yen, MR Chao, Trace analysis of methylated and hydroxymethylated cytosines in DNA by isotope-dilution LC-MS/MS: first evidence of DNA methylation in Caenorhabditis elegans. *The Biochem. J.* **465**, 39–47 (2015).
6. JH Werren, et al., Functional and evolutionary insights from the genomes of three parasitoid Nasonia species. *Sci. (New York, N.Y.)* **327**, 343–348 (2010).
7. X Wang, et al., Function and Evolution of DNA Methylation in Nasonia vitripennis. *PLOS Genet.* **9**, e1003872 (2013).
8. K Brink, CL Thomas, A Jones, TW Chan, EB Mallon, Exploring the ageing methylome in the model insect, Nasonia vitripennis. *BMC Genomics* **25**, 305 (2024).
9. DL Denlinger, Why study diapause? *Entomol. Res.* **38**, 1–9 (2008) _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1748-5967.2008.00139.x.
10. X Karp, Hormonal Regulation of Diapause and Development in Nematodes, Insects, and Fishes. *Front. Ecol. Evol.* **9** (2021) Publisher: Frontiers.
11. M Pegoraro, A Bafna, NJ Davies, DM Shuker, E Tauber, DNA methylation changes induced by long and short photoperiods in Nasonia. *Genome Res.* **26**, 203–210 (2016).
12. Y Park, H Wu, Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics* **32**, 1446–1453 (2016).
13. AE Teschendorff, S Horvath, Epigenetic ageing clocks: statistical methods and emerging computational challenges. *Nat. Rev. Genet.* **26**, 350–368 (2025) Publisher: Nature Publishing Group.
14. A Vaiserman, Developmental Tuning of Epigenetic Clock. *Front. Genet.* **9** (2018).
15. P Bateson, P Gluckman, M Hanson, The biology of developmental plasticity and the Predictive Adaptive Response hypothesis. *The J. Physiol.* **592**, 2357–2368 (2014).
16. D Gems, RS Virk, JP de Magalhães, Epigenetic clocks and programmatic aging. *Ageing Res. Rev.* **101**, 102546 (2024).
17. JP de Magalhães, Programmatic features of aging originating in development: aging mechanisms beyond molecular damage? *The FASEB J.* **26**, 4821–4826 (2012).
18. P Kang, et al., NF-κB-mediated developmental delay extends lifespan in Drosophila. *Proc. Natl. Acad. Sci.* **122**, e2420811122 (2025) Publisher: Proceedings of the National Academy of Sciences.
19. Q Wang, AAM Mohamed, M Takeda, Serotonin Receptor B May Lock the Gate of PTTH Release/Synthesis in the Chinese Silk Moth, Antheraea pernyi ; A Diapause Initiation/Maintenance Mechanism? *PLOS ONE* **8**, e79381 (2013).

DRAFT

Supporting Information for

Larval diapause slows adult epigenetic ageing in an insect model, *Nasonia vitripennis*

Erin B. Foley, Charalambos P. Kyriacou, Christian L. Thomas & Eamonn B. Mallon

Eamonn B. Mallon.
E-mail: ebm3@le.ac.uk

This PDF file includes:

Supporting text
Legends for Dataset S1 to S5
SI References

Other supporting materials for this manuscript include the following:

Datasets S1 to S5

Supporting Information Text

Methods

Rearing. *Nasonia vitripennis* used in this study were from the Leicester strain, a laboratory colony maintained at the University of Leicester for over nine years. This strain derives from AsymC, originally isolated in 1989 and subsequently cured of *Wolbachia* via heat shock treatment (1, 2).

To generate diapaused offspring, virgin females were housed at 20°C under a 16:8 h dark:light photoperiod. Host pupae were collected from Day 10 post-oviposition onwards to ensure recovery of diapaused larvae. These larvae developed to the fourth instar before being transferred to continuous darkness at 4°C for three months. After the diapause period, larvae were returned to standard rearing conditions and allowed to complete development to adulthood.

Lifespan Experiments. Diapaused and non-diapaused virgin male *Nasonia* were collected in batches within 24 hours of adult emergence (Day 0) and housed individually in plastic tubes. All individuals were maintained at 25°C and 40% relative humidity under a 12:12 h light:dark cycle. Each wasp received a daily feeding of 20 µL of 20% sucrose solution on filter paper, and mortality was recorded daily. Lifespan was defined as the number of days from adult emergence to death.

Kaplan–Meier survival analyses were performed using the survival package (v3.7) (3) and visualized with the survminer package (v0.4.9) (4). Cox proportional hazards models were fitted using the survival package. All analyses were conducted in R v4.4.1 (5).

Lifespan data were used to select sampling time points for whole-genome bisulfite sequencing (WGBS). Diapaused and non-diapaused virgin males were sampled at Days 6, 12, 18, 24, and 30, always at the same time of day. Upon collection, individuals were snap-frozen in liquid nitrogen and stored at −80°C until DNA extraction.

DNA Extraction. DNA was extracted using an adapted protocol based on the AllPrep® DNA/RNA Micro Kit (Qiagen). Each sample consisted of ten whole-body adult virgin male *Nasonia*. Four biological replicates were processed per time point (Days 6, 12, 18, 24, and 30).

DNA quality and concentration were assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific), a Qubit™ dsDNA BR Assay Kit (Thermo Fisher Scientific), and electrophoresis on a 1% agarose gel run for 40 minutes at 100 V. Samples were sent to Novogene (Beijing) for whole-genome bisulfite sequencing (WGBS). A 1% unmethylated lambda DNA spike-in was included in each sample to assess bisulfite conversion efficiency.

Whole-Genome Bisulfite Sequencing and Bioinformatic Processing. Raw sequencing data were provided in FASTQ format (submitted to Sequence Read Archive (SRA)). For samples sequenced across multiple lanes, read files were concatenated using the Unix `cat` command. Read quality was assessed using FastQC v0.12.1 (6).

A custom Snakemake pipeline (Snakemake v7.32.4; Python v3.12.1) (7) was used for preprocessing. Adapter sequences were trimmed and the first 10 bases removed from each read using Cutadapt v4.4 (8). Paired-end reads were aligned to the *Nasonia vitripennis* reference genome (Nvit_PSR1.1) (9) using Bowtie2 v2.5.1 with default parameters (10). Reads were also aligned to the unmethylated lambda genome (RefSeq accession: GCF_000840245.1) to assess bisulfite conversion efficiency.

Aligned reads were deduplicated and cytosine methylation calls extracted using Bismark v0.22.3 (11). Strand ambiguity was resolved using the `coverage2cytosine` utility in Bismark to generate destranded coverage files, which were used for downstream analyses.

Methylation Analysis. All downstream analyses were performed in R v4.4.1 (5). Destranded CpG coverage files were imported into the *MethylKit* package v1.30.0 (12) using the `methRead()` function. CpG sites with less than 10× coverage or coverage above the 99th percentile were filtered out. A binomial test was applied to each sample using the lambda genome conversion rate as the null probability of success, with a false discovery rate (FDR) threshold of $p < 0.05$ (SI Dataset S5). Only CpG sites showing significant methylation in at least one sample were retained. Percentage methylation at each site was calculated using the `percMethylation()` function.

Differential Methylation Analysis. Differential methylation analysis was performed using the *DSS* package v2.52.0 (13), which models methylation proportions using a beta-binomial generalized linear model (GLM) with an arcsine link function. A design matrix was constructed incorporating time point and treatment (diapause vs. non-diapause) as experimental factors. Linear models were fitted using the `DMLfit()` function to evaluate main effects and interactions. Differentially methylated loci (DMLs) and regions (DMRs) were identified accordingly.

Genomic features were assigned to CpG sites using a custom annotation file (GFF format) generated by Dr. Hollie Marshall using AGAT v0.10.0 (14).

Epigenetic Clock Construction. An elastic net regression model was trained to predict chronological age from DNA methylation levels across CpG sites, using data from virgin male *Nasonia vitripennis* sampled at five time points (Days 6, 12, 18, 24, and 30). Methylation data (percentage values) were filtered to retain CpG sites previously identified as differentially methylated across time using *DSS* (13).

To enhance model sparsity and robustness, CpG sites were pre-selected based on univariate Pearson correlation with age in the non-diapaused (control) group. Sites with absolute correlation ≥ 0.3 and uncorrected p-value ≤ 0.05 were retained. CpG features were then centered and scaled.

The final model was implemented using the `glmnet` algorithm with elastic net regularization ($\alpha = 0.5$), as part of the *caret* framework (15). Ten-fold cross-validation with three repeats was used for model tuning and performance estimation. The treatment group (diapause vs. control) was included as a categorical predictor, encoded as a dummy variable, and entered alongside the CpG methylation features. The optimal model was selected based on minimum root mean squared error (RMSE). Predictions were generated for all samples, and model performance was assessed using RMSE and R^2, stratified by treatment group.

Post-hoc analysis of predicted age trajectories was conducted using linear models with interaction terms, and estimated marginal means were compared between treatment groups at multiple time points using the *emmeans* package (16).

Gene Ontology Enrichment Analysis. Gene Ontology (GO) enrichment analysis was performed to assess the functional significance of genes associated with CpG sites included in the final epigenetic clock model. The 27 clock CpGs (SI Dataset S4) were mapped to nearby genes, which were then tested for enrichment against a background set comprising all genes associated with differentially methylated loci across time (SI Dataset S1). GO annotations were derived from *Nasonia vitripennis* and formatted for compatibility with *GOstats* (17).

The gene list was tested for overrepresentation of GO terms in the Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) ontologies using a conditional hypergeometric test. Analyses were performed using the *GOstats* and *GSEABase* packages. For each ontology, both over- and under-representation were tested, and GO terms with an adjusted FDR of < 0.05 (Benjamini-Hochberg) were considered significant.

To visualise and summarise redundant GO terms, semantic similarity clustering was performed using the *rrvgo* package (18). Pairwise GO term similarities were calculated using the *org.Dm.eg.db* annotation database and the "Rel" semantic similarity measure. Representative GO terms were identified by reducing the similarity matrix with a similarity threshold of 0.7. Enriched terms were visualised using treemaps, heatmaps, and scatter plots. Final results were exported in tabular form (SI Dataset S5).

All scripts used are available at https://tinyurl.com/5n6vcvsk

SI Dataset S1 (SI_Datasets.xlsx/S1_Methylated_CpGs)

CpG sites classified as methylated using a binomial test with the lambda genome conversion rate as the null probability of success. Sites passing a false discovery rate (FDR) threshold of $p < 0.05$ and significantly methylated in at least one sample were retained.

SI Dataset S2 (SI_Datasets.xlsx/S2_Age_DMPs)

7,950 CpG sites showing significant age-associated differential methylation.

SI Dataset S3 (SI_Datasets.xlsx/S3_Age_Correlated)

289 age-associated CpG sites with strong correlation to chronological age (Pearson's $|r| \geq 0.3$, uncorrected $p \leq 0.05$).

SI Dataset S4 (SI_Datasets.xlsx/S4_Clock_Coefficients)

Elastic net model coefficients for the 27 CpG sites and intercept used to estimate epigenetic age.

SI Dataset S5 (SI_Datasets.xlsx/S5_GO_terms)

Enriched gene ontology terms associated with the 27 CpG sites comprising the epigenetic clock.

References

1. JH Werren, DW Loehlin, The parasitoid wasp Nasonia: an emerging model system with haploid male genetics. *Cold Spring Harb. Protoc.* **2009**, pdb–emo134 (2009) Publisher: Cold Spring Harbor Laboratory Press.
2. DC Darling, JH Werren, Biosystematics of Nasonia (Hymenoptera: Pteromalidae): two new species reared from birds' nests in North America. *Annals Entomol. Soc. Am.* **83**, 352–370 (1990) Publisher: Oxford University Press Oxford, UK.
3. Terry M. Therneau, Patricia M. Grambsch, *Modeling Survival Data: Extending the Cox Model.* (Springer, New York), (2000).
4. A Kassambara, M Kosinski, P Biecek, *survminer: Drawing Survival Curves using 'ggplot2'.* (2021).
5. R Core Team, *R: A Language and Environment for Statistical Computing.* (R Foundation for Statistical Computing, Vienna, Austria), (2024).
6. S Andrews, et al., FastQC (2012) Place: Babraham, UK Published: Babraham Institute.
7. J Köster, S Rahmann, Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012) _eprint: https://academic.oup.com/bioinformatics/article-pdf/28/19/2520/48879301/bioinformatics_28_19_2520.pdf.
8. M Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
9. E Dalla Benetta, et al., Genome elimination mediated by gene expression from a selfish chromosome. *Sci. Adv.* **6**, eaaz9808 (2020) Publisher: American Association for the Advancement of Science.
10. B Langmead, SL Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. methods* **9**, 357–359 (2012) Publisher: Nature Publishing Group.

11. F Krueger, SR Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *bioinformatics* **27**, 1571–1572 (2011) Publisher: Oxford University Press.
12. A Akalin, et al., methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome biology* **13**, 1–9 (2012) Publisher: Springer.
13. H Feng, H Wu, Differential methylation analysis for bisulfite sequencing using DSS. *Quant. Biol.* **7**, 327–334 (2019) Publisher: Springer.
14. J Dainat, AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format.(Version v0. 7.0). *Zendo. doi* **10** (2023).
15. M Kuhn, Building Predictive Models in R Using the caret Package. *J. Stat. Softw.* **28**, 1–26 (2008).
16. RV Lenth, emmeans: Estimated marginal means, aka least-squares means, manual (2022).
17. S Falcon, R Gentleman, Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257–258 (2007) Publisher: Oxford University Press.
18. S Sayols, rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. *microPublication Biol.* (2023) Publisher: Caltech Library.