

The epigenetic and regulatory hallmarks of ageing and the role of germline-soma interaction

Jean-Charles William Marie Gabriel de Coriolis

Thesis submitted for the degree of Doctor of Philosophy

University of East Anglia
School of Biological Sciences

January 2026



©This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefore must be in accordance with current UK copyright law. In addition, any quotation or extract must include full attribution.

Thesis Abstract

This thesis investigates how germline–soma communication shapes molecular ageing in a vertebrate system, using zebrafish to integrate transcriptomic analysis, DNA methylation–based ageing biomarkers, and pharmacological intervention. The overarching aim is to determine how reproductive status influences somatic regulatory programmes and epigenetic ageing trajectories across tissues and lifespan. Chapter 2 demonstrates that germline removal induces a consistent reorganisation of somatic gene expression. Germline-free fish upregulate DNA repair, cell-cycle control, and genome maintenance pathways across multiple tissues, whereas germline-carrying fish prioritise metabolic and growth-associated processes. Following genotoxic stress, germline-free fish activate somatic repair responses more rapidly, while testes in germline-carrying fish show the strongest repair response overall, indicating tissue-specific allocation of maintenance investment. Chapter 3 develops and benchmarks over 3,000 cross-sectional epigenetic clock models, demonstrating that clock accuracy depends strongly on tissue context and modelling strategy. Chapter 4 applies optimised cross-sectional clocks across multiple tissues and identifies a robust subset of CpG sites with shared age-associated trajectories alongside extensive tissue-specific methylation patterns, indicating that epigenetic ageing comprises both conserved and tissue-restricted components. Chapter 5 adopts a longitudinal framework to track within-individual epigenetic ageing over time. Germline-free fish are epigenetically younger early in life but exhibit steeper late-life trajectories. These shifts are associated with age- and germline-dependent hypomethylation at robustly selected CpG sites and occur alongside preserved locomotor function in old age. Chapter 6 examines rapamycin effects across two experimental designs. Short-term late-life exposure reveals sex-specific epigenetic age acceleration accompanied by

reduced male fertility and altered behaviour. Lifelong exposure identifies increased ovarian pathology in genome-unstable fish alongside reduced anxiety-like behaviour in wild-type animals. Collectively, these findings establish germline status as a key regulator of somatic epigenetic ageing dynamics, highlight fundamental differences between cross-sectional and longitudinal ageing measures. This demonstrates that interventions targeting nutrient-sensing pathways have context-dependent and genotype-specific effects.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence. Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Acknowledgements

I am deeply grateful to all those who have supported and encouraged me throughout this work.

I would like to thank Simone, who gave me my first opportunity as a research technician and later supported my PhD. Although this project did not always follow a smooth course, I am grateful for the opportunities, guidance, and intellectual challenges you provided. Your insistence on rigour, clarity of thought, and independence has had a lasting influence on how I approach science, and has made me a better researcher.

I would like Sara and Alice to know how grateful I am for the advice, guidance, and support you consistently offered, always with generosity and good humour.

I wish to give special acknowledgement to Kelly, whose support has extended across twelve years of my academic life. Your encouragement, perspective, and unwavering positivity, particularly during the most difficult moments, have been indispensable, and I would not have reached this stage without you. Thank you.

Over the course of my time in the department, I have been fortunate to work alongside and form friendships with many remarkable people. To name a few: Mabel, David M., Marco, Sara, James, Elisa, Liz, Anastasia, Justin, Heather, Isaac, Charlotte, Janine, Helen, Ben G., Maria-Elena, Phill, Iain, Ben L., Matt, and Jacqueline - thank you for your friendship, wit, and kindness.

Finally, I would like to thank my family. Joe, for his love, patience, and perspective. 愛してる. My parents, for their constant energy, humour, and support. And to Shelley, Henry, and my goddaughter Beatrice, thank you for your kindness and for the support you always show me.

Table of Contents

1. Chapter 1.....	20
1.1. Conceptual and molecular foundations of ageing	20
1.2. Evolutionary perspectives of ageing.....	22
1.2.1. Hallmarks of Ageing	22
1.2.2. Key physiological theories	24
1.3. Biological functions of ageing.....	27
1.3.1. Somatic mutation accumulation and genomic instability	27
1.3.2. Protein homeostasis and the ubiquitin–proteasome system.....	29
1.3.3. Soma–germline dynamics and conserved signalling pathways	30
1.4. Epigenetic clocks as biomarkers of ageing	32
1.4.1. Development and mechanistic basis of DNA methylation clocks	33
1.4.2. Mechanisms of age-related methylation change and clock interpretation.....	34
1.4.3. Applications of epigenetic clocks.....	34
1.5. Cellular senescence and therapeutic approaches	35
1.5.1. Mechanisms of cellular senescence and the senescence- associated secretory phenotype (SASP).....	35
1.5.2. Consequences of senescent cell accumulation in tissues and immune surveillance.....	37
1.5.3. Therapeutic strategies: senolytics, senomorphics, and the central role of rapamycin	38
1.6. Zebrafish as a model for ageing research.....	40
1.7. Integration and future directions.....	42
1.8. References	43
2. Chapter 2.....	57

2.1. Abstract	57
2.2. Introduction	58
2.3. Material and methods.....	61
2.3.1. Fish lines and maintenance	61
2.3.2. Germline-free fish	61
2.3.3. Exposure to gamma irradiation.....	61
2.3.4. Sample collection	62
2.3.5. RNA extraction and sequencing	62
2.3.6. Differential gene expression analysis	62
2.3.7. Gene Ontology and chromatin analysis	63
2.4. Results.....	63
2.4.1. Germline ablation reorganises baseline somatic transcription ..	63
2.4.2. Baseline differences in DREAM complex architecture	72
2.4.3. Germline ablation accelerates and redirects the somatic response to genotoxic stress	75
2.4.4. A temporal 'directional flip' in genomic features after irradiation	79
2.4.5. Tissue-specific programmes reveal rapid adaptation and repair in GLF	80
2.4.6. Testes trajectories contextualise somatic responses	80
2.4.7. Temporal control by DREAM and hedgehog signalling	81
2.4.8. Germline ablation reshapes somatic transposable-element dynamics.....	81
2.5. Discussion	86
2.6. Data availability	88
2.7. Author contributions.....	89
2.8. Ethics declarations.....	89
2.9. Extended data	90

2.10. Abbreviations	90
2.11. References.....	91
3. Chapter 3.....	100
3.1. Abstract	100
3.2. Introduction	101
3.2.1. Approaches to clock construction	103
3.2.2. CpG pre-selection and randomness	105
3.2.3. Tissue scope: single tissue and multi-tissue models	105
3.2.4. Model of choice	106
3.2.5. Balancing predictive accuracy and biological meaning	107
3.3. Material and Methods.....	108
3.3.1. Fish Lines and Maintenance	108
3.3.2. Sample Collection	108
3.3.3. DNA Extraction.....	109
3.3.4. Library Preparation.....	109
3.3.5. RRBS sequencing.....	110
3.3.6. Data Trimming, Alignment and Methylation Extraction	110
3.3.7. Methylation data pre-processing	111
3.3.8. Epigenetic Clock Model Construction and Evaluation.....	111
3.3.9. Epigenetic clock model training and selection.....	112
3.3.10. Technical parameters for epigenetic clock construction	113
3.3.11. Analytical Rationale	114
3.4. Results.....	115
3.4.1. CpG selection strategies reveal tissue × method interactions.....	115
3.4.2. CpG dosage effects expose tissue-specific scaling limits and overfitting risks.....	119
3.4.3. Global design hierarchy prioritises feature selection over model complexity.....	120

3.4.4. Multi-tissue clocks decouple metrics and destabilise performance rankings.....	121
3.5. Discussion	126
3.5.1. CpG scaling reveals tissue-specific efficiency limits	127
3.5.2. Multi-tissue clocks require different design paradigms.....	128
3.5.3. Implications for clock benchmarking and reproducibility	129
3.5.4. Conclusion	130
3.6. Data Availability	131
3.7. Contributions	131
3.8. Ethics declaration.....	131
3.9. Extended data	132
3.10. Abbreviations	132
3.11. References.....	132
4. Chapter 4.....	141
4.1. Abstract	141
4.2. Introduction	142
4.3. Material and Methods	145
4.3.1. Genomic feature annotation any methylation quantification... 145	
4.4. Results.....	146
4.4.1. Epigenetic clock performance is strongly shaped by tissue context and CpG aggregation strategy.....	146
4.4.2. Tissue-specific differences in epigenetic age acceleration and ageing trajectories	149
4.4.3. Tissue-specific and overlapping CpG architectures.....	152
4.4.4. Functional enrichment of overlapping CpGs.....	153
4.4.5. Methylation density, stochastic drift, and entropy	157

4.5. Discussion	161
4.5.1. Tissue context strongly constrains epigenetic clock performance 161	
4.5.2. Epigenetic age acceleration reflects tissue-specific cellular histories.....	163
4.5.3. Shared CpGs encode conserved regulatory ageing signals	163
4.5.4. Divergent drift and entropy reveal tissue-specific ageing modes	165
4.5.5. Conclusion	165
4.6. Data Availability	166
4.7. Contributions	166
4.8. Ethics declaration.....	167
4.9. Extended data	167
4.10. Abbreviations	167
4.11. References.....	168
5. Chapter 5.....	177
5.1. Abstract	177
5.2. Introduction	178
5.3. Material and methods.....	183
5.3.1. Fish Lines and Maintenance	183
5.3.2. Germline-free fish	183
5.3.3. Sample Collection	184
5.3.4. DNA extractions.....	184
5.3.5. Library preparation and sequencing	184
5.3.6. Data trimming, alignment and methylation Extraction	185
5.3.7. Methylation data pre-processing	185
5.3.8. Epigenetic clock model training and selection.....	186
5.3.9. Behaviour tracking and speed analysis	187

5.4. Results.....	188
5.4.1. CpG aggregation strategy shapes epigenetic clock performance	188
5.4.2. Identification and methylation dynamics of robust CpG sites..	190
5.4.3. Functional context of robust CpG sites	193
5.4.4. Germline ablation alters somatic epigenetic ageing trajectories	196
5.4.5. Germline ablation influences stochastic methylation drift.....	199
5.4.6. Late-life behavioural correlates of epigenetic ageing.....	199
5.5. Discussion	201
5.5.1. Germline ablation reshapes somatic epigenetic ageing trajectories	201
5.5.2. Developmental gene regulation as a substrate of epigenetic ageing	203
5.5.3. Stochastic methylation drift and loss of homeostatic control ..	204
5.5.4. Molecular ageing and behavioural outcomes.....	206
5.5.5. CpG organisation and the structure of ageing signals	208
5.5.6. Conclusion	209
5.6. Data Availability	210
5.7. Contributions	210
5.8. Ethics declaration.....	210
5.9. Extended data	211
5.10. Abbreviations	211
5.11. References.....	211
6. Chapter 6.....	222
6.1. Abstract	222
6.2. Introduction	223

6.3. Material and methods.....	227
6.3.1. Shared procedures	227
6.3.2. Experiment 1: Ageing Experiment: Short-term Rapamycin exposure in late life	229
6.3.3. Experiment 2: Lifelong Exposure experiment (chronic rapamycin treatment in ABWT and msh2 background)	233
6.3.4. Rapamycin blood concentration measurements.....	237
6.4. Results.....	237
6.4.1. Experiment 1: Male reproductive success is reduced by high-dose rapamycin	238
6.4.2. Experiment 1: Low and High-dose rapamycin produce distinct locomotor behaviours in late-life	240
6.4.3. Experiment 1: Rapamycin accelerates epigenetic ageing with dose and sex-specific effects.....	243
6.4.4. Experiment 2: Chronic rapamycin increases the prevalence of macroscopic ovarian lesions in MMR-deficient fish	247
6.4.5. Experiment 2: High and low-dose rapamycin differentially affect locomotor ageing trajectories.....	247
6.4.6. Experiment 2: Rapamycin modifies stress-related behaviours in a strain-dependent manner.....	248
6.5. Discussion	251
6.5.1. Rapamycin and epigenetic ageing.....	251
6.5.2. Behavioural ageing under mTOR inhibition	253
6.5.3. Reproductive suppression and ovarian lesions in a genome-unstable context.....	254
6.5.4. Comparative context of rapamycin effects on ageing	256
6.5.5. Conclusions.....	259
6.6. Data Availability.....	259

6.7. Contributions	260
6.8. Ethics declaration.....	260
6.9. Extended data	260
6.10. Abbreviations	260
6.11. References.....	261
7. Chapter 7.....	269
7.1. Overview.....	269
7.2. Conceptual conclusions.....	270
7.2.1. Ageing trajectories are regulated, context-dependent processes 270	
7.2.2. The germline acts as a regulator of somatic ageing, not merely a cost 271	
7.2.3. Epigenetic clocks measure different aspects of ageing, not a single truth 272	
7.2.4. Interventions reshape ageing trajectories rather than uniformly slowing them.....273	
7.3. Integration with ageing theory.....	274
7.4. Conceptual questions raised by this work.....	274
7.5. Final remarks	276
8. Appendix Chapter two.....	286
8.1. Material and Methods.....	286
8.1.1. Fish husbandry details.....	286
8.1.2. CRISPR/Cas9-mediated germline ablation in zebrafish.....	286
8.1.3. sgRNA design and template preparation	287
8.1.4. In vitro transcription of sgRNA.....	287
8.1.5. Preparation of Cas9 mRNA	288

8.1.6.	Microinjection	288
8.1.7.	Verification of mutagenesis and germline loss	288
8.1.8.	Gamma irradiation procedure	289
8.1.9.	RNA extraction and library preparation	289
8.1.10.	RNA-seq processing pipeline	290
8.1.11.	Differential gene expression analyses	290
8.1.12.	Transposable element (TE) annotation and analysis	291
8.1.13.	Chromatin and gene structure analyses	291
8.2.	Supplementary Tables	292
8.3.	Supplementary Figures	336
9.	<i>Appendix Chapter three</i>	349
9.1.	Supplementary tables	349
9.2.	Supplementary Figure	362
10.	<i>Appendix Chapter Four</i>	363
10.1.	Supplementary Tables	363
10.2.	Supplementary Figures	381
11.	<i>Appendix Chapter Five</i>	384
11.1.	Supplementary Tables	384
11.2.	Supplementary Figures	393
12.	<i>Appendix Chapter Six</i>	395
12.1.	Supplementary Tables	395
12.2.	Supplementary Figures	413

TABLE OF FIGURES

FIGURE 1: GERMLINE ABSENCE ENABLES RAPID SOMATIC SHIFT FROM DAMAGE RESPONSE TO GENOMIC FIDELITY PROGRAMMES POST-IRRADIATION.....	67
FIGURE 2: GERMLINE ABLATION ELICITS A DYNAMIC EXPRESSION OF GENES AND CHROMATIN STRUCTURE IN RESPONSE TO IRRADIATION.....	69
FIGURE 3 : BIDIRECTIONAL SHIFTS IN GENE REGULATION REVEAL TEMPORALLY DISTINCT SOMATIC RESPONSES TO GERMLINE ABLATION AND IRRADIATION.	71
FIGURE 4: GLF FISH SHOW DISTINCT HEDGEHOG REGULATORY GENE AND DREAM COMPLEX TARGET GENES EXPRESSION PATTERNS AT 3HPIR.	74
FIGURE 5: TREATMENT-DEPENDENT CONSERVATION OF AGEING-RELATED GO TERMS ACROSS ZEBRAFISH STRAINS.....	78
FIGURE 6: AGE-DEPENDENT TRANSPOSABLE ELEMENT EXPRESSION IN SOMATIC TISSUES SHOWS STRAIN- AND TIMEPOINT-SPECIFIC DIRECTIONAL SHIFTS FOLLOWING IRRADIATION.	85
FIGURE 7: COMPREHENSIVE DESIGN SPACE MAPPING REVEALS TISSUE X METHOD INTERACTIONS AND CpG SCALING LIMITS.....	118
FIGURE 8: MULTITISSUE CLOCKS DECOUPLE METRICS AND EXHIBIT PERVASIVE RANK INSTABILITY.....	125
FIGURE 9: TISSUE AND CpG AGGREGATION STRONGLY INFLUENCE CLOCK PERFORMANCE SCORE.....	148
FIGURE 10: TISSUE STRONGLY IMPACTS Δ -AGE AND THE RATE OF EPIGENETIC AGEING.	151
FIGURE 11: TISSUE-SPECIFIC CpGs AND METHYLATION DYNAMICS.	156
FIGURE 12: SHARED SOMATIC CpG METHYLATION DYNAMICS ACROSS TISSUES.....	159

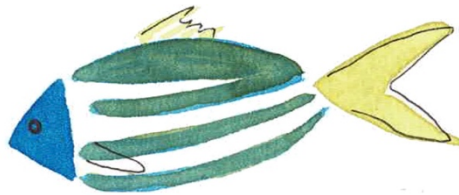
FIGURE 13: SYSTEM-LEVEL DIVERGENCE OF EPIGENETIC AGEING ACROSS TISSUES.....	160
FIGURE 14: EPIGENETIC CLOCK PERFORMANCE ACROSS CpG AGGREGATION STRATEGIES.....	189
FIGURE 15: AGE-DEPENDENT TREATMENT EFFECT IN CpG METHYLATION.	192
FIGURE 16: SYSTEM-LEVEL DIVERGENCE OF CpG METHYLATION IN CAUDAL FIN.....	195
FIGURE 17: GERMLINE ABLATION SIGNIFICANTLY ALTERS DELTA AGE TRAJECTORIES AND RESIDUAL STRUCTURE.	198
FIGURE 18: BEHAVIOURAL PHENOTYPES OF GLC AND GLF ZEBRAFISH.....	200
FIGURE 19 : REPRESENTATIVE OVARIAN MORPHOLOGIES USED TO GROSS LESION CLASSIFICATION	236
FIGURE 20: RAPAMYCIN EXPOSURE ALTERS CIRCULATING CONCENTRATIONS, REPRODUCTIVE PERFORMANCE AND OVARIAN LESION INCIDENCES.	239
FIGURE 21: BEHAVIOUR METRICS IN AGED AND LIFELONG EXPERIMENTS.....	241
FIGURE 22 RAPAMYCIN DRASTICALLY EFFECTS EPIGENETIC AGEING IN THE AGED EXPERIMENT	246
FIGURE 23: EFFECTS OF RAPAMYCIN ON ANXIETY-RELATED BEHAVIOUR IN NOVEL TANK DIVING TEST.....	250

TABLE OF TABLES

TABLE 1 : LINEAR MIXED MODEL SHOWING DOSE-DEPENDENT EFFECTS OF RAPAMYCIN ON LOCOMOTOR BEHAVIOUR.....	242
---	-----

CHAPTER ONE

Introduction: Integrative perspectives on ageing, evolutionary theories, epigenetic clocks and therapeutic advances



1. Chapter 1

1.1. Conceptual and molecular foundations of ageing

Ageing manifests as a gradual decline in physiological function, increased disease susceptibility, and elevated mortality risk across an organism's lifespan (López-Otín et al., 2013). This progressive deterioration affects fertility and multiple tissue systems, leading to progressive tissue dysfunction (Mchugh and Gil, 2018, Partridge and Barton, 1993). Traditionally, ageing was considered primarily a feature of protected environments where extrinsic mortality is minimised, such as in modern human societies or laboratory conditions (Medawar, 1952, Kirkwood, 1977). Under this view, most organisms in the wild were assumed to die from predation, disease, or environmental hazards before senescence could become apparent. However, ageing is now well-documented across wild populations of birds, mammals, and invertebrates (Ivimey-Cook and Moorad, 2020, Nussey et al., 2008), from declining reproductive success in European kestrels (Blas et al., 2009) to actuarial senescence in red deer (Nussey et al., 2006). These observations demonstrate that senescence is a pervasive biological phenomenon rather than an artefact of captivity or extended lifespan.

From an evolutionary perspective, senescence arises because the force of natural selection declines with age, allowing late-acting deleterious effects to persist in populations. Two classical genetic theories formalise this principle. Medawar's mutation accumulation (MA) theory proposes that mutations whose harmful effects are expressed only in late life can drift to high frequency because selection against them is beyond the reproductive period (Medawar, 1952). Complementarily, "Williams" antagonistic pleiotropy (AP) theory posits that alleles conferring fitness benefits early in life are favoured by selection even if they incur costs later in life, generating a trade-

Chapter 1 - Conceptual and molecular foundations of ageing

off between early-life performance and late-life survival (Williams, 1957). Together, these mechanisms provide a genetic explanation for why ageing is an intrinsic outcome of biological systems rather than an artefact of protected environments.

This evolutionary logic is reflected molecularly through conserved "hallmarks of ageing," including genomic instability (López-Otín et al., 2013), telomere attrition (Lee et al., 1998, Rudolph et al., 1999), and changes in cellular metabolism (Jinesh et al., 2025), amongst others. However, the identification of these hallmarks has not resolved fundamental conceptual disagreements, and ageing remains fragmented in its interpretation, with debate over whether it represents a pathological condition (Kennedy et al., 2014), a genetically programmed process (Kirkwood and Melov, 2011), or the cumulative results of stochastic molecular damage (Harman, 1981). This lack of consensus has contributed to disciplinary separation between evolutionary biology, molecular gerontology and biomedical research. A major development bridging the different perspectives is the emergence of epigenetic clocks, particularly those based on DNA methylation (DNAm) patterns, which provide accurate estimates of biological age distinct from chronological time (Horvath, 2013). These models have revealed molecular signatures associated with ageing, such as inflammation-linked hypermethylation, and offer a quantitative framework for assessing interventions targeting ageing, from dietary restriction to pharmacological approaches (Levine et al., 2018).

This thesis aims to integrate the evolutionary and molecular perspectives by evaluating key theories of ageing, advances in epigenetic biomarker research, and treatments targeting cellular senescence, with a particular focus on rapamycin. Through this synthesis, a comprehensive understanding of the ageing process is developed, providing insight into the mechanisms driving it.

1.2. Evolutionary perspectives of ageing

Ageing is fundamentally shaped by evolutionary forces that prioritise individual fitness, maximising lifetime reproductive success over indefinite longevity. Natural selection acts most strongly on traits that directly affect reproductive success during the prime reproductive window, and its influence progressively wanes after peak reproduction, a phenomenon framed as Haldane's "selection shadow" in which post-reproductive traits experience negligible selective pressure (Haldane, 1941). This decline allows latent deleterious mutations and maladaptive traits to accumulate unchecked in populations and manifest phenotypically later in life, leading to senescence (Medawar, 1952, Williams, 1957).

1.2.1. Hallmarks of ageing

Central to this evolutionary process are several interconnected and evolutionarily conserved molecular and cellular mechanisms, collectively termed the "hallmarks of ageing", which together drive the progressive functional decline observed across tissues and tissue systems (López-Otín et al., 2013, Kenyon, 2010). These hallmarks have been organised into three broad categories: the primary, the antagonistic and the integrative hallmarks (López-Otín et al., 2013).

Primary hallmarks comprise the initial forms of molecular damage which trigger the ageing response. These include genomic instability from unrepaired DNA lesions, point mutations, bulky adducts, and double-strand breaks that progressively compromise genome integrity (López-Otín et al., 2013, López-Gil et al., 2023). This category also includes other forms of damage, such as telomere attrition, triggering replicative senescence through chromatin shortening (Lee et al., 1998, Rudolph et al., 1999); and

Chapter 1 - Evolutionary perspectives of ageing

epigenetic alterations, encompassing global DNA hypomethylation, destabilising chromatin structure, and promoter hypermethylation silencing tumour suppressor genes (López-Otín et al., 2013).

Antagonistic hallmarks reflect adaptive cellular responses that become pathogenic when chronically activated. Proteostasis loss permits toxic protein aggregates to accumulate, as seen with amyloid- β and tau in Alzheimer's disease (López-Otín et al., 2013); as well as deregulated nutrient sensing, particularly chronic hyperactivation of insulin/IGF-1 and mTOR pathways, which causes prioritisation of anabolic growth over cellular repair and stress resistance (Kenyon, 2010, Jinesh et al., 2025).

Integrative hallmarks manifest as downstream processes that consolidate damage from primary and antagonistic mechanisms. A main example of this category is mitochondrial dysfunction. This condition, driven by electron transport chain defects, generates excess reactive oxygen species (ROS) that disrupt cellular energy metabolism and redox balance (López-Otín et al., 2013). Cellular senescence follows similar processes. Through stable p16INK4a/p21CIP1-mediated cell cycle arrest, senescent cells acquire a pro-inflammatory senescence-associated secretory phenotype (SASP), driving tissue remodelling and chronic inflammation upon accumulation (López-Otín et al., 2013). Similarly, altered intercellular communication manifests as "inflammageing" with elevated IL-6 and TNF- α cytokines, disrupting endocrine, paracrine, and neuronal signalling networks to promote tissue-specific pathologies like fibrosis and neuroinflammation (López-Otín et al., 2013).

Together, these hallmarks provide a mechanistic framework that links evolutionary theories of declining selection with the concrete cellular and tissue pathologies observed during ageing, thereby bridging conceptual models and empirical biology.

1.2.2. Key physiological theories

Among the most influential physiological models that link evolutionary genetics to soma–germline interactions and the cellular hallmarks of ageing is the Disposable Soma theory (DST), proposed by Tom Kirkwood in 1977 (Kirkwood, 1977). DST posits that organisms possess a finite pool of energetic and metabolic resources that must be allocated between two core functions: maintaining the somatic body (soma) through continuous repair of molecular damage and reproducing via the germline to ensure gene transmission to future generations (Kirkwood, 1977). In natural environments where extrinsic mortality from predation, infection, or starvation frequently reduces lifespan, natural selection is predicted to favour substantial early-life investment in reproduction over costly, open-ended somatic maintenance (Kirkwood, 2017, Kirkwood, 1977). Under this framework, the soma is effectively “disposable” once successful germline transmission has occurred, and the progressive accumulation of unrepaired somatic damage, such as lipid peroxidation, protein glycation, and DNA crosslinks, is viewed as an inevitable consequence that drives the ageing process (Kirkwood, 1977, Kirkwood, 2017, Hughes, 2010). DST therefore emphasises that the primary evolutionary driver of ageing is the maximisation of lifetime reproductive fitness under ecological constraints, in which the maintenance of perfect cellular integrity would be prohibitively expensive and confer diminishing returns beyond the reproductive period.

Building on DST, the expensive germline hypothesis refines this perspective by drawing attention to the exceptionally high cost of maintaining germline fidelity relative to somatic upkeep. Germ cells require specialised DNA repair pathways, stringent control of meiotic recombination, robust telomere maintenance, and protection from mutational and oxidative insults to

Chapter 1 - Evolutionary perspectives of ageing

preserve genomic integrity across generations (Maklakov and Immler, 2016). According to this hypothesis, preferential allocation of resources to germline protection necessarily constrains the energy available for somatic repair, thereby shaping lifespan and ageing trajectories across species (Maklakov and Immler, 2016). In this way, the expensive germline hypothesis complements DST by clarifying the germline's privileged hierarchical position in evolutionary fitness, without requiring that soma and germline always trade off in a strictly zero-sum manner.

Complementing these resource-focused models, the developmental theory of ageing proposed by Chapman and Maklakov, integrates a mechanistic account of how senescence emerges from ongoing germline–soma conflicts (Maklakov and Chapman, 2019). Late-life pathologies are interpreted as consequences of incomplete downregulation or mistiming of regulatory programmes that are beneficial earlier in life, such as growth and reproductive signalling pathways. For example, persistent activation of Hedgehog or IGF1 signalling in post-reproductive tissues can drive hyperfunction, that is continued growth, proliferation, or secretory activity, which becomes maladaptive in later life, contributing to fibrosis, hypertrophy, or tumour promotion when purifying selection is weakened (Lemaître et al., 2024). This framework therefore shifts emphasis away from simple “energetic trade-offs” and towards context-dependent physiological states, in which extrinsic mortality risk and ecological conditions modulate how strongly developmental and reproductive programmes are expressed across the lifespan.

A growing body of empirical evidence supports these physiological perspectives. Species with delayed or reduced reproductive investment, such as many bat species, live substantially longer than rodents of comparable body size and metabolic rate, consistent with stronger selection for somatic maintenance when reproduction is postponed (Austad, 1993,

Chapter 1 - Evolutionary perspectives of ageing

Lemaître et al., 2015). Experimental caloric restriction in laboratory mice, which mimics ecological food limitation, reliably extends median lifespan by 30–50% while attenuating reproductive output and downregulating growth-promoting pathways such as IIS and mTOR, in line with predictions that reduced reproductive effort allows increased allocation, or regulatory prioritisation towards repair and maintenance (Di Francesco et al., 2024, Mitchell et al., 2024). Together, DST, the expensive germline hypothesis, and the developmental theory of ageing offer complementary, rather than competing, explanations of ageing's evolutionary logic. DST and the expensive germline hypothesis highlight the fitness consequences of differential investment in soma versus germline, whereas the physiological theory explains how age-specific gene regulation and developmental programmes implement those investments at the molecular and organismal levels.

Despite their explanatory power, important questions remain regarding the detailed molecular mechanisms through which selection shapes these physiological processes. For example, the precise gene-regulatory networks that mediate germline-derived signals to somatic tissues, the roles of specific transcription factors and hormonal axes, and the extent to which these pathways exhibit plasticity in response to environmental variation are still being defined. Ongoing work is therefore increasingly integrative, combining comparative evolutionary analyses, high-throughput molecular approaches such as single-cell RNA sequencing and chromatin profiling, alongside computational modelling to connect genotype, regulatory architecture, physiological function and demographic ageing patterns within a unified framework (Lemaître et al., 2024).

1.3. Biological functions of ageing

Ageing, or senescence, is not only an evolutionary outcome but also a process that is concretely initiated through changes in cells and tissues. While evolutionary theories frame ageing in terms of declining selection and germline–soma conflicts, a mechanistic understanding of how ageing manifests at the organismal level requires careful examination of its cellular and molecular underpinnings. In this proximate view, ageing can be described as the cumulative result of progressive molecular damage, spanning DNA, proteins, lipids and organelles, together with declining repair and quality-control capacity, leading to a gradual loss of cellular, tissue and systemic function over the lifespan; and a transition from youthful homeostasis and resilience to increased frailty and vulnerability (Gladyshev et al., 2021).

1.3.1. Somatic mutation accumulation and genomic instability

A principal contributor to age-related cellular dysfunction is genomic instability, defined as the increasing burden of DNA damage and de novo somatic mutations that accrue specifically in somatic cells over time. This is conceptually distinct from the germline mutations central to Medawar's mutation accumulation theory, which act at the population level. Somatic mutations arise during mitotic divisions and from continual chemical and physical insults, and although mutation is the raw material for evolutionary adaptation at the species level (Barrick and Lenski, 2013; Wiser et al., 2013), at the level of the individual, they can have deleterious consequences. Accumulating de novo mutations can generate extensive genomic alterations, such as aneuploidy or loss of heterozygosity, that compromise genome architecture, impair gene dosage and regulatory networks, and ultimately reduce cellular viability and tissue function. The somatic mutation

Chapter 1 - Biological functions of ageing

rate is relatively constant across adult life and reflects the combined effects of intrinsic DNA replication errors (with an error rate on the order of 10^{-10} per base pair per replication), stochastic mitotic events, and persistent exposure to genotoxic agents (Lynch, 2010; Drake, 1991).

These genotoxic agents can arise from two principal sources, namely endogenous processes intrinsic to cellular metabolism and exogenous environmental exposures. Endogenous damage is generated continuously within cells, primarily through oxidative stress caused by ROS produced as by-products of mitochondrial oxidative phosphorylation. ROS such as superoxide and hydroxyl radicals attack DNA bases and the sugar-phosphate backbone, producing lesions such as single-strand breaks. It has been estimated that mammalian cells experience tens of thousands of such oxidative hits per genome per day, which can miscode during replication if unrepaired (Lindahl and Barnes, 2000; Harman, 1981). Exogenous damage, by contrast, is inflicted by environmental factors including ultraviolet (UV) radiation, genotoxic chemicals, and ionising radiation (Chatterjee and Walker, 2017).

To cope with this continuous onslaught of damage, cells have evolved an integrated DNA damage response (DDR) network. This system comprises lesion sensors that detect DNA abnormalities, signal transducers that amplify and propagate damage signals, and effector pathways that coordinate cell-cycle arrest, DNA repair, or when damage is irreparable, the induction of senescence or apoptosis (Jackson and Bartek, 2009). Within this framework, several partially specialised repair pathways operate in parallel to resolve distinct classes of DNA lesions. Base excision repair (BER) removes small base modifications, including many ROS-induced lesions. Mismatch repair (MMR) corrects replication-associated errors such as base-base mismatches and small insertions, thereby maintaining replication fidelity. Homologous recombination (HR) repairs double-strand breaks using the

sister chromatid as a template, ensuring largely error-free restoration of sequence but restricting its activity to S/G2 phases of the cell cycle. Nucleotide excision repair (NER) excises helix-distorting adducts, including UV-induced photoproducts (Harper and Elledge, 2007; Ciccia and Elledge, 2010). Collectively, these pathways are essential for maintaining genomic stability and, by extension, cellular and tissue viability. Beyond their somatic role, they are also critical for safeguarding the germline by limiting the transmission of deleterious mutations to subsequent generations, thereby shaping both individual ageing trajectories and long-term evolutionary dynamics (Richardson et al., 2004; Jackson and Bartek, 2009).

1.3.2. Protein homeostasis and the ubiquitin–proteasome system

Genomic integrity represents only one dimension of cellular homeostasis. The maintenance of a functional and balanced proteome, termed proteostasis, is equally critical for preventing age-associated cellular dysfunction. Proteostasis is maintained by an interconnected network of molecular chaperones, protein degradation systems and quality-control checkpoints that together regulate protein synthesis, folding, trafficking and turnover. Among these, the ubiquitin–proteasome system (UPS) plays a central role (Balchin et al., 2016). The UPS mediates the selective, ATP-dependent degradation of damaged, misfolded or short-lived regulatory proteins and is indispensable for fundamental cellular processes, including cell cycle progression, signal transduction, and stress adaptation (Koga et al., 2011; Navon and Ciechanover, 2009). In this pathway, target proteins are marked for destruction through the covalent attachment of polyubiquitin chains to lysine residues involving E1 ubiquitin-activating enzymes, E2 ubiquitin-conjugating enzymes, and E3 ubiquitin ligases (Hershko and

Chapter 1 - Biological functions of ageing

Ciechanover, 1998). This process prevents the accumulation of potentially toxic protein species.

With advancing age, proteasome activity declines in many tissues, leading to the accumulation of misfolded and aggregation-prone proteins. Such aggregates are a defining hallmark of numerous age-associated neurodegenerative disorders, including hyperphosphorylated amyloid- β and tau in Alzheimer's disease, and α -synuclein aggregation in Parkinson's disease. By contrast, preserved UPS function is strongly associated with healthy ageing. Fibroblasts derived from long-lived human centenarians show proteasome activity comparable to that observed in cells from much younger individuals, and display increased resistance to proteotoxic stress, suggesting that preserved proteasomal capacity contributes to both extended lifespan and compression of morbidity (Chondrogianni et al., 2000). Together, these observations position proteostasis decline as a key mechanism by which molecular damage accumulates with age and is converted into cellular dysfunction and tissue-level pathology, complementing genomic instability as a core hallmark of the ageing process.

i

1.3.3. Soma–germline dynamics and conserved signalling pathways

Despite the presence of robust DNA repair and proteostasis systems, maintaining molecular integrity is energetically expensive. Evolutionary theory predicts that the allocation of these maintenance resources differs between soma and germline, with patterns shaped by ecological and reproductive context. Experimental evidence from model organisms indicates that germline status can actively influence somatic maintenance and lifespan, highlighting a regulatory, rather than purely resource-limited, interaction between soma and germline. For example, ablation of germline stem cells in *Caenorhabditis elegans* extends lifespan by up to 60%, an effect

Chapter 1 - Biological functions of ageing

mediated by germline-derived steroid hormone signalling that activates somatic FOXO/DAF-16 transcription factors, promoting stress resistance and repair (Hsin and Kenyon, 1999). Similar interventions in *Drosophila melanogaster* enhance survival, and germline ablation in zebrafish (*Danio rerio*) increased somatic DNA repair capacity following genotoxic insult, indicating that reproductive signals can suppress somatic maintenance under normal conditions and that their removal reconfigures gene-regulatory programmes in favour of somatic protection (Partridge et al., 1999; Nöthel, 1987; Chen et al., 2020).

Reflecting this prioritisation, germ cells typically exhibit far lower mutation loads than somatic cells, often an order of magnitude fewer mutations per cell division, owing to intensified DNA repair, more stringent cell-cycle checkpoints and protective niche environments (Milholland et al., 2017). This asymmetry underscores the evolutionary imperative to safeguard heritable genetic information, while somatic maintenance is comparatively less stringently preserved (Milholland et al., 2017; Maklakov and Immler, 2016; Kirkwood and Austad, 2000).

At the signalling level, soma–germline dynamics are coordinated by conserved pathways that integrate reproductive status, environmental cues, and metabolic information to regulate longevity. The insulin/insulin-like signalling (IIS) pathway, for instance, inhibits FOXO transcription factors under nutrient-rich conditions, whereas reduced IIS activity upregulates stress-resistance, autophagy, and repair genes, extending lifespan in worms, flies, and mice (Dillin et al., 2002; Flatt et al., 2008; Kenyon, 2010a). Similarly, the mTOR signalling pathway serves as a central hub integrating amino acid availability, energy status, and growth factor signalling to balance anabolic and catabolic processes. Downregulation of mTOR, as occurs under dietary restriction or pharmacological inhibition, consistently promotes longevity

Chapter 1 - Epigenetic clocks as biomarkers of ageing

across taxa, highlighting the benefits of moderating growth and metabolic activity for long-term somatic maintenance (Kenyon, 2010a; Murphy, 2018).

Collectively, these observations show that the biological foundations of ageing arise from a dynamic interplay between persistent molecular damage, adaptive repair, quality-control systems, and regulatory networks that link reproductive status, nutrient availability and somatic maintenance. Rather than reflecting fixed trade-offs, ageing emerges from dynamic, environmentally sensitive soma–germline signalling, evolved to optimise reproductive success while shaping both lifespan and healthspan across species.

1.4. Epigenetic clocks as biomarkers of ageing

The biological foundations of ageing extend beyond the accumulation of molecular damage to include dynamic regulation of gene expression and chromatin architecture. Epigenetic modifications integrate developmental history, environmental exposures, and physiological state, providing a molecular interface between genome and environment. Among the most precise measures of biological age are epigenetic clocks, machine learning models trained on molecular markers, most commonly DNAm, to estimate biological age as a function of chronological time (Horvath, 2013, Bell et al., 2019, Zheng et al., 2024). These clocks achieve high predictive accuracy by identifying age-correlated methylation changes at thousands of cytosine–phosphate–guanine (CpG) sites across the genome, offering a robust readout of cumulative physiological wear.

Chapter 1 - Epigenetic clocks as biomarkers of ageing

1.4.1. Development and mechanistic basis of DNA methylation clocks

Epigenetic clocks are primarily based on DNAm, a chemical modification where methyl groups are added to cytosine bases, typically regulating gene activity and chromatin architecture. Age-associated methylation patterns follow both hypermethylation (gain of methyl groups) and hypomethylation (loss of methyl groups), which can be statistically modelled to predict age with high precision (Horvath and Raj, 2018, De Lima Camillo et al., 2023, Morandini et al., 2023).

Horvath's seminal 2013 pan-tissue clock, trained on 353 CpGs across 51 tissues and cell types (including blood, brain, liver and fibroblasts), achieved a median absolute error of 3.6 years, demonstrating applicability from foetal stages to centenarians (Horvath, 2013). Hannum's blood-specific clock, using 71 CpGs optimised for peripheral blood leucocytes, is particularly sensitive to inflammatory states and shifts in immune cell composition (Hannum et al., 2013). More recent 'second-generation' clocks integrate DNAm with clinical phenotypes. Levine's PhenoAge (Levine et al., 2018), combines 199 CpGs with nine plasma biomarkers (including lymphocyte percentage, mean cell volume and red cell distribution width) to predict phenotypic age and all-cause mortality (hazard ratio 1.6 per 5-year excess biological age); while Lu's GrimAge model (Lu et al., 2019), refined as GrimAge2, leverages 1,030 CpGs as proxies for seven plasma proteins (e.g., plasminogen activator inhibitor-1), outperforming prior models in lifespan prediction.

Collectively, these approaches surpass other predictors such as telomere length in prognostic utility, which suffers from limited tissue generality, high measurement variance, and weak correlation with morbidity (Levine et al., 2018).

Chapter 1 - Epigenetic clocks as biomarkers of ageing

1.4.2. Mechanisms of age-related methylation change and clock interpretation

Initially, age-associated DNAm changes were hypothesised to directly drive senescence by silencing genes involved in DNA repair or stress responses (Wilson and Jones, 1983). However, integrative multi-omics studies now indicate that DNAm shifts largely reflect downstream responses or adaptive reprogramming to primary ageing insults, such as inflammation, metabolic dysregulation or oxidative stress. (Hernandez et al., 2011, Ying et al., 2024, Xiao et al., 2019, Lu et al., 2023, Jung and Pfeifer, 2015). Critically, epigenetic clocks also exhibit differential sensitivity across ageing hallmarks, aligning more closely with nutrient-sensing and mitochondrial function (Kabacik et al., 2022), stem cell composition (Van Der Rijt et al., 2020) and intercellular communication (López-Otín et al., 2013) than genomic instability or classical senescence.

Clock accuracy is influenced by genetic variation (He et al., 2018), cell differentiation states (Kabacik et al., 2022), and environmental exposures such as diet or irradiation (Marsit, 2015). Moreover, local mutation rates, particularly at CpG dinucleotide prone to C → T transitions, contribute to epigenetic drift and tissue-specific ageing signatures (Zhou et al., 2020, Koch et al., 2023).

Accordingly, tissues show variable age-related DNAm changes, with regions of the brain such as the cerebellum exhibit comparatively slower epigenetic ageing relative to other tissues (Spiers et al., 2016, Guevara et al., 2022).

1.4.3. Applications of epigenetic clocks

Epigenetic clocks have become critical tools in ageing research, serving as quantitative endpoints in epidemiology, intervention trials, and mechanistic

Chapter 1 - Cellular senescence and therapeutic approaches

investigations. For example, GrimAge acceleration predicts 10-year mortality in large cohorts such as the UK Biobank (Lu et al., 2019), while PhenoAge2 has been applied to pharmacological trials examining metformin's potential anti-ageing effects (Higgins-Chen et al., 2022). Clocks also allow assessment of lifestyle and environmental effects. For example, smoking one pack of cigarettes advances GrimAge by ~0.7 years (Lu et al., 2019), whereas exposure to air pollution alters methylation at inflammatory loci, and caloric restriction can slow pan-tissue clocks by approximately 1% annually (Meer et al., 2018).

Despite their predictive power, epigenetic clocks vary in generalisability. Tissue or phenotype-specific models often outperform pan-tissue clocks, but lack their broad adaptability (Shireby et al., 2020). Furthermore, DNAm changes captured by clocks include both functionally relevant and neutral modifications, complicating mechanistic interpretation (Porter et al., 2021). Beyond serving as molecular biomarkers, epigenetic clocks also offer insight into evolutionary processes. Age-associated DNAm patterns likely reflect interactions between somatic maintenance and reproduction shaped by natural selection. Integrating epigenetic ageing with an evolutionary framework can thus illuminate how molecular changes relate to organismal fitness strategies and life-history optimisation.

1.5. Cellular senescence and therapeutic approaches

1.5.1. Mechanisms of cellular senescence and the senescence-associated secretory phenotype (SASP)

Cellular senescence represents a key downstream consequence of the molecular changes captured by epigenetic clocks, particularly those related to DNA damage responses, chromatin remodelling and altered intercellular communication (López-Gil et al., 2023, Herranz and Gil, 2018). It is a stable and

highly regulated physiological state induced by diverse cellular insults, including genotoxic stress, telomere attrition, oxidative damage and oncogenic activation (López-Gil et al., 2023, Herranz and Gil, 2018). Senescence is characterised by irreversible cell-cycle arrest, mediated predominantly through upregulation of cyclin-dependent kinase inhibitors via *p53*-dependent pathways (Herranz and Gil, 2018). This programme acts as a potent tumour-suppressive mechanism in early life by preventing the proliferation of damaged or pre-malignant cells (Herranz and Gil, 2018). However, with advancing age, the progressive accumulation of senescent cells, estimated to reach approximately 1–15% of cells in certain human tissues by around 80 years of age, contributes directly to functional decline and a broad spectrum of age-associated pathologies (López-Gil et al., 2023, Kanad Ghosh and Brian, 2016). Senescence, therefore, exemplifies antagonistic pleiotropy, providing early-life benefits while exerting deleterious effects late in life (Herranz and Gil, 2018, López-Gil et al., 2023).

Beyond growth arrest, senescent cells adopt a distinctive SASP profile, comprising pro-inflammatory cytokines, chemokines, growth factors, matrix-remodelling proteases and extracellular matrix components which collectively alter tissue microenvironments (Kanad Ghosh and Brian, 2016, Coppé et al., 2010). Transient SASP activity can be beneficial, supporting embryonic tissue patterning, wound repair, and reinforcing tumour suppression (Coppé et al., 2010, Soto-Gamez and Demaria, 2017, Birch and Gil, 2020). However, chronic SASP signalling during ageing promotes low-grade inflammation, pathological tissue remodelling, and disruption of tissue architecture contributing to diseases including osteoarthritis, atherosclerosis and neurodegeneration (Coppé et al., 2010, Soto-Gamez and Demaria, 2017, Huang et al., 2022).

SASP expression is governed by multi-layered signalling networks integrating persistent DDR signalling, inflammatory transcription factors,

and stress-responsive kinases (Herranz and Gil, 2018, Freund et al., 2011). DDR maintains p53 activity to initiate SASP, while NF- κ B acts as the central transcriptional regulator of SASP genes (Herranz and Gil, 2018, Freund et al., 2011). Importantly, SASP composition is highly context dependent, varying with the initiating insult, tissue and cell type, and duration of senescence, which introduces substantial heterogeneity even within a single tissue (Freund et al., 2011, Salminen et al., 2012). This variability complicates mechanistic dissection and presents challenges for targeted senotherapeutic strategies.

1.5.2. Consequences of senescent cell accumulation in tissues and immune surveillance

While transient senescence supports development, tissue repair, and tumour suppression, chronic accumulation of senescent cells drives tissue dysfunction and age-related disorders. Experimental ablation of senescent cells in progeroid and naturally aged mice delays the onset of frailty, improves tissue function, and reduces cancer incidence, confirming the causal role of senescent cells in systemic ageing (Childs et al., 2015, Kanad Ghosh and Brian, 2016).

Senescent cells resist apoptosis, disrupt tissue organisation, promote fibrosis via extracellular matrix deposition, and perpetuate inflammation via persistent SASP signalling (Kanad Ghosh and Brian, 2016). In addition, senescence can propagate to neighbouring non-senescent cells through paracrine SASP factors, a phenomenon termed the “senescence-associated bystander effect”, amplifying tissue dysfunction (Soto-Gamez and Demaria, 2017, Herranz and Gil, 2018).

Chapter 1 - Cellular senescence and therapeutic approaches

Under normal conditions, senescent cells are removed by immunosurveillance pathways mediated by natural killer (NK) cells, cytotoxic T-cells and macrophages, which detect stress-induced ligands and altered self-antigens to induce apoptosis (Herranz and Gil, 2018, Wang et al., 2024a, Wang et al., 2024b). Ageing impairs these processes through immunosenescence and inflammageing, characterised by diminished NK cytotoxicity, T-cell exhaustion and SASP-driven expansion of immunosuppressive myeloid populations, collectively impairing senescent cell clearance (Wang et al., 2024a, Wang et al., 2024b, Huang et al., 2022). As a result, senescent cell burden increases in older individuals, accelerating functional decline and compromising tissue regenerative capacity (Wang et al., 2024a, Wang et al., 2024b).

1.5.3. Therapeutic strategies: senolytics, senomorphics, and the central role of rapamycin

Recognition of senescent cells as drivers of ageing has motivated the development of senotherapeutic strategies, which can be grouped into two categories. Senolytics are agents that selectively induce apoptosis in senescent cells by targeting their dependence on pro-survival pathways, thereby reducing senescent cell burden and ameliorating pathology in preclinical models (Childs et al., 2015, Kanad Ghosh and Brian, 2016). The combination of dasatinib and quercetin (D+Q) is currently under early-phase clinical investigation for age-related conditions (Pulakat and Chen, 2020, Justice et al., 2019).

Senomorphics (or senostatics) modulate deleterious components of the SASP without eliminating senescent cells, reducing chronic inflammation and paracrine damage (Soto-Gamez and Demaria, 2017, Huang et al., 2022). Representative senomorphics include rapamycin and metformin, which

Chapter 1 - Cellular senescence and therapeutic approaches

modulate metabolic and inflammatory pathways to suppress SASP output (Soto-Gamez and Demaria, 2017, Saliev and Singh, 2025). Early clinical studies of rapamycin in older adults report improved vaccine responses, physical performance and systemic inflammatory profiles (Mannick et al., 2014).

Central to many senomorphic interventions is mTOR, a conserved kinase integrating nutrient, energy and growth factor signals to regulate cell growth, metabolism and autophagy (Saxton and Sabatini, 2017). mTOR functions in two complexes, mTORC1, promoting anabolic processes and inhibiting autophagy, and mTORC2, regulating cytoskeletal organisation and insulin/AKT signalling (Saxton and Sabatini, 2017, Weichhart, 2018). Chronic mTORC1 hyperactivation in adulthood drives hallmarks of ageing, including impaired proteostasis, deregulated nutrient sensing, mitochondrial dysfunction, and cancer susceptibility, representing another example of antagonistic pleiotropy at the signalling level (Blagosklonny, 2010, Weichhart, 2018).

Rapamycin, a macrolide derived from *Streptomyces hygroscopicus*, selectively inhibits mTORC1, attenuating SASP output by limiting NF- κ B-driven transcription, enhancing autophagy and mitophagy, improving mitochondrial quality control, and supporting proteostasis (Herranz et al., 2015, Laberge et al., 2015). Preclinical studies show rapamycin extends lifespan and healthspan across yeast, nematodes, flies, and mice, even when administered in mid-to-late life, improving survival, immune function, metabolic regulation, and cognitive performance (Harrison et al., 2009, Miller et al., 2011, Kaeberlein, 2014, Saxton and Sabatini, 2017).

Clinical use of rapamycin is, however, constrained by dose-dependent adverse effects. Doses of rapamycin that cause immunosuppressive effects, as used in tissue transplantation, impair wound healing and increase infection risk, while chronic exposure can induce glucose intolerance and

dyslipidaemia, likely through partial mTORC2 inhibition (Saxton and Sabatini, 2017, Konopka et al., 2023). These concerns have prompted exploration of low-dose or intermittent regimens, which appear to retain many of the beneficial effects on immune function and physical performance while mitigating toxicity. However, long-term data in healthy older populations remain limited (Mannick et al., 2014).

1.6. Zebrafish as a model for ageing research

Zebrafish (*Danio rerio*) have emerged as a versatile vertebrate model for ageing combining many experimental advantages of invertebrate systems with key features of mammalian biology, including gradual, multi-system senescence and conserved genetic pathways. They are small, inexpensive to maintain, highly fecund, and amenable to large-scale genetic and pharmacological screens, making them particularly suitable for mechanistic studies of ageing and the evaluation of interventions at scale (Gerhard, 2003, Kishi et al., 2009, Gilbert et al., 2014). Unlike short-lived invertebrates, zebrafish exhibit vertebrate-typical tissues and tissue systems, such as a complex brain and adaptive immune system, allowing direct study of age-related decline in structures and functions that are closely analogous to those affected in humans (Gerhard, 2003, Kishi et al., 2009).

A major strength of zebrafish lies in their external development and optical transparency during embryonic and larval stages, enabling in vivo imaging of cellular and molecular processes, including DNA damage responses, stem-cell dynamics and senescence-associated markers, within intact organisms over time (Kishi et al., 2009, Adams and Kafaligonul, 2018). Their relatively short lifespan, approximately two years under laboratory conditions, facilitates longitudinal analysis of gradual ageing trajectories,

Chapter 1 - Zebrafish as a model for ageing research

including progressive changes in musculoskeletal morphology, fin regeneration and endocrine function, and circadian rhythm (Kishi et al., 2009, Gilbert et al., 2014). Ageing zebrafish also display many classical mammalian ageing phenotypes, such as spinal curvature, accumulation of senescence-associated proteins, reduced regenerative capacity, and declining reproductive output, underscoring the conservation of core ageing processes (Kishi et al., 2009, Adams and Kafaligonul, 2018).

Beyond these general advantages, zebrafish are uniquely suited to experimental tests of evolutionary theories of ageing, particularly those concerning soma–germline interactions. Their amenability to genetic manipulation enables precise dissection of germline–soma dynamics *in vivo*, providing vertebrate insights into physiological trade-offs central to models such as the expensive germline hypothesis (Chen et al., 2020). A key resource for such work, and the primary germline ablation model employed in this thesis, is the dead end 1 (*dnd1*) mutant line, generated through targeted disruption of the RNA-binding protein *dnd1*, which is essential for primordial germ cell (PGC) specification, survival and migration during early embryogenesis (Slanchev et al., 2005, Li et al., 2017, Chen et al., 2020). In this strain, *dnd1* loss-of-function prevents PGC formation beyond the first few cell divisions, yielding viable, germline-free adults that develop normally through somatic lineages. This model therefore enables controlled investigation of soma-only ageing trajectories, providing critical insights into the interplay between reproductive status and somatic maintenance (Slanchev et al., 2005, Li et al., 2017, Chen et al., 2020).

1.7. Integration and future directions

The multidisciplinary study of ageing has advanced through the integration of complementary perspectives spanning evolutionary theory, molecular biomarkers and therapeutic interventions, creating a cohesive narrative from ultimate causation to proximate mechanisms. Evolutionary frameworks, including mutation accumulation, antagonistic pleiotropy, disposable soma theory, developmental theory of ageing, and the expensive germline hypothesis, collectively explain senescence as an adaptive consequence of declining late-life selection pressures and germline-soma conflicts shaped by extrinsic mortality regimes. At the molecular level, epigenetic clocks provide precise, quantifiable tools to estimate biological age and track the progression of ageing, linking genotypic influences, environmental exposures, and physiological decline, thereby translating abstract evolutionary predictions into measurable epigenomic phenomena. Therapeutically, rapamycin and related mTOR inhibitors provide proof-of-concept that targeting evolutionarily conserved ageing pathways can yield tangible benefits, suppressing SASP-driven inflammation, restoring autophagy, mitophagy, and proteostasis, and extending healthspan and lifespan across diverse model organisms, including yeast (*Saccharomyces cerevisiae*), worms (*Caenorhabditis elegans*), flies (*Drosophila melanogaster*) and mice (HET3).

Despite these converging advances, critical knowledge gaps remain. Foremost among them is elucidating the precise molecular mechanisms of germline-soma interactions and their direct influence on lifespan, biological age and healthspan. Furthermore, the causal versus correlative relationship between age-associated DNAm changes and phenotypic ageing requires mechanistic clarification to refine biomarkers, optimise predictive models, and identify optimal windows for intervention.

This thesis addresses these gaps by employing the *dnd1* germline-ablated zebrafish model to dissect germline-soma signalling dynamics under controlled genotoxic stress, integrating phenotypic, epigenetic clock, and transcriptomic analyses to establish causal links between germline status, DNA repair capacity and biological ageing trajectories. Through this vertebrate system, the work tests core predictions of unified evolutionary frameworks, quantifies the plasticity of physiological resource allocation, and evaluates the translational potential of mTOR modulation for healthspan extension. In doing so, it bridges molecular mechanisms with evolutionary theory, illuminating actionable pathways to mitigate age-related decline.

1.8. References

ADAMS, M. M. & KAFALIGONUL, H. 2018. Zebrafish-A Model Organism for Studying the Neurobiological Mechanisms Underlying Cognitive Brain Aging and Use of Potential Interventions. *Front Cell Dev Biol*, 6, 135.

AUSTAD, S. N. 1993. Retarded senescence in an insular population of Virginia opossums (*Didelphis virginiana*). *Journal of Zoology*, 229, 695-708.

BALCHIN, D., HAYER-HARTL, M. & HARTL, F. U. 2016. In vivo aspects of protein folding and quality control. *Science*, 353, aac4354.

BELL, C. G., LOWE, R., ADAMS, P. D., BACCARELLI, A. A., BECK, S., BELL, J. T., CHRISTENSEN, B. C., GLADYSHEV, V. N., HEIJMANS, B. T., HORVATH, S., IDEKER, T., ISSA, J.-P. J., KELSEY, K. T., MARIONI, R. E., REIK, W., RELTON, C. L., SCHALKWYK, L. C., TESCHENDORFF, A. E., WAGNER, W., ZHANG, K. & RAKYAN, V. K. 2019. DNA methylation aging clocks: challenges and recommendations. *Genome Biology*, 20.

Chapter 1 - References

BIRCH, J. & GIL, J. 2020. Senescence and the SASP: many therapeutic avenues. *Genes & Development*, 34, 1565-1576.

BLAGOSKLONNY, M. V. 2010. Why men age faster but reproduce longer than women: mTOR and evolutionary perspectives. *Aging*, 2, 265-273.

BLAS, J., SERGIO, F. & HIRALDO, F. 2009. Age-related improvement in reproductive performance in a long-lived raptor: a cross-sectional and longitudinal study. *Ecography*, 32, 647-657.

CHEN, H.-Y., JOLLY, C., BUBLYS, K., MARCU, D. & IMMLER, S. 2020. Trade-off between somatic and germline repair in a vertebrate supports the expensive germ line hypothesis. *Proceedings of the National Academy of Sciences*, 117, 8973-8979.

CHILDS, B. G., DURIK, M., BAKER, D. J. & VAN DEURSEN, J. M. 2015. Cellular senescence in aging and age-related disease: from mechanisms to therapy. *Nature Medicine*, 21, 1424-1435.

COPPÉ, J.-P., DESPREZ, P.-Y., KRTOLICA, A. & CAMPISI, J. 2010. The Senescence-Associated Secretory Phenotype: The Dark Side of Tumor Suppression. *Annual Review of Pathology: Mechanisms of Disease*, 5, 99-118.

DE LIMA CAMILLO, L. P., ASIF, M. H., HORVATH, S., LARSCHAN, E. & SINGH, R. 2023. Histone mark age of human tissues and cells. Cold Spring Harbor Laboratory.

DI FRANCESCO, A., DEIGHAN, A. G., LITICHEVSKIY, L., CHEN, Z., LUCIANO, A., ROBINSON, L., GARLAND, G., DONATO, H., VINCENT, M., SCHOTT, W., WRIGHT, K. M., RAJ, A., PRATEEK, G. V., MULLIS, M., HILL, W. G., ZEIDEL, M. L., PETERS, L. L., HARDING, F., BOTSTEIN, D., KORSTANJE, R., THAISS, C. A., FREUND, A. & CHURCHILL, G. A. 2024. Dietary restriction impacts health and lifespan of genetically diverse mice. *Nature*, 634, 684-692.

Chapter 1 - References

FREUND, A., PATIL, C. K. & CAMPISI, J. 2011. p38MAPK is a novel DNA damage response-independent regulator of the senescence-associated secretory phenotype. *The EMBO Journal*, 30, 1536-1548.

GERHARD, G. S. 2003. Comparative aspects of zebrafish (*Danio rerio*) as a model for aging research. *Exp Gerontol*, 38, 1333-41.

GILBERT, M. J., ZERULLA, T. C. & TIERNEY, K. B. 2014. Zebrafish (*Danio rerio*) as a model for the study of aging and exercise: physical ability and trainability decrease with age. *Exp Gerontol*, 50, 106-13.

GLADYSHEV, V. N., KRITCHEVSKY, S. B., CLARKE, S. G., CUERVO, A. M., FIEHN, O., DE MAGALHÃES, J. P., MAU, T., MAES, M., MORITZ, R. L., NIEDERNHOFER, L. J., VAN SCHAFTINGEN, E., TRANAH, G. J., WALSH, K., YURA, Y., ZHANG, B. & CUMMINGS, S. R. 2021. Molecular damage in aging. *Nature Aging*, 1, 1096-1106.

GUEVARA, E. E., HOPKINS, W. D., HOF, P. R., ELY, J. J., BRADLEY, B. J. & SHERWOOD, C. C. 2022. Epigenetic ageing of the prefrontal cortex and cerebellum in humans and chimpanzees. *Epigenetics*, 17, 1774-1785.

HALDANE, J. B. S. 1941. New paths in genetics. *Allen & Unwin*.

HANNUM, G., GUINNEY, J., ZHAO, L., ZHANG, L., HUGHES, G., SADDA, S., KLOTZLE, B., BIBIKOVA, M., FAN, J.-B., GAO, Y., DECONDE, R., CHEN, M., RAJAPAKSE, I., FRIEND, S., IDEKER, T. & ZHANG, K. 2013. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell*, 49, 359-367.

HARMAN, D. 1981. The aging process. *Proceedings of the National Academy of Sciences*, 78, 7124-7128.

HARRISON, D. E., STRONG, R., SHARP, Z. D., NELSON, J. F., ASTLE, C. M., FLURKEY, K., NADON, N. L., WILKINSON, J. E., FRENKEL, K., CARTER, C. S., PAHOR, M., JAVORS, M. A., FERNANDEZ, E. & MILLER, R. A. 2009. Rapamycin

Chapter 1 - References

fed late in life extends lifespan in genetically heterogeneous mice. *Nature*, 460, 392-395.

HE, Z., ZHANG, R., JIANG, F., HOU, W. & HU, C. 2018. Role of genetic and environmental factors in DNA methylation of lipid metabolism. *Genes Dis*, 5, 9-15.

HERNANDEZ, D. G., NALLS, M. A., GIBBS, J. R., AREPALLI, S., VAN DER BRUG, M., CHONG, S., MOORE, M., LONGO, D. L., COOKSON, M. R., TRAYNOR, B. J. & SINGLETON, A. B. 2011. Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Human Molecular Genetics*, 20, 1164-1172.

HERRANZ, N., GALLAGE, S., MELLONE, M., WUESTEFELD, T., KLOTZ, S., HANLEY, C. J., RAGUZ, S., ACOSTA, J. C., INNES, J., ANDREW, BANITO, A., GEORGILIS, A., MONTOYA, A., WOLTER, K., DHARMALINGAM, G., FAULL, P., CARROLL, T., MARTÍNEZ-BARBERA, J. P., CUTILLAS, P., REISINGER, F., HEIKENWALDER, M., MILLER, A., RICHARD, WITHERS, D., ZENDER, L., THOMAS, G. J. & GIL, J. 2015. mTOR regulates MAPKAPK2 translation to control the senescence-associated secretory phenotype. *Nature Cell Biology*, 17, 1205-1217.

HERRANZ, N. & GIL, J. 2018. Mechanisms and functions of cellular senescence. *J Clin Invest*, 128, 1238-1246.

HERSHKO, A. & CIECHANOVER, A. 1998. The ubiquitin system. *Annu Rev Biochem*, 67, 425-79.

HIGGINS-CHEN, A. T., THRUSH, K. L., WANG, Y., MINTEER, C. J., KUO, P.-L., WANG, M., NIIMI, P., STURM, G., LIN, J., MOORE, A. Z., BANDINELLI, S., VINKERS, C. H., VERMETTEN, E., RUTTEN, B. P. F., GEUZE, E., OKHUIJSEN-PFEIFER, C., VAN DER HORST, M. Z., SCHREITER, S., GUTWINSKI, S., LUYKX, J. J., PICARD, M., FERRUCCI, L., CRIMMINS, E. M., BOKS, M. P., HÄGG, S., HU-

Chapter 1 - References

SELIGER, T. T. & LEVINE, M. E. 2022. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nature Aging*, 2, 644-661.

HORVATH, S. 2013. DNA methylation age of human tissues and cell types. *Genome Biology*, 14, R115.

HORVATH, S. & RAJ, K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*, 19, 371-384.

HUANG, W., HICKSON, L. J., EIRIN, A., KIRKLAND, J. L. & LERMAN, L. O. 2022. Cellular senescence: the good, the bad and the unknown. *Nat Rev Nephrol*, 18, 611-627.

HUGHES, K. A. 2010. Mutation and the evolution of ageing: from biometrics to system genetics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 1273-1279.

IVIMEY-COOK, E. & MOORAD, J. 2020. The diversity of maternal-age effects upon pre-adult survival across animal species. *Proceedings of the Royal Society B: Biological Sciences*, 287, 20200972.

JACKSON, S. P. & BARTEK, J. 2009. The DNA-damage response in human biology and disease. *Nature*, 461, 1071-1078.

JINESH, S., ÖZÜPEK, B. & ADITI, P. 2025. Premature aging and metabolic diseases: the impact of telomere attrition. *Frontiers in Aging*, 6.

JUNG, M. & PFEIFER, G. P. 2015. Aging and DNA methylation. *BMC Biology*, 13, 7.

JUSTICE, J. N., NAMBIAR, A. M., TCHKONIA, T., LEBRASSEUR, N. K., PASCUAL, R., HASHMI, S. K., PRATA, L., MASTERNAK, M. M., KRITCHEVSKY, S. B., MUSI, N.

Chapter 1 - References

- & KIRKLAND, J. L. 2019. Senolytics in idiopathic pulmonary fibrosis: Results from a first-in-human, open-label, pilot study. *EBioMedicine*, 40, 554-563.
- KABACIK, S., LOWE, D., FRANSEN, L., LEONARD, M., ANG, S.-L., WHITEMAN, C., CORSI, S., COHEN, H., FELTON, S., BALI, R., HORVATH, S. & RAJ, K. 2022. The relationship between epigenetic age and the hallmarks of aging in human cells. *Nature Aging*, 2, 484-493.
- KAEBERLEIN, M. 2014. Rapamycin and ageing: when, for how long, and how much? *J Genet Genomics*, 41, 459-63.
- KANAD GHOSH AND BRIAN, C. C. 2016. The Senescence-Associated Secretory Phenotype: Critical Effector in Skin Cancer and Aging. *Journal of Investigative Dermatology*, 136, 2133-2139.
- KENNEDY, B. K., BERGER, S. L., BRUNET, A., CAMPISI, J., CUERVO, A. M., EPEL, E. S., FRANCESCHI, C., LITHGOW, G. J., MORIMOTO, R. I., PESSIN, J. E., RANDO, T. A., RICHARDSON, A., SCHADT, E. E., WYSS-CORAY, T. & SIERRA, F. 2014. Geroscience: Linking Aging to Chronic Disease. *Cell*, 159, 709-713.
- KENYON, C. J. 2010. The genetics of ageing. *Nature*, 464, 504-512.
- KIRKWOOD, T. B. & MELOV, S. 2011. On the programmed/non-programmed nature of ageing within the life history. *Curr Biol*, 21, R701-7.
- KIRKWOOD, T. B. L. 1977. Evolution of ageing. *Nature*, 270, 301-304.
- KIRKWOOD, T. B. L. 2017. The Disposable Soma Theory: Origins and Evolution. *Cambridge University Press, Cambridge*, The evolution of senescence in the tree of life.
- KISHI, S., SLACK, B. E., UCHIYAMA, J. & ZHDANOVA, I. V. 2009. Zebrafish as a genetic model in biological and behavioral gerontology: where development meets aging in vertebrates--a mini-review. *Gerontology*, 55, 430-41.

Chapter 1 - References

KOCH, Z., LI, A., EVANS, D. S., CUMMINGS, S. & IDEKER, T. 2023. Somatic mutation as an explanation for epigenetic aging. Cold Spring Harbor Laboratory.

KONOPKA, A. R., LAMMING, D. W., GRASSO, B. A., MARRAH, R. C., DECKER, S., GARG, N., PARK, Y., LIM, S. Y., SIMCOX, J. A., GREEN, C. L., GRUNOW, I., YADEV, A. K., MARRAH, R. C., GRASSO, B. A., DECKER, S., PABICH, S., MANDELROT, D., WALLHAUS, T. R., WIEBEN, O., OSMAN, F., CHAPPELL, R. J., ONG, I. M., ALISCH, R. S., SIMCOX, J. A., ELLIEHAUSEN, C. J., MINTON, D. M., TRAUTMAN, M. E., SPAHIC, A., BENDLIN, B. B., ASTHANA, S. & TERSKIKH, A. 2023. Blazing a trail for the clinical use of rapamycin as a geroprotecTOR. *GeroScience*, 45, 2769-2783.

LABERGE, R.-M., SUN, Y., ORJALO, A. V., PATIL, C. K., FREUND, A., ZHOU, L., CURRAN, C., SAMUEL, DAVALOS, A. R., WILSON-EDELL, K. A., LIU, S., LIMBAD, C., DEMARIA, M., LI, P., HUBBARD, G. B., IKENO, Y., JAVORS, M., DESPREZ, P.-Y., BENZ, C. C., KAPAHI, P., NELSON, P. S. & CAMPISI, J. 2015. MTOR regulates the pro-tumorigenic senescence-associated secretory phenotype by promoting IL1A translation. *Nature Cell Biology*, 17, 1049-1061.

LEE, H.-W., BLASCO, M. A., GOTTLIEB, G. J., HORNER, J. W., GREIDER, C. W. & DEPINHO, R. A. 1998. Essential role of mouse telomerase in highly proliferative organs. *Nature*, 392, 569-574.

LEMAÎTRE, J.-F., BERGER, V., BONENFANT, C., DOUHARD, M., GAMELON, M., PLARD, F. & GAILLARD, J.-M. 2015. Early-late life trade-offs and the evolution of ageing in the wild. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20150209.

LEMAÎTRE, J.-F., MOORAD, J., GAILLARD, J.-M., MAKLAKOV, A. A. & NUSSEY, D. H. 2024. A unified framework for evolutionary genetic and physiological theories of aging. *PLOS Biology*, 22, e3002513.

Chapter 1 - References

LEVINE, M. E., LU, A. T., QUACH, A., CHEN, B. H., ASSIMES, T. L., BANDINELLI, S., HOU, L., BACCARELLI, A. A., STEWART, J. D., LI, Y., WHITSEL, E. A., WILSON, J. G., REINER, A. P., AVIV, A., LOHMAN, K., LIU, Y., FERRUCCI, L. & HORVATH, S. 2018. An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, 10, 573-591.

LI, Q., FUJII, W., NAITO, K. & YOSHIZAKI, G. 2017. Application of *dead end*-knockout zebrafish as recipients of germ cell transplantation. *Molecular Reproduction and Development*, 84, 1100-1111.

LÓPEZ-GIL, L., PASCUAL-AHUIR, A. & PROFT, M. 2023. Genomic Instability and Epigenetic Changes during Aging. *International Journal of Molecular Sciences*, 24, 14279.

LÓPEZ-OTÍN, C., BLASCO, M. A., PARTRIDGE, L., SERRANO, M. & KROEMER, G. 2013. The hallmarks of aging. *Cell*, 153, 1194-217.

LU, A. T., FEI, Z., HAGHANI, A., ROBECK, T. R., ZOLLER, J. A., LI, C. Z., LOWE, R., YAN, Q., ZHANG, J., VU, H., ABLAEVA, J., ACOSTA-RODRIGUEZ, V. A., ADAMS, D. M., ALMUNIA, J., ALOYSIUS, A., ARDEHALI, R., ARNESON, A., BAKER, C. S., BANKS, G., BELOV, K., BENNETT, N. C., BLACK, P., BLUMSTEIN, D. T., BORS, E. K., BREEZE, C. E., BROOKE, R. T., BROWN, J. L., CARTER, G. G., CAULTON, A., CAVIN, J. M., CHAKRABARTI, L., CHATZISTAMOU, I., CHEN, H., CHENG, K., CHIAVELLINI, P., CHOI, O. W., CLARKE, S. M., COOPER, L. N., COSSETTE, M. L., DAY, J., DEYOUNG, J., DIROCCO, S., DOLD, C., EHMKE, E. E., EMMONS, C. K., EMMRICH, S., ERBAY, E., ERLACHER-REID, C., FAULKES, C. G., FERGUSON, S. H., FINNO, C. J., FLOWER, J. E., GAILLARD, J. M., GARDE, E., GERBER, L., GLADYSHEV, V. N., GORBUNOVA, V., GOYA, R. G., GRANT, M. J., GREEN, C. B., HALES, E. N., HANSON, M. B., HART, D. W., HAULENA, M., HERRICK, K., HOGAN, A. N., HOGG, C. J., HORE, T. A., HUANG, T., IZPISUA BELMONTE, J. C., JASINSKA, A. J., JONES, G., JOURDAIN, E., KASHPUR, O., KATCHER, H., KATSUMATA, E., KAZA, V., KIARIS, H., KOBOR, M. S., KORDOWITZKI, P., KOSKI, W. R., KRÜTZEN,

Chapter 1 - References

- M., KWON, S. B., LARISON, B., LEE, S. G., LEHMANN, M., LEMAITRE, J. F., LEVINE, A. J., LI, C., LI, X., LIM, A. R., LIN, D. T. S., LINDEMANN, D. M., LITTLE, T. J., MACORETTA, N., MADDOX, D., MATKIN, C. O., MATTISON, J. A., MCCLURE, M., MERGL, J., et al. 2023. Universal DNA methylation age across mammalian tissues. *Nature Aging*, 3, 1144-1166.
- LU, A. T., QUACH, A., WILSON, J. G., REINER, A. P., AVIV, A., RAJ, K., HOU, L., BACCARELLI, A. A., LI, Y., STEWART, J. D., WHITSEL, E. A., ASSIMES, T. L., FERRUCCI, L. & HORVATH, S. 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, 11, 303-327.
- MAKLAKOV, A., ALEXEI & IMMLER, S. 2016. The Expensive Germline and the Evolution of Ageing. *Current Biology*, 26, R577-R586.
- MAKLAKOV, A. A. & CHAPMAN, T. 2019. Evolution of ageing as a tangle of trade-offs: energy versus function. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20191604.
- MANNICK, J. B., DEL GIUDICE, G., LATTANZI, M., VALIANTE, N. M., PRAESTGAARD, J., HUANG, B., LONETTO, M. A., MAECKER, H. T., KOVARIK, J., CARSON, S., GLASS, D. J. & KLINKSTEIN, L. B. 2014. mTOR inhibition improves immune function in the elderly. *Sci Transl Med*, 6, 268ra179.
- MARSIT, C. J. 2015. Influence of environmental exposure on human epigenetic regulation. *Journal of Experimental Biology*, 218, 71-79.
- MCHUGH, D. & GIL, J. 2018. Senescence and aging: Causes, consequences, and therapeutic avenues. *Journal of Cell Biology*, 217, 65-77.
- MEDAWAR, P. B. 1952 An Unsolved Problem of Biology. *H. K. Lewis, London*.
- MEER, M. V., PODOLSKIY, D. I., TYSHKOVSKIY, A. & GLADYSHEV, V. N. 2018. A whole lifespan mouse multi-tissue DNA methylation clock. *Elife*, 7.

Chapter 1 - References

- MILHOLLAND, B., DONG, X., ZHANG, L., HAO, X., SUH, Y. & VIJG, J. 2017. Differences between germline and somatic mutation rates in humans and mice. *Nature Communications*, 8, 15183.
- MILLER, R. A., HARRISON, D. E., ASTLE, C. M., BAUR, J. A., BOYD, A. R., DE CABO, R., FERNANDEZ, E., FLURKEY, K., JAVORS, M. A., NELSON, J. F., ORIHUELA, C. J., PLETCHER, S., SHARP, Z. D., SINCLAIR, D., STARNES, J. W., WILKINSON, J. E., NADON, N. L. & STRONG, R. 2011. Rapamycin, But Not Resveratrol or Simvastatin, Extends Life Span of Genetically Heterogeneous Mice. *The Journals of Gerontology: Series A*, 66A, 191-201.
- MITCHELL, S. E., SIMPSON, M., COULET, L., GOUEDARD, S., HAMBLY, C., MORIMOTO, J., ALLISON, D. B. & SPEAKMAN, J. R. 2024. Reproduction has immediate effects on female mortality, but no discernible lasting physiological impacts: A test of the disposable soma theory. *Proceedings of the National Academy of Sciences*, 121.
- MORANDINI, F., RECHSTEINER, C., PEREZ, K., PRAZ, V., LOPEZ GARCIA, G., HINTE, L. C., VON MEYENN, F. & OCAMPO, A. 2023. ATAC-clock: An aging clock based on chromatin accessibility. *GeroScience*, 46, 1789-1806.
- NUSSEY, D. H., COULSON, T., FESTA-BIANCHET, M. & GAILLARD, J.-M. 2008. Measuring senescence in wild animal populations: towards a longitudinal approach. *Functional Ecology*, 22, 393-406.
- NUSSEY, D. H., KRUUK, L. E., DONALD, A., FOWLIE, M. & CLUTTON-BROCK, T. H. 2006. The rate of senescence in maternal performance increases with early-life fecundity in red deer. *Ecol Lett*, 9, 1342-50.
- PARTRIDGE, L. & BARTON, N. H. 1993. Optimally, mutation and the evolution of ageing. *Nature*, 362, 305-311.

Chapter 1 - References

- PORTER, H. L., BROWN, C. A., ROOPNARINESINGH, X., GILES, C. B., GEORGESCU, C., FREEMAN, W. M. & WREN, J. D. 2021. Many chronological aging clocks can be found throughout the epigenome: Implications for quantifying biological aging. *Aging Cell*, 20.
- PULAKAT, L. & CHEN, H. H. 2020. Pro-Senescence and Anti-Senescence Mechanisms of Cardiovascular Aging: Cardiac MicroRNA Regulation of Longevity Drug-Induced Autophagy. *Front Pharmacol*, 11, 774.
- RUDOLPH, K. L., CHANG, S., LEE, H.-W., BLASCO, M., GOTTLIEB, G. J., GREIDER, C. & DEPINHO, R. A. 1999. Longevity, Stress Response, and Cancer in Aging Telomerase-Deficient Mice. *Cell*, 96, 701-712.
- SALIEV, T. & SINGH, P. B. 2025. Targeting Senescence: A Review of Senolytics and Senomorphics in Anti-Aging Interventions. *Biomolecules*, 15, 860.
- SALMINEN, A., KAUPPINEN, A. & KAARNIRANTA, K. 2012. Emerging role of NF- κ B signaling in the induction of senescence-associated secretory phenotype (SASP). *Cell Signal*, 24, 835-45.
- SAXTON, R. A. & SABATINI, D. M. 2017. mTOR Signaling in Growth, Metabolism, and Disease. *Cell*, 168, 960-976.
- SHIREBY, G. L., DAVIES, J. P., FRANCIS, P. T., BURRAGE, J., WALKER, E. M., NEILSON, G. W. A., DAHIR, A., THOMAS, A. J., LOVE, S., SMITH, R. G., LUNNON, K., KUMARI, M., SCHALKWYK, L. C., MORGAN, K., BROOKES, K., HANNON, E. & MILL, J. 2020. Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex. *Brain*, 143, 3763-3775.
- SLANCHEV, K., STEBLER, J., DE LA CUEVA-MÉNDEZ, G. & RAZ, E. 2005. Development without germ cells: the role of the germ line in zebrafish sex differentiation. *Proc Natl Acad Sci U S A*, 102, 4074-9.

Chapter 1 - References

- SOTO-GAMEZ, A. & DEMARIA, M. 2017. Therapeutic interventions for aging: the case of cellular senescence. *Drug Discov Today*, 22, 786-795.
- SPIERS, H., HANNON, E., WELLS, S., WILLIAMS, B., FERNANDES, C. & MILL, J. 2016. Age-associated changes in DNA methylation across multiple tissues in an inbred mouse model. *Mech Ageing Dev*, 154, 20-3.
- VAN DER RIJT, S., MOLENAARS, M., MCINTYRE, R. L., JANSSENS, G. E. & HOUTKOOPE, R. H. 2020. Integrating the Hallmarks of Aging Throughout the Tree of Life: A Focus on Mitochondrial Dysfunction. *Frontiers in Cell and Developmental Biology*, 8.
- WANG, B., HAN, J., ELISSEEFF, J. H. & DEMARIA, M. 2024a. The senescence-associated secretory phenotype and its physiological and pathological implications. *Nature Reviews Molecular Cell Biology*, 25, 958-978.
- WANG, D., DOU, L., SUI, L., XUE, Y. & XU, S. 2024b. Natural killer cells in cancer immunotherapy. *MedComm (2020)*, 5, e626.
- WEICHHART, T. 2018. mTOR as Regulator of Lifespan, Aging, and Cellular Senescence: A Mini-Review. *Gerontology*, 64, 127-134.
- WILLIAMS, G. C. 1957. PLEIOTROPY, NATURAL SELECTION, AND THE EVOLUTION OF SENESCENCE. *Evolution*, 11, 398-411.
- WILSON, V. L. & JONES, P. A. 1983. DNA Methylation Decreases in Aging But Not in Immortal Cells. *Science*, 220, 1055-1057.
- XIAO, F.-H., WANG, H.-T. & KONG, Q.-P. 2019. Dynamic DNA Methylation During Aging: A "Prophet" of Age-Related Outcomes. *Frontiers in Genetics*, 10.

Chapter 1 - References

YING, K., LIU, H., TARKHOV, A. E., SADLER, M. C., LU, A. T., MOQRI, M., HORVATH, S., KUTALIK, Z., SHEN, X. & GLADYSHEV, V. N. 2024. Causality-enriched epigenetic age uncouples damage and adaptation. *Nature Aging*, 4, 231-246.

ZHENG, Z., LI, J., LIU, T., FAN, Y., ZHAI, Q.-C., XIONG, M., WANG, Q.-R., SUN, X., ZHENG, Q.-W., CHE, S., JIANG, B., ZHENG, Q., WANG, C., LIU, L., PING, J., WANG, S., GAO, D.-D., YE, J., YANG, K., ZUO, Y., MA, S., YANG, Y.-G., QU, J., ZHANG, F., JIA, P., LIU, G.-H. & ZHANG, W. 2024. DNA methylation clocks for estimating biological age in Chinese cohorts. *Protein & Cell*, 15, 575-593.

ZHOU, Y., HE, F., PU, W., GU, X., WANG, J. & SU, Z. 2020. The Impact of DNA Methylation Dynamics on the Mutation Rate During Human Germline Development. *G3 Genes/Genomes/Genetics*, 10, 3337-3346.

Chapter Two

Germline ablation affects somatic gene expression, resulting in improved somatic maintenance

Jean-Charles de Coriolis¹, Ghazal Alavioon², Wilfried Haerty³, Emmanouil Tsakoumis⁴, Monika Schmitz⁴, Alice M. Godden¹, Raheleh Rahabari⁵, Alexei A. Maklakov¹, Simone Immler¹

¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, NR4 6TX, Norwich, UK

² Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, 752 36 Uppsala, Sweden

³ Earlham Institute, Norwich Research Park, NR4 7UZ, Norwich, UK

⁴ Department of Ecotoxicology, Evolutionary Biology Centre, Uppsala University, 752 36 Uppsala, Sweden

⁵ Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK



2. Chapter 2

2.1. Abstract

Genomes are continuously challenged by DNA damage arising from both endogenous metabolic processes and exogenous insults. Although repair pathways counteract this damage, their efficiency declines with age. Evolutionary theory predicts that the germline is prioritised for genomic fidelity over somatic tissues, implying the existence of regulatory signals that differentially modulate germline and somatic maintenance. Here, we show that germline ablation in male zebrafish (*Danio rerio*) enhances somatic genome protection across five tissues (brain, intestine, kidney, spleen and fin). Transcriptomic profiling revealed that germline-free males exhibit baseline upregulation of DNA repair pathways, cell-cycle checkpoint regulators and chromatin integrity genes, together with an accelerated and amplified transcriptional response to γ -irradiation. These changes were accompanied by reduced repression by the DREAM complex, transient activation of Hedgehog signalling and suppression of transposable element activity, particularly LINEs and LTRs, under genotoxic stress. In contrast, germline-carrying controls preferentially upregulated metabolic and biosynthetic programmes, indicating a divergent allocation of transcriptional and energetic resources. We further identified temporal shifts in gene length and exon density consistent with more energy-efficient transcriptional strategies in the germline-ablated soma. Tissue-specific analyses revealed pronounced enhancement of repair-associated pathways in brain and kidney, alongside immune remodelling in intestine and spleen, demonstrating that soma–germline interactions are both tissue-specific and functionally heterogeneous. Notably, testes mounted stronger and more rapid DNA damage responses than any somatic tissue, underscoring their privileged status in genome maintenance.

Together, these results provide direct experimental evidence that germline-derived signals actively suppress somatic maintenance programmes. Germline removal reconfigures transcriptional regulation in favour of genome protection, stress responsiveness and chromatin stability in somatic tissues. This supports evolutionary frameworks in which ageing emerges from regulatory interactions between germline fidelity and somatic investment, linking molecular DNA repair dynamics to broader theories of germline cost and somatic decay.

2.2. Introduction

Cellular genomes are continuously challenged by endogenous metabolic processes and exogenous insults that generate DNA lesions. These are counteracted by specialised DNA repair pathways that preserve genomic integrity (Schumacher et al., 2021). Repair, however, is inherently imperfect, and residual errors accumulate as mutations that compromise genome stability and cellular function. Across the animal kingdom, the germline exhibits higher-fidelity repair than somatic tissues, whereas the soma relies more heavily on error-prone mechanisms, resulting in substantially higher mutation rates. In somatic tissues, declining repair capacity contributes to the age-associated accumulation of mutations and genomic instability (Moore et al., 2021, Gorbunova et al., 2007). Beyond oncogenesis, unrepaired DNA damage interferes with transcription and replication, thereby disrupting core cellular processes and promoting downstream dysfunction in proteostasis, mitochondrial activity and signalling networks (Schumacher et al., 2021). DNA damage is therefore not only a driver of mutational burden but a central contributor to organismal ageing.

Accumulating evidence indicates that the germline actively modulates somatic maintenance, including DNA repair capacity (Moses et al., 2023).

Chapter 2 - Introduction

Germline ablation increases stress resistance and longevity in multiple systems (Moses et al., 2023), and under genotoxic stress the germline can promote enhanced somatic protection (Ermolaeva et al., 2013). These findings are often framed within the disposable soma theory, which proposes a trade-off between investment in reproduction and somatic maintenance (Kirkwood, 1977, Kirkwood and Holliday, 1979). Under this model, organisms may favour energetically cheaper but less accurate repair in somatic tissues to preserve resource for high-fidelity germline maintenance (Maklakov and Immler, 2016). Consistent with this view, germline ablation accelerates somatic repair after genotoxic stress in zebrafish (*Danio rerio*) (Chen et al., 2020).

However, several studies suggest that germline-soma coupling cannot be explained by resource allocation alone (Maklakov and Immler, 2016, Flatt and Partridge, 2018). In *Caenorhabditis elegans*, removal of the somatic gonad abolishes reproduction but does not extend lifespan, whereas germline ablation does, implicating endocrine signalling rather than reproductive cost in the regulation of somatic ageing. Similarly, in the short-lived killifish *Nothobranchius furzeri*, germline ablation enhances somatic maintenance in females and extends lifespan in males, whereas simply reducing resource-intensive reproduction does not (Moses et al., 2023). Together, these findings argue that germline-derived signals actively regulate somatic repair and stress responses, independently of immediate energetic trade-offs.

These observations point to a regulatory model in which germline-soma signalling pathways dynamically tune somatic genome maintenance according to environmental context. When conditions compromise germline function or delay reproduction, enhanced somatic protection may be favoured to preserve survival until reproduction becomes viable again (Ermolaeva et al., 2013). From a life-history perspective, environmental suppression of fertility has long been proposed to strengthen selection on

Chapter 2 - Introduction

somatic robustness and stress tolerance by shifting resource allocation away from reproduction and towards maintenance (Kirkwood and Holliday, 1979; Stearns, 1992), a pattern supported by recent empirical evidence in zebrafish (Irish et al., 2025). Resource-allocation and endocrine-signalling hypotheses are therefore complementary rather than competing, with germline maintenance costs likely to be highly context dependent.

Transposable elements (TEs) provide a sensitive indicator of genome maintenance capacity. TE mobilisation represents an important endogenous source of DNA damage, including double-strand breaks generated during retrotransposition, and TE activity is strongly modulated by environmental and physiological stress (Kazazian and Goodier, 2002, Slotkin and Martienssen, 2007, Levin and Moran, 2011). TE repression is markedly stronger in the germline than in somatic tissues (Rebollo et al., 2012), reflecting the higher value placed on heritable genome integrity. Consequently, comparing TE activity between soma and germline across environmental and genetic context offers a sensitive window into differential investment in genome maintenance.

Here, we investigated how the germline regulates somatic genome maintenance and DNA repair across five somatic tissues in male zebrafish. We examined both baseline conditions and responses to non-lethal γ -irradiation, a defined genotoxic stress that elicits robust responses in both germline and soma. Using transcriptome profiling and TE expression analyses in brain, intestine, kidney, spleen, caudal fin, and, in testes where available, we compared germline-free males (GLF; homozygous *dnd1* knockout) with germline-carrying full-sib controls (GLC; homozygous wild-type). Tissues were collected 3 h and 24 h post-irradiation (hpir) to capture early and intermediate response dynamics. We show that germline ablation broadly enhances somatic genome protection under both standard and stressful conditions, accelerates transcriptional responses to DNA damage,

Chapter 2 - Material and methods

and suppresses TE expression in somatic tissues, supporting a model in which germline-derived signals actively constrain somatic maintenance.

2.3. Material and methods

2.3.1. Fish lines and maintenance

Wild-type AB zebrafish were obtained from the Zebrafish International Resource Centre (ZIRC, University of Oregon, Eugene, OR) and maintained under standard laboratory conditions at the SciLifeLab, Evolutionary Biology Centre, Uppsala University, Sweden. Fish were housed in 3-L tanks at 28°C under a 12:12-h light–dark cycle and fed live *Artemia* and dried flake food (Sparos) three to five times daily, depending on age.

2.3.2. Germline-free fish

CRISPR/Cas9-mediated knockout of the germline-specific gene dead-end (*dnd1*) was performed to generate GLF males, whereas heterozygous carriers (*wt/dnd1*) produced GLC wild-type siblings. Loss of *dnd1* results in the absence of primordial germ cells, leading to sterile males with normal somatic gonadal development.

.

2.3.3. Exposure to gamma irradiation

Twenty-four sexually mature male GLC and GLF fish (>6 months old) were exposed to 20 Gy γ -irradiation from a Cs-137 source at Stockholm University, Sweden. Fish were individually housed in six-well plates containing 16.8 mL system water per well, sealed, and irradiated for ~25 min.

Chapter 2 - Material and methods

2.3.4. Sample collection

Five somatic tissues (intestine, kidney, spleen, brain, caudal fin) and testes (GLC only) were collected at 3h and 24h post-irradiation (hpir) from three fish per group-time, resulting in 24 individuals in total, yielding 126 tissue samples, with three replicates per tissue per condition. Tissues were dissected on ice and flash-frozen in liquid nitrogen (Figure S1).

2.3.5. RNA extraction and sequencing

RNA was extracted using TRIzol, cleaned with TURBO DNA-free (Ambion), and libraries were prepared with NEBNext Ultra II Directional RNA-seq (poly-A selection, 200–300 bp inserts) at Novogene (UK). Sequencing was performed on Illumina NovaSeq 6000 (150-bp paired-end reads).

2.3.6. Differential gene expression analysis

Raw RNA-seq reads were processed using the nf-core/rnaseq pipeline v3.9 executed within the Nextflow workflow framework v22.04.5 (Ewels et al, 2020, Di Tommaso et al., 2017). Read alignment to the reference genome was performed using STAR (Dobin et al., 2013), and transcript abundance was quantified using Salmon (Patro et al., 2017). Differential expression analysis was conducted using edgeR v3.38 based on negative binomial models, with p-values adjusted for multiple testing using the Benjamini–Hochberg false discovery rate procedure (Robinson et al., 2009). Volcano plots were generated using the EnhancedVolcano package (<https://bioconductor.org/packages/release/bioc/html/EnhancedVolcano.html>).

2.3.7. Gene Ontology and chromatin analysis

Gene Ontology (GO) enrichment analysis was performed using pathfindR v2.1.0 (Ulgen et al., 2019), using zebrafish GO annotations derived from MsigDB gene set collections (Liberzon et al., 2015, Liberzon et al., 2011). To enable cross-species comparison, human orthologues were identified using BiomaRt v1.0.7 (Durinck et al., 2009), and analogous GO enrichment analyses were conducted on the corresponding human gene sets. Genomic coordinates of differentially expressed genes were intersected with chromatin accessibility and chromatin state tracks (ATAC-seq) to assess associations with heterochromatic regions (see Supplementary Methods). Exonic and intronic lengths were calculated from Ensembl gene annotations, release 107 (Cunningham et al., 2022).

2.4. Results

2.4.1. Germline ablation reorganises baseline somatic transcription

Across the five somatic tissues examined (brain, caudal fin, spleen, intestine and kidney), GLF and GLC males displayed extensive baseline transcriptional divergence. In total, we identified 6,855 differentially expressed genes (DEGs), with 4,498 upregulated in GLC and 2,357 upregulated in GLF (Tables S1, S2). Although more than 270 Gene Ontology (GO) terms were shared between strains, these were predominantly associated with transcriptional regulation, genome maintenance and protein synthesis, and were largely driven by higher expression in GLC soma (Wilcoxon $V = 1,701.5$, $p < 0.0001$; Figure 1). Semantic similarity analysis revealed that gene sets supporting shared GO terms overlapped only weakly, with low median Resnik similarity values for non-irradiated samples (median = 0.2; interquartile range = 0.08), indicating that similar functional categories were frequently underpinned by distinct molecular programmes in each strain. The highest similarities were observed for double-strand break repair and ribosome biogenesis, whereas

Chapter 2 - Results

microtubule organisation, DNA replication and chromatin remodelling showed particularly low concordance (Figure S2, S3).

At the functional level, GLF-upregulated genes were enriched for processes related to cell-cycle progression and chromatin-linked DNA integrity, including telomere maintenance and DNA packaging. In contrast, GLC-upregulated genes emphasised metabolic activity and adaptive stress responses, such as mitochondrial gene expression and ubiquitin-mediated proteolysis (Figure 1). These opposing functional signatures were largely driven by genes located in consistently closed chromatin regions. In these regions, strain-specific differences in expression were strongly associated with systemic differences in gene architecture, with GLC favouring longer, exon-poor genes and GLF favouring short, exon-rich genes. In contrast, in traditionally open chromatin regions, GLF preferentially upregulated longer exon-poor genes (Figure 2).

These global somatic trends were preserved at the tissue level. Brain and kidney showed enrichment for genome maintenance pathways, intestine and spleen were dominated by immune-related programmes, and the caudal fin was characterised by structural and checkpoint control processes (Figure S3). Among the most strongly differentially expressed baseline genes, GLF soma upregulated *armc9*, implicated in ciliogenesis and cilia-mediated signalling, and *nlrcl1*, associated with antiviral immunity. GLC soma, on the other hand, showed higher expression of immune and differentiation-related genes, including *mhc2dab* and *sppl2* (Figure 3, Table 3, 4, 5).

In addition to functional differences, the global properties of DEGs differed markedly between strains (Figure 3). Mixed-effects modelling identified a significant interaction between contrast and strain for gene expression level ($\chi^2 = 2,252$, $df = 17$, $p < 0.0001$), gene length ($\chi^2 = 1,538$, $df = 17$, $p < 0.0001$) and exonic density ($\chi^2 = 301,334$, $df = 17$, $p < 0.0001$). In GLC, DEGs were typically

Chapter 2 - Results

longer, exhibited lower exonic density and were expressed at lower levels relative to the transcriptome average (Tables S6, S7). In contrast, GLF soma preferentially engaged shorter, exon-rich genes. This pattern is consistent with reduced transcriptional expenditure and increased splicing efficiency, as proposed by models linking shorter gene length and higher exon-to-intron ratio to lower energetic cost for equivalent transcriptional output (Hausser et al., 2019, Heyn et al., 2015, Jonkers et al., 2014).

Chapter 2 - Results

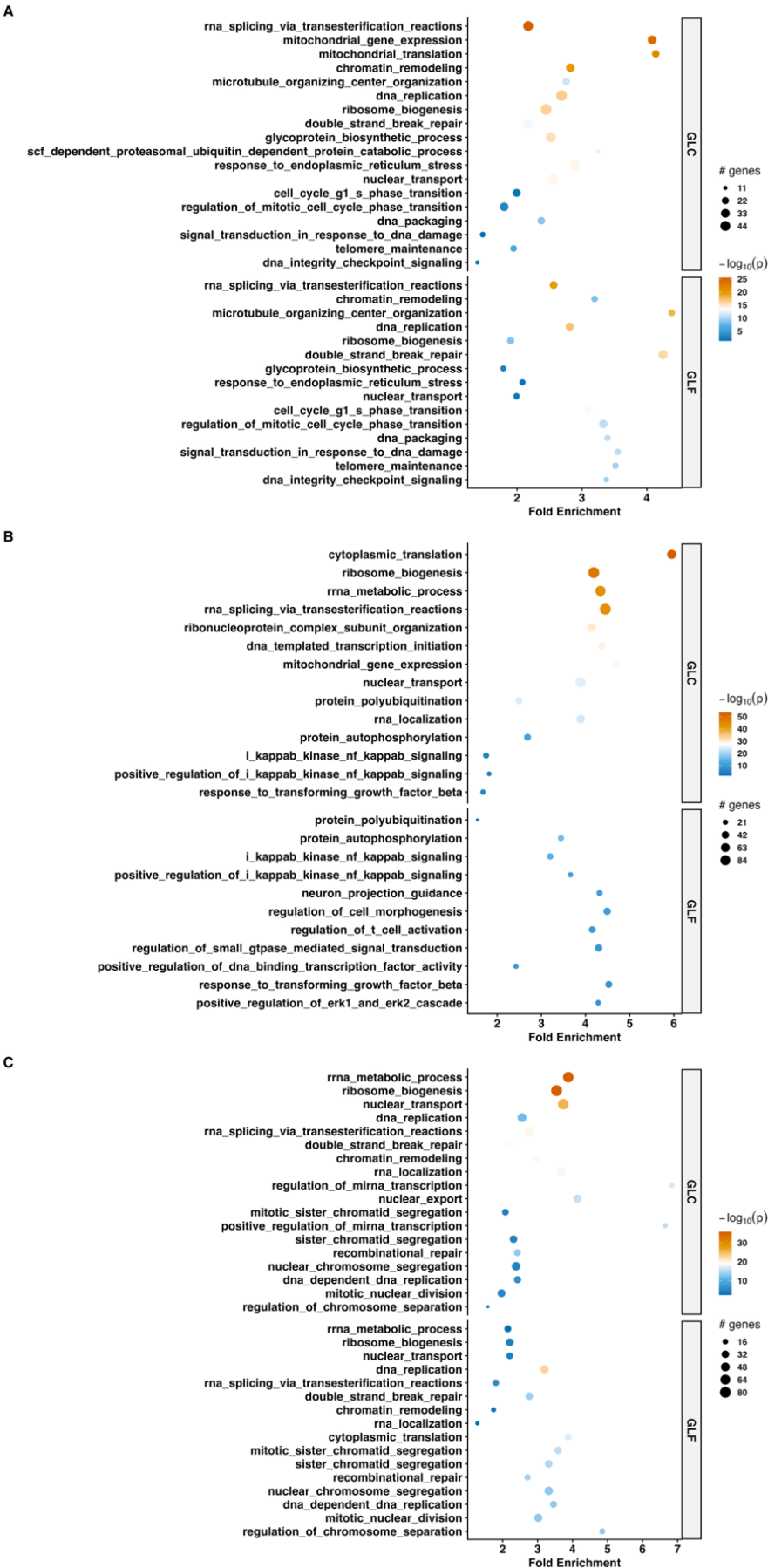


Figure 1: Germline absence enables rapid somatic shift from damage response to genomic fidelity programmes post-irradiation.

GO term enrichment across three timepoints post-irradiated **A)** non-irradiated (nonIR), **B)** 3 hours post-irradiated (3hPIR), and **C)** 24 hours post-irradiated (24hPIR), stratified by host strain. Each panel displays significantly enriched GO terms grouped by strain. Circle colour denotes statistical significance ($-\log_{10}$ p-value; orange = high significance, blue = low), and circle size reflects the number of genes associated with each GO term.

Chapter 2 - Results

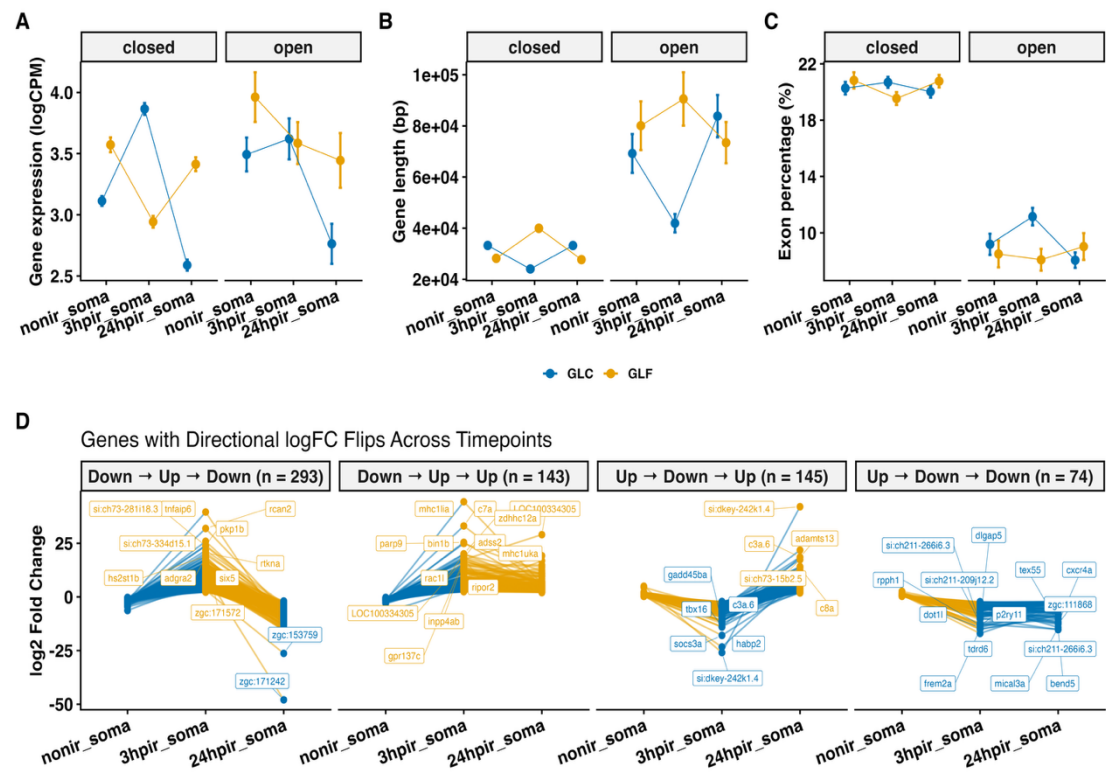


Figure 2: Germline ablation elicits a dynamic expression of genes and chromatin structure in response to irradiation.

A) Displays gene expression (logCPM), **B)** gene length, and **C)** exon percentage, respectively, stratified by chromatin accessibility status (open vs. closed chromatin). Each plot compares the soma of GLC vs GLF across the same post-irradiation timepoints. Colours denote strain identity. Points denote group means, and error bars represent standard error. **D)** Line plots display temporal expression trajectories of genes ($\text{FDR} < 0.05$) that exhibit directional flips in logFC across timepoints (nonir_soma, 3hpir_soma, 24hpir_soma). Genes were categorised into four flip patterns: (i) Down $\rightarrow$ Up $\rightarrow$ Up, (ii) Down $\rightarrow$ Up $\rightarrow$ Down, (iii) Up $\rightarrow$ Down $\rightarrow$ Down, and (iv) Up $\rightarrow$ Down $\rightarrow$ Up, based on the sign of logFC at each timepoint. Each facet represents one flip pattern. Lines connect the expression values for individual genes, coloured by fish strain. Points indicate mean logFC per gene per timepoint. Labels denote the top 15 genes per flip category, ranked by absolute logFC. Label placement was optimised to minimise overlap using repulsion-based adjustment.

Chapter 2 - Results

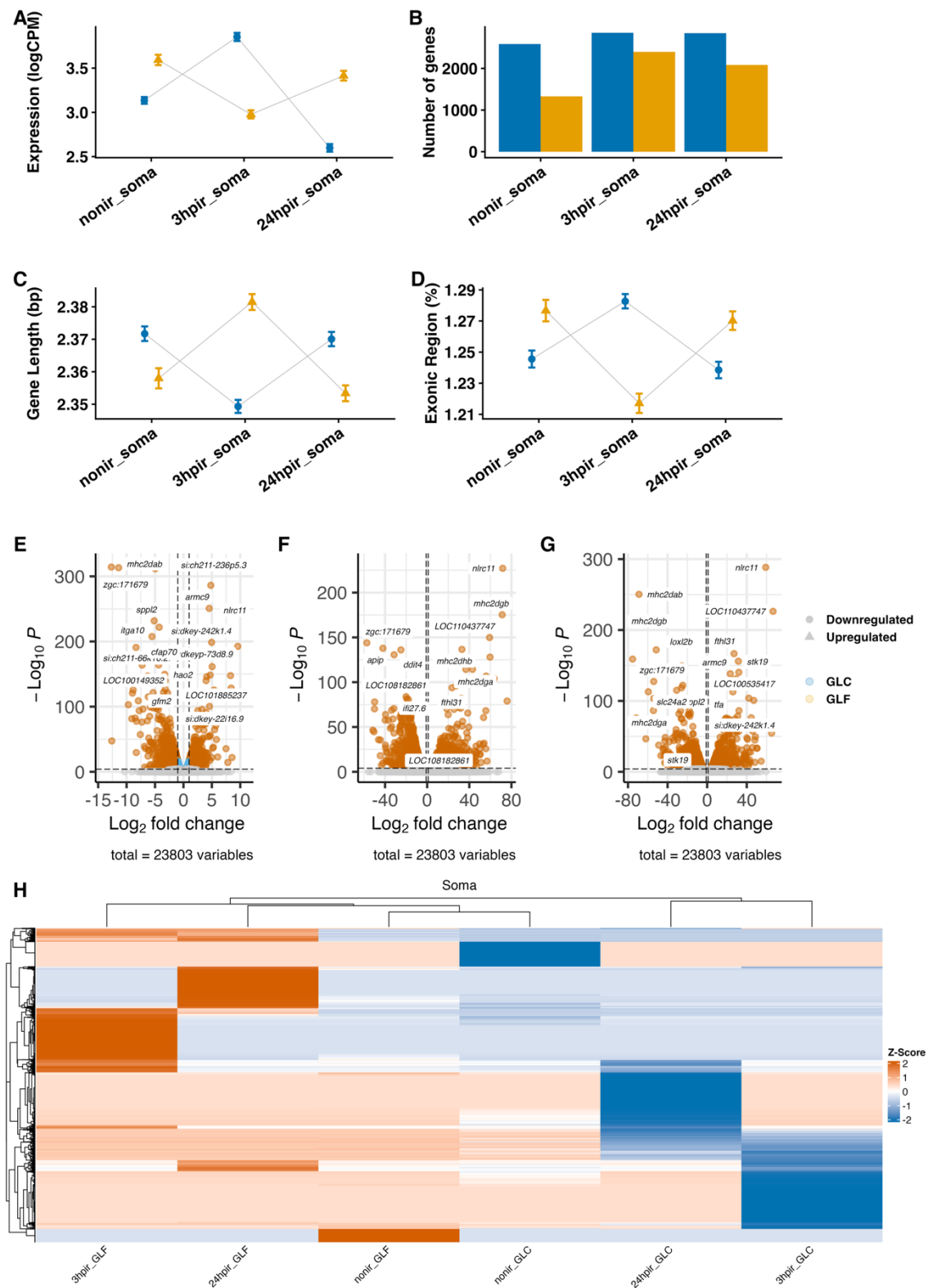


Figure 3 : Bidirectional shifts in gene regulation reveal temporally distinct somatic responses to germline ablation and irradiation.

Panels **(A-D)** explore the interaction between germline ablation (GLC vs GLF) and irradiation (nonir, 3hpir & 24hpir) on the soma, where nonir_soma refers to contrast C1 (nonir GLC vs nonir GLF), 3hpir_soma to contrast C2.3 (3hpir GLC vs 3hpir GLF) & 24hpir_soma to contrast C2.24 (24hpir GLC vs 24hpir GLF). **A)** Shows gene expression (logCPM), **B)** Number of differentially expressed genes, **C)** Gene length (bp) and **D)** Percentage exonic region of differentially expressed genes. Points denote group means, and error bars represent standard errors. Volcano plots **(E-G)**, where points represent a single gene, the x-axis shows the log₂ fold change in expression between GLC and GLF fish, and the y-axis displays the -log₁₀(p-value), which indicates statistical significance. Genes with a positive logFC are expressed more in GLF tissues, while downregulated ones are expressed more in GLC tissues. Non-significant genes are depicted in grey. Horizontal and vertical dashed lines represent the significance threshold (-log₁₀(p-value) = 1.3) and fold-change thresholds (log₂FC = ±0.6), respectively. **E)** Shows the differential expression of contrast C1, **F)** shows the differential expression of contrast C2.3, and **G)** shows the differential expression of contrast C2.24. **(H)** Gene expression values (logCPM) were normalised per gene across the soma of GLC and GLF fish to compute Z-scores, representing relative expression levels. Positive Z-scores (shaded orange) indicate expression above the gene's mean, whereas negative Z-scores (shaded blue) indicate expression below the mean. Rows correspond to genes, columns represent samples grouped by strain and timepoint post-irradiation.

2.4.2. Baseline differences in DREAM complex architecture

The conserved DREAM repressor complex coordinates cell-cycle and DNA repair by silencing proliferation-associated genes during quiescence and following DNA damage, including *pcna*, *ccna2*, *rad51*, and *cdk1* (Bujarrabal-Dueso et al., 2023). In non-irradiated soma, its transcriptional configuration differed markedly between strains (Figure 4). GLC expressed three core DREAM components (*lin54*, *rbf1*, *e2f5*), consistent with an intact and potentially active repressor complex. In contrast, GLF expressed only *e2f4*, indicating a partial or weakened DREAM assembly.

In parallel, GLF showed elevated expression of canonical DREAM target genes (*pcna*, *ccna2*, *rad51*; Figure 4, S4), consistent with reduced DREAM-mediated repression and a higher basal engagement of replication and DNA repair programmes in the soma (Sadasivam and DeCaprio, 2013).

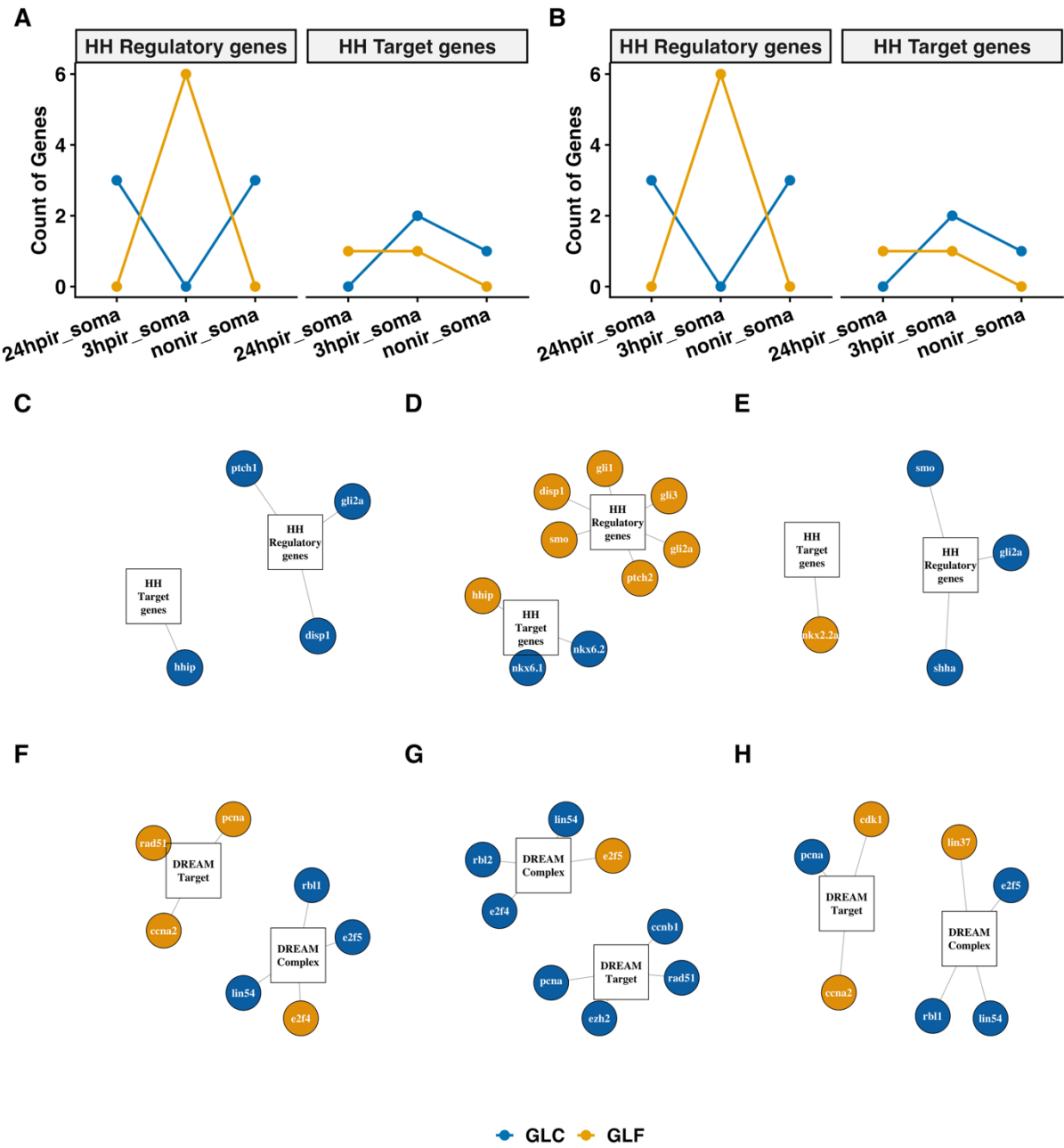


Figure 4: GLF fish show distinct hedgehog regulatory gene and DREAM complex target genes expression patterns at 3hpir.

Panels **(A-B)** Line plots showing the number of differentially expressed genes (FDR < 0.05) across timepoints (nonir_soma, 3 hpir_soma, and 24hpir_soma) for GLC and GLF fish. Gene counts are grouped by gene type (Hedgehog regulatory vs target; DREAM complex vs DREAM targets; see facet panels). Lines and points represent the dynamics of DE gene counts across developmental timepoints for each fish, coloured according to either GLC or GLF fish. **A)** Hedgehog complex, **B)** Dream Complex. Panels **(C-H)** display network visualisations of the relationships between differentially expressed genes (FDR < 0.05) and their classification as regulatory or target gene (for Hedgehog **(C, D, E)** and Dream pathways **(F, G, H)**, for each timepoint-specific contrasts. Panel **(C & F)** non-irradiated soma, **(D & G)** Soma 3hpir, **(E & H)** Soma 24hpir. Circles represent individual genes, and squares represent gene type categories. Nodes are coloured by fish population. Edges connect genes to their respective functional category. Multiline labels are used for clarity in type nodes; gene node labels are shown in white. Networks were arranged using the Fruchterman–Reingold layout algorithm to emphasise topological separation between categories.

2.4.3. Germline ablation accelerates and redirects the somatic response to genotoxic stress

To assess somatic responses to DNA damage, males were exposed to 20 Gy γ -irradiation with a ^{137}Cs source and sampled at 3 hpir and 24 hpir. Both strains mounted strong transcriptional responses, but these differed in timing, functional composition and gene usage (Figure 3). At 3 hpir, 8,252 DEGs were differentially expressed (3,779 higher in GLF; 4,473 higher in GLC), and at 24 hpir, 7,991 DEGs were detected (3,372 higher in GLF; 4,619 higher in GLC; Table S2).

Although many GO terms were shared between strains (208 at 3 hpir; 262 at 24 hpir), the genes supporting these categories showed limited overlap, particularly immediately after irradiation (Figure S2, S3). At 3 hpir, similarity confined to core damage-response processes such as chromosome segregation, DSB repair and RNA catabolism, whereas processes such as protein autophosphorylation, nuclear localisation and I- κ B kinase/NF- κ B signalling were supported by largely distinct gene sets. By 24 hpir, gene-level overlap increased for ribosomal large subunit biogenesis and RNA metabolism indicating partial convergence of translational control, while DNA repair, chromatin regulation and nucleocytoplasmic trafficking remained strain specific (Semantic similarity median was 0.2 (IQR 0.07) at 3 hpir and 24 hpir; Figures 1, 3).

Among highly divergent function categories, the “ageing” GO term showed strong temporal and strain dependence. It was supported exclusively by GLC genes at baseline, shared between strains immediately after irradiation, and by 24 hpir became largely restricted to GLC genes (Figure 5). This pattern suggests that irradiation transiently synchronises ageing-related transcriptional programmes before reinstating strain-specific control.

Chapter 2 - Results

Functionally, the early response in GLC (3 hpir) was dominated by biosynthetic housekeeping processes, including cytoplasmic translation and ribosome biogenesis (Figure 1). In contrast, GLF upregulated pathways involved in signal transduction, immune modulation and structural remodelling, consistent with a faster transition into adaptive and reparative states (Figure 1). By 24 hpir, GLC remained biased towards transcriptional and biosynthetic activity, whereas GLF shifted strongly towards DNA replication, mitosis, homologous recombination and proteasomal degradation (Figures 1, 5). Together, these patterns indicate that germline ablation accelerates the engagement of repair and turnover pathways while redirecting the balance away from sustained biosynthetic investment.

Chapter 2 - Results

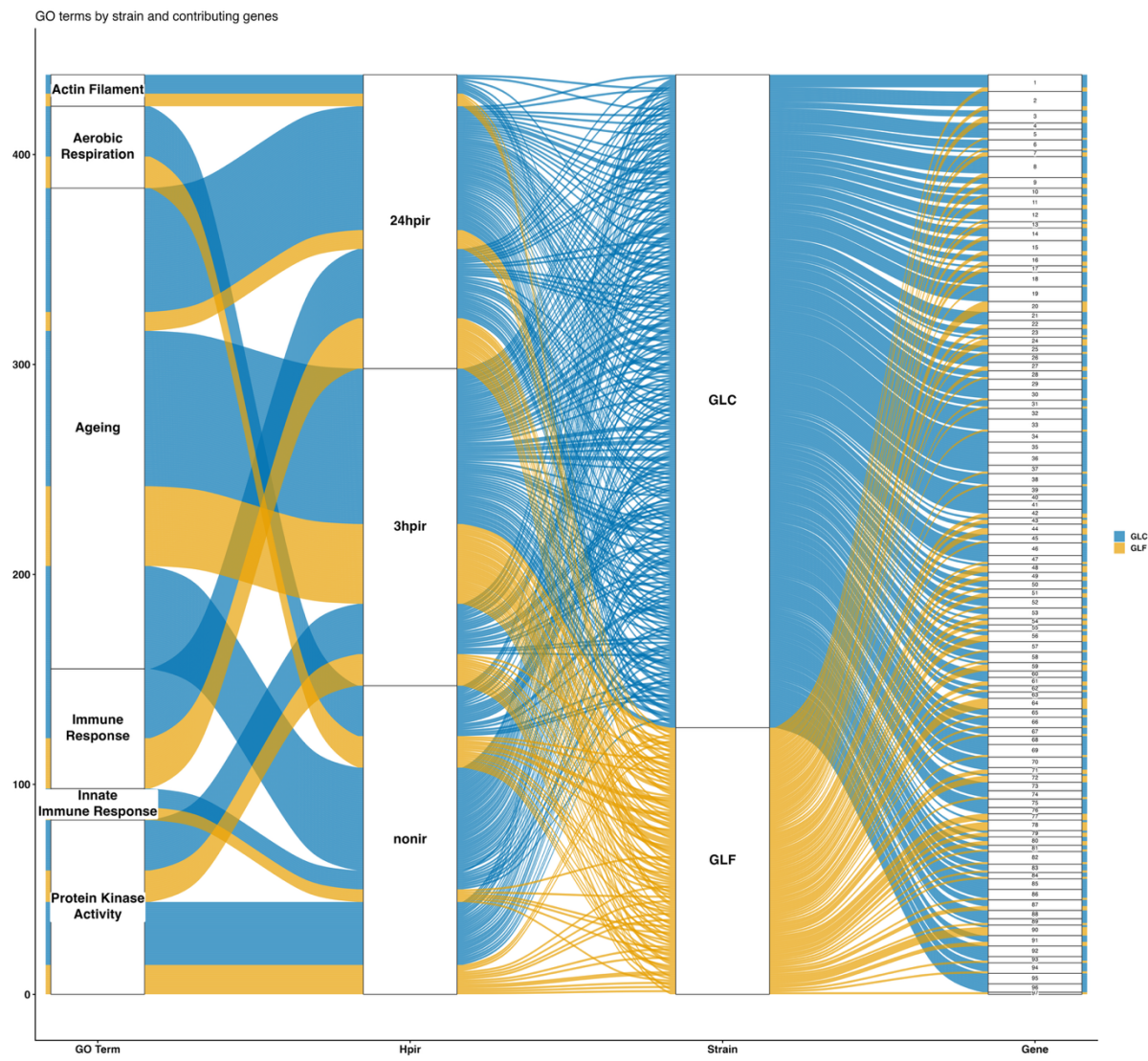


Figure 5: Treatment-dependent conservation of ageing-related GO terms across zebrafish strains.

Alluvial plot illustrating the relationships between enriched Gene Ontology (GO) terms, experimental condition, strain, and contributing gene sets. Each stratum on the left represents a GO biological process term, with flows tracking their association across irradiation condition/timepoint, strain, and underlying gene sets. Ribbon widths are proportional to the number of genes contributing to each connection, highlighting shared versus strain-specific contributions. While major ageing-related GO terms are observed in both strains, GLC fish show stronger representation and higher gene-level contribution to ageing-associated processes compared with GLF. Colours denote strain-level clustering.

2.4.4. A temporal 'directional flip' in genomic features after irradiation

Beyond functional categories, gene architecture and expression properties also showed systemic and reversible strain differences after irradiation (Tables S6, S7; Figures 2, 3). At baseline, relative to GLF, GLC DEGs were characterised by lower expression, lower exonic density and longer gene length. At 3 hpir, this relationship inverted, with GLF engaging longer, exon-poorer and more weakly expressed genes. By 24 hpir, these properties reverted towards the original baseline configuration. Thus, irradiation induced a transient “directional flip” in the genomic architecture of the transcriptional response.

These reversals were not uniformly driven by all tissues. Brain and kidney contributed most strongly to expression divergence; brain, caudal fin and spleen dominated gene-length divergence; and brain and intestine primarily shaped exonic-density differences.

Genes that switched from being higher in GLF at 3 hpir to higher in GLC at 24 hpir were enriched for tissue regeneration (Figures 1, 2), muscle maintenance and cellular catabolism, consistent with earlier activation of reparative and metabolic restructuring in GLF soma (Schiaffino and Reggiani, 2011). Among the most prominently temporally flipping genes were MHC class II haplotypes (*mhc2dgb*, *mhc2dhb*, *mhc2dga*), which consistently shifted from GLF-biased expression at 3 hpir to GLC-biased expression at 24 hpir across tissues (Table S5). Notably, *mhc2dab* remained uniquely high in GLC soma (Figures 2, 3), highlighting stable strain-specific immune regulation.

Chapter 2 - Results

2.4.5. Tissue-specific programmes reveal rapid adaptation and repair in GLF

At 3 hpir, GLC tissues broadly upregulated biosynthetic and proliferative metabolism, including cytoplasmic translation, ribosome biogenesis and DNA replication. In contrast, GLF tissues preferentially upregulated pathways associated with tissue repair, immune signalling and stress response (Figure 1), reinforcing the idea of earlier adaptive reprogramming. This distinction is evident at the tissue level. In GLF, the brain showed enrichment for coagulation and ERK signalling, the intestine and kidney for DNA damage response and NF- κ B signalling, and the caudal fin for hormone-mediated pathways. By 24 hpir, GLF brain, intestine and spleen strongly upregulated recombinational repair, chromosome segregation and mitotic division, while the caudal fin showed signatures of metabolic adaptation (Figure 1). In contrast, GLC tissues at 24 hpir continued to emphasise mitochondrial energy metabolism, ribosome biogenesis and RNA splicing (Figure 1), indicating prolonged prioritisation of biosynthetic and energetic homeostasis rather than rapid engagement of repair and turnover.

2.4.6. Testes trajectories contextualise somatic responses

In GLC testes, irradiation induced a stronger and more dynamic transcriptional response than in any somatic tissue. At 3 hpir, 6,710 genes were upregulated and 9,118 downregulated, and at 24 hpir, 6,989 were upregulated and 7,736 downregulated, with widespread reversals between the two timepoints (Table S2; Figure S5).

At 3 hpir, testes rapidly activated DNA repair, chromatin remodelling and transcriptional control while simultaneously repressing many pathways that were induced in GLC soma, including ribosome biogenesis, RNA splicing and mitochondrial translation (Table S8). By 24 hpir, testes upregulated ribosome

Chapter 2 - Results

biogenesis, telomere maintenance, mitochondrial gene expression and DSB repair while downregulating immune signalling, returning towards homeostasis more rapidly than any somatic tissue (Table S8).

2.4.7. Temporal control by DREAM and hedgehog signalling

Irradiation reshaped the baseline configuration of the DREAM repressor network (Figure 4). At 3 hpir, GLC upregulated four DREAM targets (*ezh2*, *pcna*, *ccnb1*, *rad51*), whereas GLF showed little activation of this axis. By 24 hpir, the pattern reversed: GLF expressed *ccna2* and *cdk1*, while GLC expressed *pcna*, approximating the respective baseline configurations observed under non-irradiated conditions. This dynamic is consistent with DREAM-mediated temporal control of cell-cycle and DNA repair gene expression following genotoxic stress (Figure 4).

Hh signalling, which regulates stem-cell maintenance, proliferation and regenerative capacity in adult tissues and mediates germline-soma communication (Jing et al., 2023, Jia et al., 2015), also showed strain-specific activation (Figure 4). In GLF, Hh regulators (*disp1*, *ptch2*, *gli1*, *gli2a*, *gli3*, *smo*) were transiently induced in the spleen at 3 hpir, followed by broader somatic upregulation of *gli2a*, *ssha* and *smo* at 24 hpir. This pattern indicates a short-lived, context-dependent activation of Hh signalling, consistent with a transient regenerative or niche-supportive response in the absence of the germline.

2.4.8. Germline ablation reshapes somatic transposable-element dynamics

TEs exhibited strong strain- and time-specific behaviour that closely paralleled gene expression dynamics (Figures 6, S7). At baseline, 522 TEs were

Chapter 2 - Results

differentially expressed (239 higher in GLF; 283 in GLC; Table S9). After irradiation, 229 TEs were differentially at 3 hpir and 256 at 24 hpir (Table S9).

A directional switch mirrored the gene-expression response; GLF reduced TE expression at 3 hpir and increased it at 24 hpir, whereas GLC showed the opposite pattern (Figure 6). This indicates that germline ablation alters not only the timing of gene activation but also the temporal control of repetitive-element activity.

There was also a TE class-specific response. At baseline, GLF showed a positive correlation between LINE copy-number and diversity. After irradiation, both strains converged on reduced diversity with increased copy number (Figure S6). At 3 hpir, GLF LINE expression declined, and differentially expressed repeats were located closer to genes on opposite strands, consistent with increased regulatory interference in the absence of the germline (Figure S6). GLC uniquely expressed L2-1_DR at this timepoint (Figure S6). LTR dynamics were largely driven by Gypsy elements and showed an inverse relationship between copy number and diversity across conditions. DNA transposons were particularly elevated in GLF at 24 hpir and were characterised by increased copy number and younger inferred insertion age.

When DE genes and DE TEs were intersected, GLC showed greater overlap at baseline (Figure S7), suggesting tighter basal coupling between gene expression and TE activity, potentially through shared chromatin accessibility or epigenetic control. In GLF, this overlap was weaker at baseline, consistent with stronger TE repression. After irradiation however, overlap increased specifically on chromosome 4 in GLF and was dominated by LTR elements (Figure S7). Given that LTRs are stress-responsive and can function as enhancers or alternative promoters (Chuong et al., 2016, Feschotte, 2008),

Chapter 2 - Results

this pattern is consistent with local rewiring of regulatory networks during the DNA damage response.

No comparable chromosome-wide structure was observed for DE genes or DE TEs in isolation, indicating that the coordinated response emerges specifically from their interaction rather than from independent chromosomal effects.

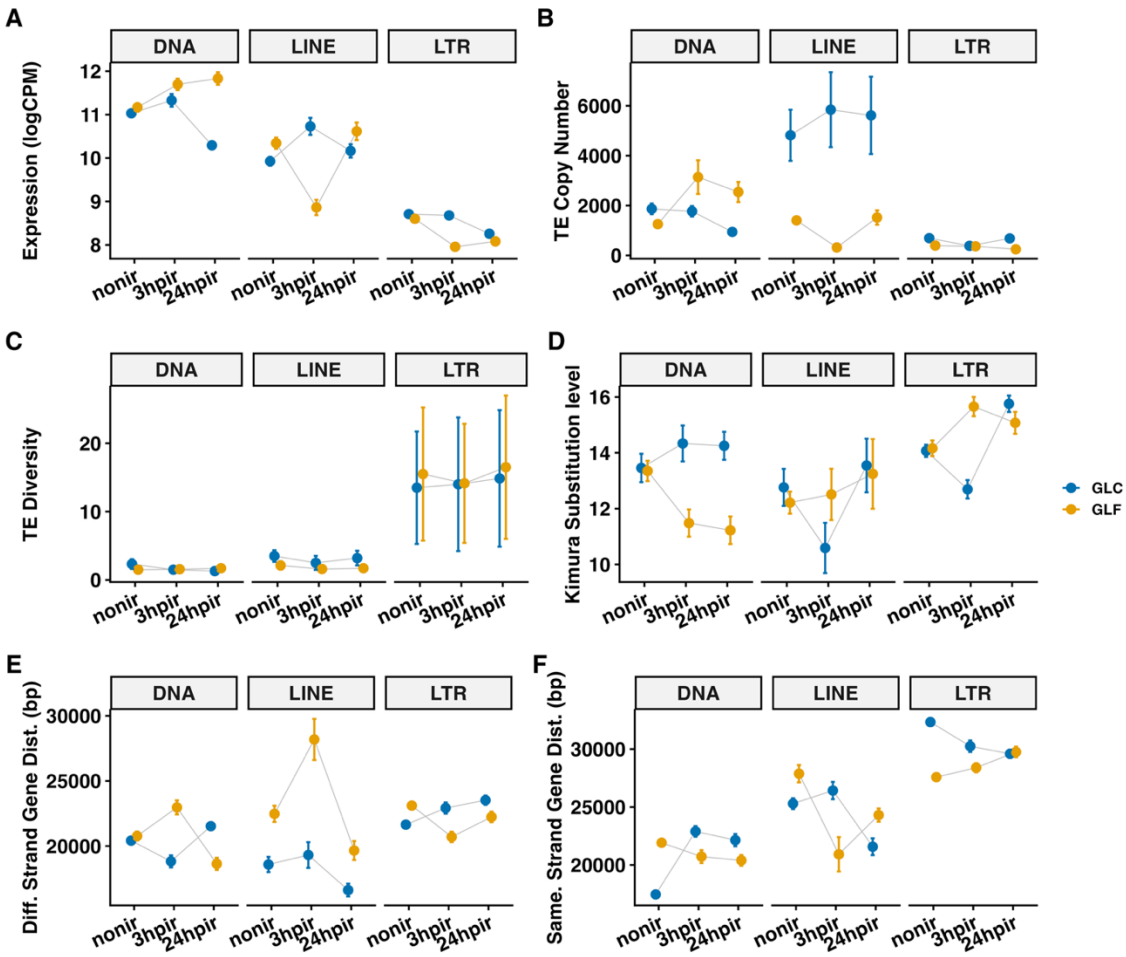


Figure 6: Age-dependent transposable element expression in somatic tissues shows strain- and timepoint-specific directional shifts following irradiation.

A) presents the mean log counts per million (CPM) of significantly expressed TE Repeats (FDR < 0.05). **B)** the mean TE Copy number of each repeat per class, **C)** the number of TE subfamilies differentially expressed, **D)** the mean synonymous substitution rate (Ks) for the same TE Repeats, used as a proxy for TE evolutionary age. **E)** the distance of expressed TE to the closest gene on the opposite strand (in bp), and **F)** shows the distance to genes on the same strand. Points denote group means, and error bars represent standard errors. Strain identity is denoted by colour and maintained consistently across panels. All panels examined TE metrics across non-irradiated, 3 hours post-irradiation (3hpir), and 24 hours post-irradiation (24hpir) conditions.

2.5. Discussion

In the absence of the germline, somatic tissues expressed a greater number of shorter, exon-rich genes involved in cell communication, proliferation and genome integrity, a strategy consistent with reduced energetic cost and increased splicing efficiency (Pai et al., 2017, Hausser et al., 2019). At baseline, DREAM complex repression appears attenuated in GLF fish, permitting selective activation of nucleotide-excision and homologous-recombination repair pathways. Following irradiation, DREAM activity is re-established, restraining proliferation and restoring cell-cycle control (Bujarrabal-Dueso et al., 2023, Hwang and Kim, 2025). This dynamic likely limits the propagation of highly damaged cells while maintaining genomic integrity (Agrawal and Sengupta, 2025, Schmidt et al., 2024).

Germline ablation also produced a faster and more clearly phased response to genotoxic stress. GLF fish mounted an early adaptive response dominated by signal transduction, immune modulation and structural remodelling, before transitioning by 24 h to replication-coupled repair and proteostasis. In contrast, GLC fish preferentially activate biosynthetic and transcriptional programmes early and delay engagement of genome-maintenance modules. Temporal ‘directional flips’ in gene length, exonic density and expression profiles, and tissue drivers (particularly the brain), underscore a rapid and systematic reprioritisation of resources in GLF. Although protein synthesis and translational control emerged as shared functional categories between strains, the underlying gene usage differed substantially, indicating that germline ablation reshapes the regulatory implementation of protein-related programmes rather than uniformly altering proteostasis capacity.

Primary cilia act as signalling hubs that integrate pathways central to growth, metabolism and tissue homeostasis, and accumulating evidence links ciliary dysfunction to ageing-related phenotypes, including neurodegeneration,

Chapter 2 - Discussion

metabolic dysregulation and impaired regenerative capacity (Anvarian et al., 2019, Morrison, 2021, Palla et al., 2022). Upregulation of *armc9* in GLF, therefore, suggests enhanced ciliary maintenance and signalling capacity. In parallel, preservation of chromatin organisation during DNA repair is essential for transcriptional fidelity, and elevated expression of chromatin regulators in GLF may buffer epigenetic drift and reinforce long-term genomic stability (Teschendorff et al., 2013, Rodrigues et al., 2014).

Ageing is also accompanied by immune decline, particularly through reduced MHC class II expression and impaired antigen presentation (Wang and Liu, 2022, Bertucci-Richter and Parrott, 2023). In GLF, dynamic regulation of MHC loci, characterised by early activation followed by coordinated resolution, suggests efficient detection and clearance of damaged cells, followed by restoration of homeostasis. This pattern is consistent with established links between senescence, MHC class II expression, and tissue remodelling during ageing (Sinning et al., 2023, Herrero et al., 2001), as well as broader DNA damage-immune system crosstalk (Bujarrabal-Dueso et al., 2023, Schmidt et al., 2024).

In vivo, DREAM target engagement inverts after irradiation, consistent with transient relaxation of cell-cycle to facilitate high-fidelity repair before re-establishing proliferative control (Bujarrabal-Dueso et al., 2023, Schmidt et al., 2024). Hh signalling is essential for tissue regeneration and metabolic regulation but becomes pro-ageing when chronically dysregulated and shows a similarly precise pattern; suppressed at baseline in GLF and transiently induced under genotoxic stress. This mirrors tightly controlled activation in regenerative contexts (Rajan et al., 2023, Male et al., 2020, Sun et al., 2020) and potentially limiting chronic, deleterious signalling (Baechle et al., 2023).

Chapter 2 - Data availability

Collectively, across tissues and molecular layers, germline ablation reprogrammes the soma toward rapid, immune-coupled adaptation and high-fidelity repair coordinated by pathway-level regulators such as DREAM and Hh. Despite distinct strain identities, GLF fish adopt elements of a “germline-like” temporal logic, supporting models in which resource reallocation from reproduction to somatic maintenance enhances resilience and mitigates age-associated decline. This framework aligns with systemic regulators of ageing, including mTOR and NF- κ B signalling. Pharmacological mTORC1 inhibition (such as rapamycin) reduces NF- κ B activity, DNA damage, and immune dysfunction in ageing models (Lamming, 2016, Dumas and Lamming, 2020), paralleling several features observed in GLF and reinforcing disposable-soma trade-offs.

Thus, germline ablation establishes an energy-efficient somatic programme characterised by rapid immune-coupled adaptation and timely deployment of high-fidelity DNA repair. Temporal control by DREAM and Hedgehog pathways, integrating both gene and TE regulation, underpins a faster and more robust recovery from genotoxic stress in GLF than in GLC. TE dynamics, including class-specific behaviour and age-associated shifts, both reflect and potentially facilitate this rewiring. Together, these findings position germline status as a systems-level modulator of somatic homeostasis and ageing and suggest that transient induction of germline-like repair programmes may offer a route to enhanced genome stability without compromising somatic integrity.

2.6. Data availability

All sequencing data, including RNA-seq data, generated for this project are available under accession PRJEB102813 (under embargo until acceptance for

Chapter 2 - Author contributions

publication), <https://www.ebi.ac.uk/ena/browser/view/PRJEB102813>. All scripts and supplementary data are available at the git repository : <https://github.com/asdfjohnny1/germline-ablation-transcriptomics.git>.

Detailed methodology and settings for bioinformatic pipelines used can be found in supplementary methods.

2.7. Author contributions

The studies were conceptualised and designed by S.I., J.C.D.C. and G.A.

Data analysis and visualisation were carried out by J.C.D.C.

Samples were collected and processed by G.A., E.T and M.S.

A.M.G. provided bioinformatic support.

All code was conceived and generated by J.C.D.C.

Supervision was provided by S.I., W.H., R.R., A.A.M.

Funding was provided by S.I.

The first draft of the manuscript was written by J.C.D.C., S.I. and A.M.

All authors revised and approved the manuscript.

2.8. Ethics declarations

Competing interests: All authors declare no competing interests

All zebrafish work was completed in accordance with institutional, local and national laws and governing bodies.

2.9. Extended data

See file “Appendix Chapter 2” for supplementary methods, figures, tables and legends.

All files and processed data files are available on the GitHub page : <https://github.com/asdfjohnny1/germline-ablation-transcriptomics.git>.

2.10. Abbreviations

TEs – Transposable Elements

GLC – Germline Carrying

GLF – Germline Free

DEGs – Differentially expressed Genes

GO – Gene Ontology

3 hpir – 3 hours post irradiation

24 hpir – 24 hours post irradiation

Hh – Hedgehog

Chr4 – Chromosome 4

Dnd1 – Dead-end 1 gene

2.11. References

AGRAWAL, R. & SENGUPTA, S. 2025. p53 regulates DREAM complex-mediated repression in a p21-independent manner. *The EMBO Journal*, 44, 2279-2297.

ANVARIAN, Z., MYKYTYN, K., MUKHOPADHYAY, S., PEDERSEN, L. B. & CHRISTENSEN, S. T. 2019. Cellular signalling by primary cilia in development, organ function and disease. *Nat Rev Nephrol*, 15, 199-219.

BAECHLE, J. J., CHEN, N., MAKHIJANI, P., WINER, S., FURMAN, D. & WINER, D. A. 2023. Chronic inflammation and the hallmarks of aging. *Mol Metab*, 74, 101755.

BERTUCCI-RICHTER, E. M. & PARROTT, B. B. 2023. The rate of epigenetic drift scales with maximum lifespan across mammals. *Nature Communications*, 14.

BUJARRABAL-DUESO, A., SENDTNER, G., MEYER, D. H., CHATZINIKOLAOU, G., STRATIGI, K., GARINIS, G. A. & SCHUMACHER, B. 2023. The DREAM complex functions as conserved master regulator of somatic DNA-repair capacities. *Nature Structural & Molecular Biology*, 30, 475-488.

CHEN, H.-Y., JOLLY, C., BUBLYS, K., MARCU, D. & IMMLER, S. 2020. Trade-off between somatic and germline repair in a vertebrate supports the expensive germ line hypothesis. *Proceedings of the National Academy of Sciences*, 117, 8973-8979.

CHUONG, E. B., ELDE, N. C. & FESCHOTTE, C. 2016. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science*, 351, 1083-1087.

CUNNINGHAM, F., ALLEN, J. E., ALLEN, J., ALVAREZ-JARRETA, J., AMODE, M. R., ARMEAN, I. M., AUSTINE-ORIMOLOYE, O., AZOV, A. G., BARNES, I.,

Chapter 2 - References

BENNETT, R., BERRY, A., BHAI, J., BIGNELL, A., BILLIS, K., BODDU, S., BROOKS, L., CHARKHCHI, M., CUMMINS, C., DA RIN FIORETTO, L., DAVIDSON, C., DODIYA, K., DONALDSON, S., EL HOUDAIGUI, B., EL NABOULSI, T., FATIMA, R., GIRON, C. G., GENEZ, T., MARTINEZ, J. G., GUIJARRO-CLARKE, C., GYMER, A., HARDY, M., HOLLIS, Z., HOURLIER, T., HUNT, T., JUETTEMANN, T., KAIKALA, V., KAY, M., LAVIDAS, I., LE, T., LEMOS, D., MARUGÁN, J. C., MOHANAN, S., MUSHTAQ, A., NAVEN, M., OGEH, D. N., PARKER, A., PARTON, A., PERRY, M., PILIŽOTA, I., PROSOVETSKAIA, I., SAKTHIVEL, M. P., SALAM, A. I. A., SCHMITT, B. M., SCHUILENBURG, H., SHEPPARD, D., PÉREZ-SILVA, J. G., STARK, W., STEED, E., SUTINEN, K., SUKUMARAN, R., SUMATHIPALA, D., SUNER, M. M., SZPAK, M., THORMANN, A., TRICOMI, F. F., URBINA-GÓMEZ, D., VEIDENBERG, A., WALSH, T. A., WALTS, B., WILLHOFT, N., WINTERBOTTOM, A., WASS, E., CHAKIACHVILI, M., FLINT, B., FRANKISH, A., GIORGETTI, S., HAGGERTY, L., HUNT, S. E., GR, I. I., LOVELAND, J. E., MARTIN, F. J., MOORE, B., MUDGE, J. M., MUFFATO, M., PERRY, E., RUFFIER, M., TATE, J., THYBERT, D., TREVANION, S. J., DYER, S., HARRISON, P. W., HOWE, K. L., YATES, A. D., ZERBINO, D. R. & FLICEK, P. 2022. Ensembl 2022. *Nucleic Acids Res*, 50, D988-d995.

DI TOMMASO, P., CHATZOU, M., FLODEN, E. W., BARJA, P. P., PALUMBO, E. & NOTREDAME, C. 2017. Nextflow enables reproducible computational workflows. *Nat Biotechnol*. United States.

DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29, 15-21.

DUMAS, S. N. & LAMMING, D. W. 2020. Next Generation Strategies for Geroprotection via mTORC1 Inhibition. *The Journals of Gerontology: Series A*, 75, 14-23.

Chapter 2 - References

- DURINCK, S., SPELLMAN, P. T., BIRNEY, E. & HUBER, W. 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*, 4, 1184-91.
- ERMOLAEVA, M. A., SEGREF, A., DAKHOVNIK, A., OU, H.-L., SCHNEIDER, J. I., UTERMÖHLEN, O., HOPPE, T. & SCHUMACHER, B. 2013. DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance. *Nature*, 501, 416-420.
- EWELS PA, P. A., FILLINGER S, PATEL H, ALNEBERG J, WILM A, GARCIA MU, DI TOMMASO P, NAHNSEN S. 2020. The nf-core framework for community-curated bioinformatics pipelines. *NAt Biotechnol*.
- FESCHOTTE, C. 2008. Transposable elements and the evolution of regulatory networks. *Nat Rev Genet*, 9, 397-405.
- FLATT, T. & PARTRIDGE, L. 2018. Horizons in the evolution of aging. *BMC Biol*, 16, 93.
- GORBUNOVA, V., SELUANOV, A., MAO, Z. & HINE, C. 2007. Changes in DNA repair during aging. *Nucleic Acids Res*, 35, 7466-74.
- HAUSSER, J., MAYO, A., KEREN, L. & ALON, U. 2019. Central dogma rates and the trade-off between precision and economy in gene expression. *Nat Commun*, 10, 68.
- HERRERO, C., MARQUÉS, L., LLOBERAS, J. & CELADA, A. 2001. IFN- γ -dependent transcription of MHC class II IA is impaired in macrophages from aged mice. *Journal of Clinical Investigation*, 107, 485-493.
- HEYN, P., KALINKA, A. T., TOMANCAK, P. & NEUGEBAUER, K. M. 2015. Introns and gene expression: cellular constraints, transcriptional regulation, and evolutionary consequences. *Bioessays*, 37, 148-54.

Chapter 2 - References

- HWANG, Y.-J. & KIM, M. J. 2025. Emerging Role of the DREAM Complex in Cancer and Therapeutic Opportunities. *International Journal of Molecular Sciences*, 26, 322.
- IRISH, S. D., KIMBERLEY, A., IMMLER, S., MOORAD, J. & MAKLAKOV, A. A. 2025. Mutation accumulation underpins evolution of lifespan extension by dietary restriction. Cold Spring Harbor Laboratory.
- JIA, Y., WANG, Y. & XIE, J. 2015. The Hedgehog pathway: role in cell differentiation, polarity and proliferation. *Arch Toxicol*, 89, 179-91.
- JING, J., WU, Z., WANG, J., LUO, G., LIN, H., FAN, Y. & ZHOU, C. 2023. Hedgehog signaling in tissue homeostasis, cancers and targeted therapies. *Signal Transduction and Targeted Therapy*, 8.
- JONKERS, I., KWAK, H. & LIS, J. T. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife*, 3, e02407.
- KAZAZIAN, H. H., JR. & GOODIER, J. L. 2002. LINE drive. retrotransposition and genome instability. *Cell*, 110, 277-80.
- KIRKWOOD, T. B. L. 1977. Evolution of ageing. *Nature*, 270, 301-304.
- KIRKWOOD, T. B. L. & HOLLIDAY, R. 1979. The evolution of ageing and longevity. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205, 531-546.
- LAMMING, D. W. 2016. Inhibition of the Mechanistic Target of Rapamycin (mTOR)–Rapamycin and Beyond. *Cold Spring Harbor Perspectives in Medicine*, 6, a025924.
- LEVIN, H. L. & MORAN, J. V. 2011. Dynamic interactions between transposable elements and their hosts. *Nat Rev Genet*, 12, 615-27.

Chapter 2 - References

LIBERZON, A., BIRGER, C., THORVALDSDÓTTIR, H., GHANDI, M., MESIROV, J. P. & TAMAYO, P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*, 1, 417-425.

LIBERZON, A., SUBRAMANIAN, A., PINCHBACK, R., THORVALDSDÓTTIR, H., TAMAYO, P. & MESIROV, J. P. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*, 27, 1739-40.

MAKLAKOV, A., ALEXEI & IMMLER, S. 2016. The Expensive Germline and the Evolution of Ageing. *Current Biology*, 26, R577-R586.

MALE, I., OZACAR, A. T., FAGAN, R. R., LORING, M. D., SHEN, M.-C., PACE, V. A., DEVINE, C. A., LAWSON, G. E., LUTSERVITZ, A. & KARLSTROM, R. O. 2020. Hedgehog Signaling Regulates Neurogenesis in the Larval and Adult Zebrafish Hypothalamus. *eneuro*, 7, ENEURO.0226-20.

MOORE, L., CAGAN, A., COORENS, T. H. H., NEVILLE, M. D. C., SANGHVI, R., SANDERS, M. A., OLIVER, T. R. W., LEONGAMORNLEERT, D., ELLIS, P., NOORANI, A., MITCHELL, T. J., BUTLER, T. M., HOOKS, Y., WARREN, A. Y., JORGENSEN, M., DAWSON, K. J., MENZIES, A., O'NEILL, L., LATIMER, C., TENG, M., VAN BOXTEL, R., IACOBUZIO-DONAHUE, C. A., MARTINCORENA, I., HEER, R., CAMPBELL, P. J., FITZGERALD, R. C., STRATTON, M. R. & RAHBARI, R. 2021. The mutational landscape of human somatic and germline cells. *Nature*, 597, 381-386.

MORRISON, C. G. 2021. Primary cilia and the DNA damage response: linking a cellular antenna and nuclear signals. *Biochem Soc Trans*, 49, 829-841.

MOSES, E., ATLAN, T., SUN, X., FRANEK, R., SIDDIQUI, A., MARINOV, G. K., SHIFMAN, S., ZUCKER, D. M., ORON-GOTTESMAN, A., GREENLEAF, W. J., COHEN, E., RAM, O. & HAREL, I. 2023. The germline regulates longevity and somatic repair in a sex-specific manner. Cold Spring Harbor Laboratory.

Chapter 2 - References

- PAI, A. A., HENRIQUES, T., MCCUE, K., BURKHOLDER, A., ADELMAN, K. & BURGE, C. B. 2017. The kinetics of pre-mRNA splicing in the *Drosophila* genome and the influence of gene architecture. *Elife*, 6.
- PALLA, A. R., HILGENDORF, K. I., YANG, A. V., KERR, J. P., HINKEN, A. C., DEMETER, J., KRAFT, P., MOONEY, N. A., YUCEL, N., BURNS, D. M., WANG, Y. X., JACKSON, P. K. & BLAU, H. M. 2022. Primary cilia on muscle stem cells are critical to maintain regenerative capacity and are lost during aging. *Nature Communications*, 13.
- PATRO, R., DUGGAL, G., LOVE, M. I., IRIZARRY, R. A. & KINGSFORD, C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*, 14, 417-419.
- RAJAN, A. M., ROSIN, N. L., LABIT, E., BIERNASKIE, J., LIAO, S. & HUANG, P. 2023. Single-cell analysis reveals distinct fibroblast plasticity during tenocyte regeneration in zebrafish. *Sci Adv*, 9, eadi5771.
- REBOLLO, R., ROMANISH, M. T. & MAGER, D. L. 2012. Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet*, 46, 21-42.
- ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2009. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 139-140.
- RODRIGUES, H. F., SOUZA, T. A., GHIRALDINI, F. G., MELLO, M. L. & MORAES, A. S. 2014. Increased age is associated with epigenetic and structural changes in chromatin from neuronal nuclei. *J Cell Biochem*, 115, 659-65.
- SADASIVAM, S. & DECAPRIO, J. A. 2013. The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nat Rev Cancer*, 13, 585-95.

Chapter 2 - References

- SCHIAFFINO, S. & REGGIANI, C. 2011. Fiber types in mammalian skeletal muscles. *Physiol Rev*, 91, 1447-531.
- SCHMIDT, A., ALLMANN, S., SCHWARZENBACH, C., SNYDER, P., CHEN, J.-X., NAGEL, G., SCHÖNEIS, A., RASENBERGER, B., BELI, P., LOEWER, A., HOFMANN, G., THOMAS, TOMICIC, T., MAJA & CHRISTMANN, M. 2024. The p21CIP1-CDK4-DREAM axis is a master regulator of genotoxic stress-induced cellular senescence. *Nucleic Acids Research*, 52, 6945-6963.
- SCHUMACHER, B., POTHOF, J., VIJG, J. & HOEIJMAKERS, J. H. J. 2021. The central role of DNA damage in the ageing process. *Nature*, 592, 695-703.
- SINNING, J., FUNK, N. D., SOERENSEN-ZENDER, I., WULFMEYER, V. C., LIAO, C. M., HALLER, H., HINZE, C., SCHMIDT-OTT, K. M., MELK, A. & SCHMITT, R. 2023. The aging kidney is characterized by tubuloinflammation, a phenotype associated with MHC-II gene expression. *Frontiers in Immunology*, 14.
- SLOTKIN, R. K. & MARTIENSSEN, R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*, 8, 272-85.
- SUN, M. R., CHUNG, H. M., MATSUKI, V., FINK, D. M., STEBBINS, M. J., PALECEK, S. P., SHUSTA, E. V. & LIPINSKI, R. J. 2020. Sonic Hedgehog Signaling in Cranial Neural Crest Cells Regulates Microvascular Morphogenesis in Facial Development. *Frontiers in Cell and Developmental Biology*, 8.
- TESCHENDORFF, A. E., WEST, J. & BECK, S. 2013. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Human Molecular Genetics*, 22, R7-R15.
- ULGEN, E., OZISIK, O. & SEZERMAN, O. U. 2019. pathfindR: An R Package for Comprehensive Identification of Enriched Pathways in Omics Data Through Active Subnetworks. *Frontiers in Genetics*, 10.

Chapter 2 - References

WANG, L. & LIU, X. 2022. Comprehensive Analysis of the Expression and Prognosis for the DREAM Complex in Human Cancers. *Frontiers in Genetics*, 13.

CHAPTER THREE

Defining the determinants of a robust epigenetic clock

Jean-Charles de Coriolis¹, Neil A. Robertson^{2,3}, Nelly Olova², Raheleh Rahabari⁴,
Tamir Chandra^{2,3}, Alexei A. Maklakov¹, Simone Immler¹

¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, NR4 6TX, Norwich, UK

² MRC Human Genetics Unit, Institute of Genetics and Cancer. The University of Edinburgh, Edinburgh, UK

³ MAYO Clinic, Rochester, Minnesota, USA

⁴ Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK



3. Chapter 3

3.1. Abstract

Epigenetic clocks are widely used quantifiers of biological ageing that estimate age from coordinated DNA methylation changes across the genome. Although increasingly applied across tissues and species, the methodological choices in feature definition and aggregation shape clock performance, but this remains poorly understood. Here, we systematically benchmark over 4,000 epigenetic clock models constructed from reduced representation bisulphite sequencing data across brain, caudal fin, intestine, and pooled multi-tissue datasets in zebrafish. The aim was to quantify how methodological choices influence clock accuracy, stability and biological interpretability, rather than defining a single “best” clock. Models span three CpG aggregation strategies (Point, Tile, and Cluster), multiple regression frameworks, and feature-selection thresholds. In single tissues, Tile-based clocks consistently outperform both Point- and Cluster-based models, particularly in brain and intestine, with whole-region tiling outperforming CpG-island-restricted approaches. Feature-scaling analyses reveal tissue-specific limits on informative CpG number, indicating diminishing returns beyond defined CpG budgets. Mixed-effects modelling identifies CpG aggregation strategy and feature filtering as the dominant determinants of performance. By contrast, multi-tissue clocks show reduced accuracy and marked instability across tissues. In pooled multi-tissue models, R^2 -based CpG pre-filtering degrades performance, and most pooled models fall into low-performing or unstable rank classes. Cross-tissue stability analyses demonstrate that CpGs optimised for single-tissue prediction rarely generalise when tissues are pooled, revealing a fundamental trade-off between tissue-specific precision and cross-tissue robustness.

Together, these results show that epigenetic clock performance is driven more by how methylation features are defined and aggregated than by model architecture, and that clocks require scale- and tissue-aware design principles rather than universal feature sets.

3.2. Introduction

Ageing is a progressive, systemic process characterised by the gradual loss of cellular and physiological integrity. It arises from the cumulative impact of molecular damage, stochastic events and epigenetic drift, leading to functional decline and increased susceptibility to disease and death (López-Otín et al., 2013). Epigenetic changes, particularly those involving DNA methylation (DNAm), have attracted attention as measurable and dynamic indicators of biological ageing (Horvath, 2013, Hannum et al., 2013), providing a molecular substrate for developing quantitative models of ageing rate – epigenetic clocks.

Epigenetic clocks, as introduced in Section 1.4 of the Introduction of this thesis, are statistical models that estimate biological age from DNAm data. They typically rely on a subset of CpG sites whose methylation levels correlate predictably with chronological age (Horvath, 2013, Hannum et al., 2013). The construction of an epigenetic clock involves three major steps: (i) preprocessing and normalisation of DNAm data; (ii) identification or selection of CpGs associated with age; and (iii) model fitting using regression or machine learning techniques. Model-fitting approaches used in epigenetic clock construction include penalised linear regression methods, as well as non-linear and machine-learning frameworks including random forests, support vector machines, and neural networks.

Amongst these, penalised regression models are the most commonly applied. Lasso regression (L1 regularisation) reduces model complexity by shrinking the coefficients of many CpG sites exactly to zero, thereby

excluding them from age prediction (Horvath, 2013; Hannum 2013). Elastic Net regression extends this approach by combining L1 and L2 regularisation, improving coefficient stability in the presence of strong correlations among CpG sites, a characteristic feature of high-dimensional DNAm data in which the number of CpGs greatly exceeds the number of samples (Zou and Hastie, 2005). The resulting “epigenetic age” provides a continuous measure of molecular ageing, and its deviation from chronological age, termed age acceleration, has been linked to mortality, disease risk, and exposure to physiological or environmental stressors (Levine et al., 2018; Lu et al., 2019; Hillary et al., 2020).

DNAm involves the enzymatic addition of a methyl group to cytosine residues, predominantly at CpG dinucleotides. In vertebrates, DNAm regulates gene expression, chromatin structure and genome stability (Deaton and Bird, 2011, Jones et al., 2015). During ageing, many vertebrate species, particularly mammals, exhibit a characteristic pattern of DNAm change involving a gradual global hypomethylation alongside site-specific hypermethylation at discrete regulatory regions, including promoters and enhancers (Hernandez-Meza et al., 2021). This pattern has been most consistently documented in humans and other mammals, while evidence from non-mammalian vertebrates suggests greater tissue and species-specific variability (Rakyan et al., 2010, Field et al., 2018). Global hypomethylation is thought to arise from imperfect maintenance of methylation during DNA replication and cumulative epigenetic drift (Field et al., 2018), whereas localised hypermethylation at regulatory elements may reflect targeted chromatin remodelling or selective constraints acting on gene regulatory networks during ageing (Rakyan et al., 2010, Hernandez-Meza et al., 2021).

The directionality and localisation of these changes make DNAm a suitable metric for modelling chronological age. However, methylation patterns

associated with age are not uniformly distributed. Age-related CpGs are often enriched near developmental genes and regulatory elements (Rakyan et al., 2010, Zhou et al., 2022), suggesting that methylation drift may involve partial reactivation or repression of key transcriptional networks. These observations lend biological plausibility to the use of DNAm as a molecular clock but also highlight that statistical correlations may not always correspond to causal mechanisms (Jiang and Guo, 2020, Aquino et al., 2018).

3.2.1. Approaches to clock construction

Although the goal of most of these epigenetic clocks is to predict age from DNAm data, approaches differ in how DNAm is represented across the genome. Across the literature, three main, broad CpG aggregation frameworks are used, here termed Point, Tile and Cluster clocks, each offering a different balance between resolution, noise tolerance and interpretability.

Point clocks use the methylation levels of individual CpG sites as predictors (Li et al., 2022, Hannum et al., 2013), where each site acts as a discrete variable in a penalised regression model. This approach captures fine-scale variation and often yields high predictive accuracy when training large datasets (Li et al., 2022, Hannum et al., 2013, Lehne et al., 2015, Zhang et al., 2019). However, because single-CpG measurements can be affected by local biological variability, batch effects, or population and tissue-specific differences, Point clocks may exhibit sensitivity to dataset composition. For example, CpGs selected as predictive in one training dataset may not generalise to other datasets from different tissues or technical approaches (Li et al., 2022, Lehne et al., 2015, Zhang et al., 2019).

Chapter 3 - Introduction

Tile clocks aggregate CpG methylation values over fixed genomic windows, typically ranging from ~ 300-500bp, randomly spread across the genome. Averaging within these defined tiles provides a regional estimate of methylation, reducing stochastic noise at individual CpGs and potentially capturing high-order regulatory states (Chrusciel et al., 2022, Stubbs et al., 2017, Meer et al., 2018, Higgins-Chen et al., 2022). By integrating multiple neighbouring sites, Tile-based aggregation can improve model stability relative to Point clocks, but it can also complicate biological interpretation, because window size and spacing are arbitrarily defined and influence how local chromatin and regulatory context are represented.

Like Tile clocks, Cluster clocks aggregate CpGs to reduce stochastic noise and stabilise age predictions (Ester et al., 1996, Simpson et al., 2023). However, instead of averaging CpGs within fixed genomic windows, Cluster clocks define groups based on data-driven clustering algorithms, such as DBSCAN, which aggregate CpGs with correlated methylation patterns across samples. The mean methylation of each cluster is then used as a predictor in the age model (Simpson et al., 2023, Ester et al., 1996). These clusters often capture co-methylation modules that correspond to regulatory domains or chromatin regions, providing greater biological coherence than Point clocks while reducing sensitivity to single-site variability. Cluster clocks, therefore, aim to balance stability and interpretability, although aggregation can sometimes reduce raw predictive accuracy when clusters encompass heterogeneous methylation behaviour.

Together, these approaches reflect a continuum between fine-grained statistical precision and biologically informed generalisation.

3.2.2. CpG pre-selection and randomness

A common feature of clock construction is CpG pre-selection. Typically, only CpGs showing significant association with age or sufficient variability are retained prior to model training. Feature filtering reduces dimensionality and computational burden but introduces bias and limits interpretability, as selected CpGs may reflect noise, tissue composition or cohort-specific effects rather than universal ageing signals.

Higgins Chen et al. (2022) and Mathur et al. (2024) demonstrated that models trained on randomly chosen CpG subsets can perform similarly to canonical clocks, yielding comparable mean absolute error (MAE) and R^2 (Mathur et al., 2024; Higgins-Chen et al., 2022). This suggests that predictive power derives from the global correlation structure of the methylome rather than the biological importance of individual loci. Consequently, while clocks remain useful biomarkers, their constituent CpGs should not automatically be interpreted as mechanistically or functionally important in ageing, since their inclusion may reflect correlated methylation patterns rather than causal roles in biological processes.

3.2.3. Tissue scope: single tissue and multi-tissue models

Epigenetic clocks can be developed at different biological scales, ranging from single-tissue models that capture local ageing trajectories to multi-tissue clocks that reflect systemic ageing across whole individuals. Single-tissue epigenetic clocks are trained on DNAm data from a specific tissue or cell type, capturing local ageing trajectories (Meer et al., 2018). These models often achieve high accuracy within the same tissue but generalise poorly to others due to differences in methylation dynamics across tissues (Apsley et al., 2024). By contrast, multi-tissue clocks are trained across diverse tissues to identify CpG sites whose methylation changes are consistent across the

whole organisms or a group of organisms (Horvath, 2013, Apsley et al., 2024). These models are more versatile but typically display reduced precision, as they must accommodate heterogeneous tissue-specific ageing patterns (Horvath, 2013, Levine et al., 2018, McEwen et al., 2020). The distinction between these two approaches reflects the broader balance between specificity and generality in biological modelling. Understanding how tissue scope influences clock performance provides insight into whether DNAm age represents systemic or tissue-specific ageing processes.

3.2.4. Model choice

The choice of modelling approach for an epigenetic clock is closely related to the tissue scope of the data. Single-tissue clocks generally involve more homogeneous DNAm patterns, whereas multi-tissue clocks must accommodate correlated CpG dynamics across diverse tissues, influencing both feature selection and model regularisation. Most epigenetic clocks employ regularised regression approaches, predominantly Lasso (L1) or Elastic Net (L1 + L2) regression, to simultaneously perform feature selection and prevent overfitting (Zou and Hastie, 2005). Lasso enforces sparsity, producing compact models, but performs poorly when predictors are highly correlated (Hannum et al., 2013). Elastic net models, on the other hand, are advantageous because they handle groups of correlated predictors by allowing them to be selected together, whilst still maintaining a sparse model structure (Horvath, 2013, Friedman et al., 2010, Zou and Hastie, 2005). Hyperparameters (α for mixing ratio and λ for penalty strength) are typically optimised via cross-validation. In some cases, alternative machine learning approaches, such as gradient boosting machines (Li et al., 2018), support vector machines (Fan et al., 2022) or neural networks (Martínez-Enguita et al., 2025), are used for epigenetic clock construction, but penalised regression

remains the dominant choice due to its interpretability, variable selection properties and stability in high-dimensional genomic data.

3.2.5. Balancing predictive accuracy and biological meaning

A central challenge in epigenetic clock research lies in reconciling predictive accuracy with biological interpretability. High R^2 values indicate strong correlation with chronological age but may result from statistical overfitting or latent confounding (Dabrowski et al., 2023). Conversely, metrics such as MAE, which quantify absolute prediction error, may better reflect biologically relevant variation in ageing rate, but clocks optimised for interpretability or functional coherence often exhibit higher MAE and underperform on predictive benchmarks (Higgins-Chen et al., 2022). Few studies explicitly consider this trade-off. Most report predictive statistics without evaluating how methodological choices, such as feature aggregation, pre-selection or model regularisation, affect biological inference. As a result, there is a lack of consensus on what impacts the construction of robust clocks, whether one that predicts age with minimal error or one that reflects biological coherent processes. In practice, these goals are difficult to achieve simultaneously. Epigenetic clocks can be very good statistical models or very good biological representations, but rarely both. Recognising this distinction is essential for interpreting clock outputs and improving their utility as biomarkers for biological ageing.

Despite the growing number of published epigenetic clocks, the methodological reasoning behind their construction is also often under-described. Decisions about feature definition, CpG aggregation and regression strategy are rarely justified, leaving open questions about why certain steps are performed and how they influence model performance. The present study attempts to address this gap by comparing multiple strategies

Chapter 3 - Material and Methods

for building epigenetic clocks, including Point, Tile and Cluster aggregation of CpG sites, across single and multi-tissue contexts in the model organism, zebrafish (*Danio rerio*). By using RRBS, we evaluated how data representation, CpG selection and modelling approaches influenced both predictive accuracy and biological interpretability. By examining these factors systematically, our study aims to clarify the methodological foundations of clock construction and to promote a more transparent understanding of how methylation data encode biological time.

3.3. Material and Methods

3.3.1. Fish Lines and Maintenance

Wild-type zebrafish (AB strain) were obtained from the Zebrafish International Resource Centre (ZIRC, University of Oregon, Eugene, USA) and maintained under standard laboratory conditions in the Controlled Environment Facility, University of East Anglia, UK. All experimental fish used in this study were the product of the facility's strict outbreeding programme to preserve genetic diversity.

Fish were housed in 3L tanks under a 12: 12 light: dark photoperiod at 28°C. Adult fish were fed three times daily with live *Artemia* (ZM System) and dried food (Sparos, Adult zebrafish diet, nutrition in aquaculture). Juvenile fish were fed four times daily to support growth and maintain consistent nutritional conditions across developmental stages.

3.3.2. Sample Collection

Fish were euthanised using Aquacalm (Syndel) in accordance with institutional and UK Home Office guidelines. Samples of three somatic

Chapter 3 - Material and Methods

tissues were collected from 70 individual fish, brain, caudal fin and intestine, yielding a total of 196 samples (67 brain, 65 fins and 64 intestine). Fish ages spanned eleven groups from 3 to 22 months (Table S10). To account for increased variability in methylation at younger ages, sampling was more intensive in early life stages.

3.3.3. DNA Extraction

Genomic DNA was extracted individually from each tissue sample using a modified Qiagen DNeasy Blood and Tissue Kit protocol (Qiagen, 69506). Samples were digested in proteinase K at 56 °C for 1-2 h with intermittent vortexing until complete lysis was achieved. Lysates were processed through silica spin columns according to the manufacturer's instructions, and DNA was eluted in 30 µL EB buffer. DNA concentrations and purity were evaluated using a Qubit fluorometer and a NanoDrop spectrophotometer prior to library preparation.

3.3.4. Library Preparation

A total of 196 RRBS libraries were prepared in two independent batches. Batch 1 (n = 99) was processed at the University of Edinburgh, and Batch 2 (n=97) at the University of East Anglia, using the TECAN Ovation® RRBS methyl-Seq System Kit (TECAN, 9522-A01), following an optimised version of the manufacturer's protocol (see detailed procedure at <https://www.protocols.io/view/ovation-rrbs-methyl-seq-library-preparation-protocol-kxygk3kydg8j/v1>).

Briefly, genomic DNA was digested with *MspI*, ligated to methylated adapters, end-repaired, and subjected to bisulphite conversion. Libraries were then purified via magnetic beads, PCR amplified, and size selected.

Chapter 3 - Material and Methods

Fragment size distribution and library quality were assessed using an Agilent TapeStation, and final quantification was performed prior to equimolar pooling.

3.3.5. RRBS sequencing

Batch 1 libraries were sequenced on an Illumina NovaSeq 6000 platform using an S2 flow cell across six lanes, 2 x 150 bp paired-end reads generating approximately 4 billion reads per lane. Batch 2 libraries were sequenced on an Illumina NovaSeq X platform using a 10B flow cell over three lanes, producing 2 x 150 bp paired-end reads, generating approximately 2.5 billion reads per lane. All sequencing was performed at the Wellcome Sanger Institute, Cambridge, UK. Raw data were demultiplexed and delivered as gzipped FASTQ files.

3.3.6. Data Trimming, Alignment and Methylation Extraction

Raw reads were processed using the NuGEN Technologies pipeline (<https://github.com/nugentechnologies/NuMetRRBS>). Adapter and low-quality bases were trimmed with Trim Galore v 0.6.5 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and diversity adaptor artefacts were removed using a NuGEN-provided Python script (<https://github.com/nugentechnologies/NuMetRRBS/trimRRBSdiversityAdaptCustomers.py>). Cleaned reads were aligned to the zebrafish reference genome (GRCz11; https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_000002035.4/), using Bismark v 0.25.1 (Krueger and Andrews, 2011), and duplicate reads were

Chapter 3 - Material and Methods

identified and marked with the NuDup tool in NuGEN (<https://github.com/tecangenomics/nudup>).

Methylation calling was performed using the Bismark methylation extractor (Krueger and Andrews, 2011), generating cytosine methylation reports in “.cov” format. Methylation metrics (methylated and unmethylated counts per CpG) were used as input for downstream analyses.

3.3.7. Methylation data pre-processing

Methylation data were imported into R v 4.3.1 (R Core Team, 2022) using the methylKit package v 1.3 (Akalin et al., 2012). Coverage files were read using the “methRead()” function, applying a minimum per-CpG coverage threshold of five reads (*mincov* =5). CpG sites were retained only if present in at least 90% of samples within each tissue (*min_samples* <- *as.interger(ceiling(n_samples * 0.9))*).

Methylation percentage was calculated as methylated reads / total reads x 100. No transformation was applied to these percentages. The resulting tissue-specific methylation matrices formed the basis for subsequent clock training across all analytical configurations.

3.3.8. Epigenetic Clock Model Construction and Evaluation

To assess how analytical design influences predictive performance, epigenetic clocks were constructed under a factorial framework combining multiple pre-processing, filtering and modelling strategies. Each model estimated epigenetic age from DNA methylation data derived from either individual CpG (Point), aggregated genomic windows (Tile), or correlated CpGs (Cluster).

Chapter 3 - Material and Methods

Five major methodological aspects were systematically changed: (I) Feature inclusion thresholds involved minimal sample coverage per CpG / CpG island of 0 (no filter), ≥ 5 , ≥ 10 or ≥ 20 CpGs. (II) Age-association filtering where a minimum univariate CpG-age correlation (R^2) was performed at 0 (no filter), ≥ 0.05 , ≥ 0.1 , ≥ 0.2 , or ≥ 0.3 . (III) Feature aggregation at three genomic scales, namely Point, Tile and Cluster. (IV) Regression model, either Lasso or Elastic net models. (V) Window size (Epsilon; Epse) of CpG site/island was set to 12 different sizes (bp), 1 (point), 50, 100, 500, 750, 1000, 2000, 3000, 5000, 22500, 25000.

This full-factorial design produced a comprehensive model space incorporating multiple features and parameter combinations, allowing systematic evaluation of robustness, interpretability and predictive power of the different clock models (Table S11). The top selected CpGs for each tissue can be found in Table S12.

3.3.9. Epigenetic clock model training and selection

The complete model space included 3 x CpG aggregation groups, 2 x regression models, 5 x R^2 Thresholds, 4 x Feature filters, 12 x Epses (Tile windows) and 4 x Tissues / Tissue-groupings (Table S11), resulting in 4,320 distinct models. For each model, the full set of samples was randomly divided into a training set (80% of individuals) and a held-out validation set (20%), with the training set used for model fitting, optimisation of CpG selection, and hyperparameter tuning, and the held-out validation set for unbiased evaluation of predictive performance.

Model optimisation was performed using a nested iterative procedure comprising 100 outer iterations and 50 inner iterations per outer iteration. Each outer iteration corresponds to a new random division of the samples

into training (80%) and held-out validation (20%) sets, capturing variability in the training sample composition. Within each outer iteration, 50 inner iterations were performed in which random subsets of candidate CpGs were selected from the available pool and used to fit the regression model. Each inner iteration produced predicted ages and associated performance metrics (R^2 and MAE), which were standardised across inner iterations. A combined score integrating scaled R^2 , inverted MAE, and a penalisation for models with large numbers of CpGs was used to rank inner iterations and identify the best-performing set of CpGs for that outer iteration.

After all 100 outer iterations were completed, the single best-performing outer iteration, representing both the most robust training/validation split and the optimised subset of CpGs, was selected for final model construction. A final epigenetic clock was then refitted using the full training set and only the selected CpGs from this best outer iteration. Predicted ages for the held-out validation set were computed using this final model to assess its generalisability. For each model, the number and identity of both candidate and selected CpGs, the fitted cross-validation object, and predicted ages were kept for further quantification. This procedure was repeated independently across all tissues and analytical conditions, enabling quantitative comparison of predictive performance across the full parameter space.

3.3.10. Technical parameters for epigenetic clock construction

For point clocks, each CpG site was treated as an independent predictor, with all available after filtering CpGs retained across samples. Tile clocks summarised DNAm across genomic windows of varying size (50 bp–25 kb), with tiles placed using two strategies: (1) covering the whole reduced representation region, capturing global methylation patterns without bias,

and (2) confined to CpG-rich regions corresponding to annotated CpG islands. The CpG island approach reduces randomness in tile placement, focusing feature selection on biologically relevant regions, and decreases the total number of tiles, reducing model complexity and computational requirements. CpG island annotations for the *Danio rerio* reference genome GRCz11 were retrieved from the UCSC Genome Browser Table Browser and exported as a BED file using default settings for chromosome coordinates. Cluster clocks grouped CpGs based on correlated DNAm patterns using a DBSCAN-based clustering pipeline, with mean cluster DNAm used as the predictive feature.

Where CpG sites were missing from individual samples, values were imputed using the median methylation level within tissues to preserve matrix completeness. Regularisation strength (λ) was optimised using 10-fold cross-validation to minimise mean squared error.

3.3.11. Analytical Rationale

By systematically varying filtering thresholds, feature scales and regularisation schemes, we characterised the sensitivity of model performance to both statistical and biological sources of variation.

Including both Lasso and Elastic Net regression enabled evaluation of how penalty structure affects feature retention, particularly for spatially correlated CpGs. Comparing Point, Tile and Cluster aggregations provided insight into the genomic window sizes most informative for biological age estimation. Collectively, this framework offers a comprehensive and reproducible basis for assessing the methodological determinants of epigenetic clock behaviour across somatic tissues.

3.4. Results

After sequencing and quality control, 14 samples failed QC, resulting in the generation of a total of 182 tissue samples, comprising 62 brain, 61 caudal fin, and 59 intestine samples, forming the input dataset for all downstream analyses. All samples passing initial sequencing and alignment quality control were retained as candidates for epigenetic clock construction. Subsequent filtering steps, including CpG coverage thresholds, CpG completeness requirements, and feature selection, were applied within each modelling configuration, such that the number of retained CpGs and effective sample-feature matrices varied across models (All analytical steps can be found in the GitHub repository, discussed in section 3.10 of this thesis).

In total, 4,320 distinct epigenetic clock configurations were trained, spanning multiple CpG aggregation strategies, regression frameworks, and filtering parameterisations. Following model training and validation, 3,036 models met performance and quality criteria and were retained for downstream analyses (Table S11). Unless stated otherwise, all results below refer to this model set defined in the methods (3.3.8, 3.3.9 & 3.3.10), with comparisons interpreted across configurations rather than assuming a shared underlying feature space.

3.4.1. CpG selection strategies reveal tissue × method interactions

We evaluated all 3,036 epigenetic clock models across brain, caudal fin, intestine, and multi-tissue (pooled brain + caudal fin + intestine) datasets, testing three CpG aggregation strategies: Point, Tile, and Cluster (Figure 7A). Model performance, summarised as a combined score integrating rescaled R^2 and inverse MAE, differed substantially depending on both tissue and CpG grouping.

Chapter 3 - Results

Across all datasets, Tile clocks consistently outperformed the other two CpG aggregation methods (Figure 7B). A linear mixed-effects model confirmed a highly significant effect of CpG grouping on performance ($\chi^2 = 97$, $df = 2$, $p < 0.0001$; Table S13). Pairwise comparisons revealed that Cluster and Point performed similarly (Cluster vs Point: *Estimate* = -0.03, $z = -0.9$, $p = 0.6$), but both were worse than Tile clocks (Cluster vs Tile: *Estimate* = -0.1, $z = -8.9$, $p < 0.0001$; Point vs Tile: *Estimate* = -0.08, $z = -2.4$, $p = 0.04$). This pattern held when examining single-tissue clocks individually. In brain and intestine, Tile clocks outperformed both Point and Cluster strategies (brain: Cluster vs Tile: *Estimate* = -0.3, $z = -13.4$, $p < 0.0001$; Point vs Tile: *Estimate* = -0.2, $z = -3.4$, $p = 0.002$; intestine: Cluster vs Tile: *Estimate* = -0.2, $z = -8.5$, $p < 0.0001$) (Figure S8). In multi-tissue clocks, however, Cluster aggregation slightly outperformed Tile clocks (Cluster vs Tile: *Estimate* = 0.1, $z = 4.7$, $p < 0.0001$), highlighting a tissue $\times$ CpG grouping interaction (Figure S8).

These results indicate that while Tile clocks are generally superior for single-tissue models, optimal CpG selection depends on the tissue context. Strategies optimised for individual tissues do not necessarily generalise to multi-tissue datasets, emphasising the need for context-specific optimisation when designing epigenetic clocks.

Chapter 3 - Results

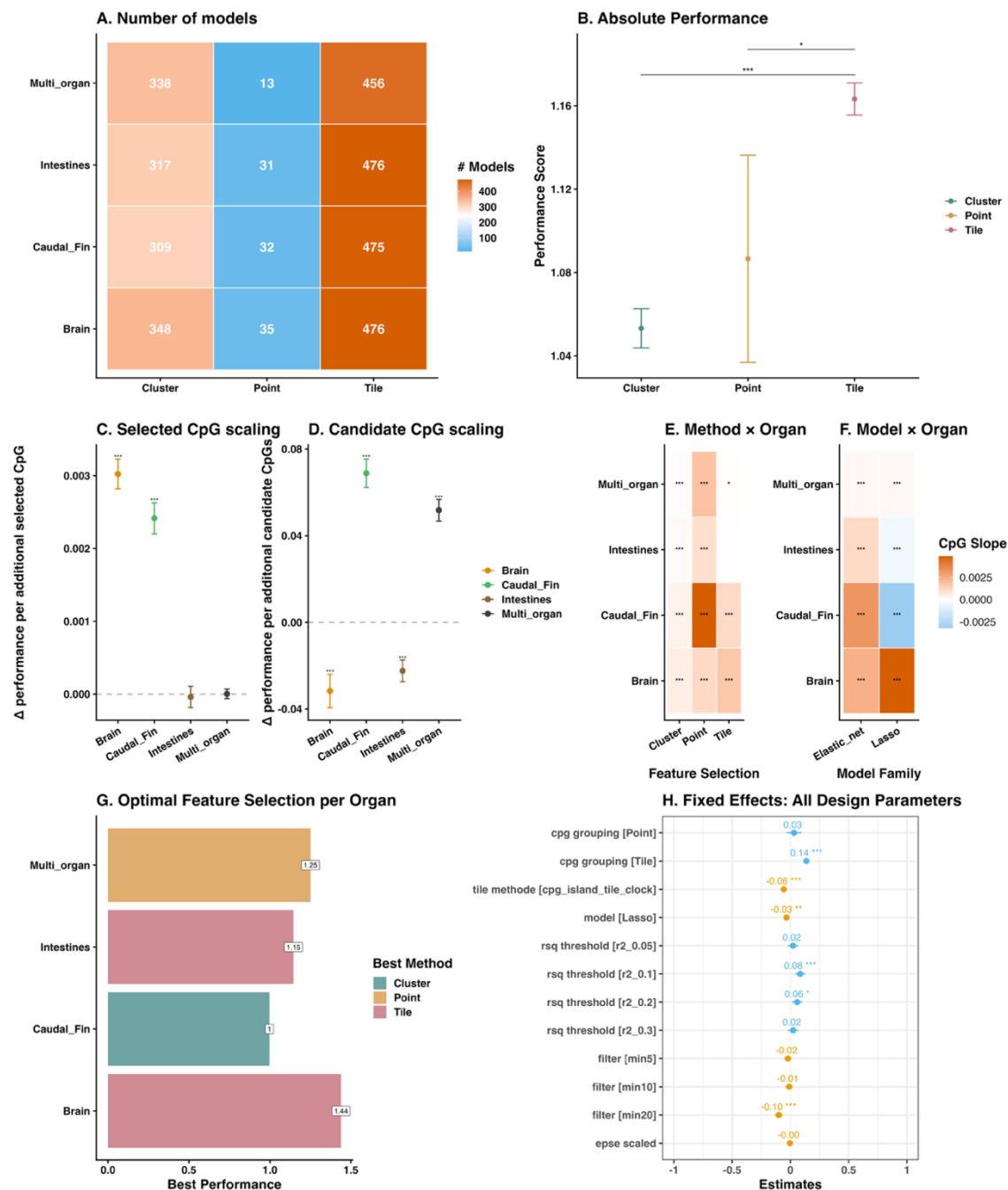


Figure 7: Comprehensive design space mapping reveals tissue x method interactions and CpG scaling limits.

A) Heatmap of model counts across CpG grouping strategies (Point, Tile, Cluster) and tissues (brain, caudal fin, intestine, multi-tissue). **B)** Performance scores (rescaled R^2 + inverse MAE) by CpG grouping, faceted by tissue, showing Tile dominance over Point and Cluster methods. **(C-D)** Tissue-specific scaling of performance with selected **C)** and candidate **D)** CpG counts. **(E-F)** Heatmaps of performance scaling slopes (performance ~ selected CpG count) for tissue × CpG grouping **E)** and regression models, Lasso and Elastic Net **F)**. **G)** Optimal CpG grouping per tissue (Tile: brain/intestine; Cluster: fin; Point: multi-tissue). **H)** Mixed-effects model coefficients ranking all design parameters (CpG grouping > tiling > Elastic Net > R^2 filtering). Significance is indicated with asterisks, with * = $p < 0.05$, ** $p < 0.001$ & *** $p < 0.0001$. Point indicate the mean and error bars represent standard errors.

3.4.2. CpG dosage effects expose tissue-specific scaling limits and overfitting risks

We next examined how the number of selected CpGs, influenced clock performance across tissues (Figure 7C, Table S14). Overall, performance scaled heterogeneously with selected CpG count. In brain (LMM: *Estimate* = 0.003, $z = 29.2$, $p < 0.0001$) and caudal fin (*Estimate* = 0.002, $z = 22$, $p < 0.0001$), additional selected CpGs improved predictive accuracy, indicating these tissues benefit from higher CpG complexity (Table S14). By contrast, intestine and multi-tissue clocks showed no significant scaling with selected CpGs, suggesting these systems operate in a lower-complexity regime where adding more features provides minimal marginal gain (Figure 7C).

Examining the number of candidate CpGs imputed into the regression model further highlighted tissue-specific dynamics (Figure 7D; Table S13). Increasing the candidate pool available for the regression model improved performance for multi-tissue (*Estimate* = 0.05, $z = 20.2$, $p < 0.0001$) and caudal fin clocks (*Estimate* = 0.07, $z = 20.7$, $p < 0.0001$), but paradoxically reduced performance in brain (*Estimate* = -0.03, $z = -8$, $p < 0.0001$) and intestine (*Estimate* = -0.002, $z = -8.7$, $p < 0.0001$), indicating that excessive candidate expansion can introduce noise and overfitting in some tissues.

Tissue $\times$ CpG grouping heatmaps revealed that positive scaling with selected CpGs was near-universal across combinations, with the exception of Tile clocks in intestine, which showed no significant effect (LMM: *Estimate* = 1.2, $z = 1$, $p = 1$; Figure 7E). Model-specific effects were also observed: Lasso models displayed negative scaling in caudal fin (*Estimate* = -0.004, $z = -15.6$, $p < 0.0001$) and intestine clocks (*Estimate* = -0.0004, $z = -4.4$, $p = 0.01$), whereas all other tissue-model combinations showed improved performance with increasing selected CpGs (Figure 7E, 7F).

Chapter 3 - Results

Taken together, these patterns define tissue-specific “CpG budgets”, described here as the minimum number of CpGs required to achieve near-maximal predictive performance within a given tissue. Median performance was achieved with 65 CpGs (IQR: 42–94), and models with ~50–100 CpGs captured 95% of maximum performance, highlighting diminishing returns and overfitting risks when expanding the candidate pool excessively.

3.4.3. Global design hierarchy prioritises feature selection over model complexity

To assess the relative contribution of all analytical choices simultaneously, we fitted a comprehensive linear mixed-effects model incorporating CpG aggregation strategy, tiling method, regression model family, R^2 pre-filtering threshold, feature filtering strategy, and tile size (Epse) (Figure 7H). Among these parameters, CpG aggregation strategy emerged as the dominant determinant of performance. Tile-based clocks significantly outperformed both Point and Cluster approaches ($\chi^2 = 97$, $df = 2$, $p < 0.0001$), and within Tile models, whole-RRBS tiling consistently outperformed CpG-island-restricted tiling ($\chi^2 = 139$, $df = 1$, $p = 0.0002$).

Regression model family had a smaller but significant effect, with Elastic Net models outperforming Lasso models overall ($\chi^2 = 8.6$, $df = 1$, $p = 0.03$), consistent with the improved stability of Elastic Net in the presence of correlated methylation features. Moderate R^2 pre-filtering thresholds (0.1–0.2) improved predictive performance across models ($\chi^2 = 23$, $df = 4$, $p < 0.0001$), whereas stricter (0.05 & 0.3) or no filtering reduced accuracy, reflecting a balance between noise removal and excessive feature exclusion. In contrast, tile size, spanning 50 bp to 25 kb, had no detectable effect on model performance once CpG grouping and filtering were accounted for (χ^2

Chapter 3 - Results

$= 0.4$, $df = 1$, $p = 0.5$), indicating that tiling granularity plays a minor role relative to feature selection strategy.

Tissue-stratified analyses revealed predictable shifts in optimal design choices (Figure 7G). Tile-based clocks were consistently superior in brain and intestine, Cluster clocks performed best in caudal fin, and Point clocks achieved the highest performance in multi-tissue models (Figure 7G). Despite these tissue-specific differences, the global hierarchy of design parameters was preserved across contexts, with CpG grouping and feature filtering exerting substantially larger effects than regression architecture or tiling resolution. Together, these results demonstrate that epigenetic clock performance is driven primarily by how methylation features are defined and filtered, rather than by increased model complexity or finer genomic partitioning.

3.4.4. Multi-tissue clocks decouple metrics and destabilise performance rankings

Across all models, higher predictive concordance (R^2) was associated with lower absolute prediction error (MAE), indicating that models explaining more variance in age tended to make more accurate predictions overall ($t = 7$, $df = 3304$, $p < 0.0001$; Figure 8A). However, this relationship depended strongly on the CpG aggregation strategy (Figure 8A). When stratified by CpG grouping, Cluster-based models showed a significant negative association between R^2 and MAE (LMM: *Estimate* = -0.07 , $z = -1.9$, $p < 0.001$), consistent with coherent improvements across both metrics. In contrast, Point and Tile clocks exhibited positive associations (Point: *Estimate* = 0.5 , $z = 2.8$, $p < 0.0001$; Tile: *Estimate* = 0.2 , $z = 6.5$, $p < 0.0001$), indicating that gains in explained variance were accompanied by increased absolute error in these models. This divergence suggests that metric optimisation behaves

Chapter 3 - Results

differently across CpG grouping strategies and that R^2 and MAE are not universally interchangeable indicators of performance.

The impact of R^2 -based CpG pre-filtering further underscored tissue-specific effects (Figure 8B). Increasing the R^2 threshold significantly improved model performance in brain clocks (LMM: *Estimate* = 0.6, z = 5.9, p < 0.0001 per unit increase in threshold), whereas the same filtering strategy significantly reduced performance in multi-tissue models (Multi-tissue *Estimate* = -0.5, z = -5, p < 0.0001). Caudal fin and intestine clocks showed no significant sensitivity to R^2 thresholding, with performance remaining stable across filter levels (caudal fin: *Estimate* = -0.09, z = -0.7, p = 0.4; intestine: *Estimate* = -0.1, z = -1.2, p = 0.2). This narrow performance range indicates robustness to preprocessing choices in these tissues.

Assessment of cross-tissue stability highlights differences in performance consistency. When models were classified by mean rank and rank variability across tissues (Figure 8C, Table S15), Tile-based clocks were the most consistently high-performing and stable, with 36.5% of models in the “High-performing stable” category and only 14.3% classified as “Low-performing stable.” Cluster-based clocks showed intermediate stability, with 14.4% high-performing stable models and the largest fraction of low-performing models (32–33%). Point-based clocks were the least stable, with 40% of models considered “High-performing unstable” despite only 8.6% being high-performing stable (Table S15). Tissue-level analysis similarly revealed tissue-specific patterns: brain clocks were enriched for high-performing stable models (24.3%), while multi-tissue clocks displayed the greatest instability, with 69.6% falling into the low-performing unstable category (Table S15). These results indicate that both CpG aggregation strategy and tissue context strongly influence model robustness, with Tile aggregation enhancing stability across tissues.

Chapter 3 - Results

Collectively, these analyses demonstrate that both metric coupling (R^2 vs MAE) and the benefits of R^2 -based CpG filtering are context-dependent. They also underscore the importance of explicitly assessing clock stability across tissues and CpG strategies, reinforcing the need for tissue-aware optimisation and multi-model benchmarking when designing epigenetic clocks for multi-tissue applications.

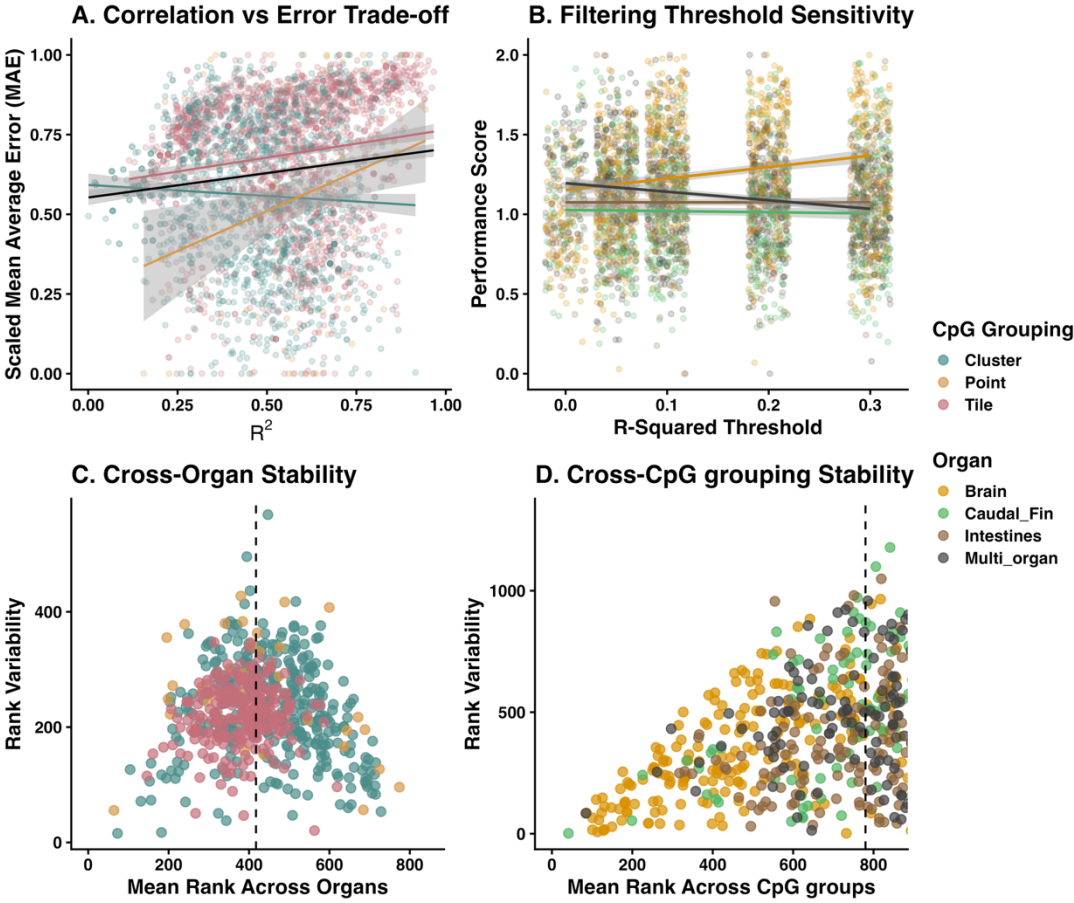


Figure 8: Multitissue clocks decouple metrics and exhibit pervasive rank instability.

(A) Relationship between predictive concordance (R^2) and scaled mean absolute error (MAE) for individual models, coloured by CpG aggregation method (Point, Tile, Cluster). The overall trend is coloured in black. **(B)** Effect of R^2 -based CpG pre-filtering on model performance across tissues. Points show individual models, with smoothing lines indicating tissue-specific trends in sensitivity to thresholding. Brain clocks are strongly affected, while fin and intestine clocks show minimal change. **(C)** Cross-tissue stability of models summarised by CpG aggregation method. Each point represents a model, with mean rank across tissues (x-axis) versus rank variability (y-axis). Colour indicates CpG grouping. Dashed lines show median thresholds for high/low performance and stability. **(D)** Cross-CpG grouping stability across tissues. Points show mean rank versus rank variability per tissue (colour). Dashed lines indicate median thresholds.

3.5. Discussion

Our systematic evaluation of >3,000 cross-sectional epigenetic clock variants identified CpG aggregation strategy as the dominant determinant of predictive performance in tissue-specific settings. Tile-based clocks consistently outperformed both Point- and Cluster-based models, particularly in the brain and the intestine, with whole-region tiling outperforming CpG-island-restricted approaches. These findings extend prior observations that regional methylation summaries outperform single-CpG predictors by reducing stochastic noise and capturing coordinated epigenetic regulation (Higgins-Chen et al., 2022, Chrusciel et al., 2022) and demonstrate that these benefits are strongest when tissue context is preserved.

Most widely used epigenetic clocks, including the Horvath human multi-tissue clock (Horvath, 2013), Hannum’s human blood clock (Hannum et al., 2013), and subsequent mouse (Meer et al., 2018) and pan-mammalian clocks (Lu et al., 2019, Lu et al., 2023), are Point-based, relying on individual CpGs selected for cross-tissue or cross-species stability. This design choice has proven effective for broad applicability but necessarily prioritises robustness over tissue-specific resolution. Our results show that, when tissue homogeneity is maintained, aggregating CpGs into genomic tiles yields substantially higher accuracy, likely by integrating local regulatory context and buffering measurement noise intrinsic to RRBS-based methylation data.

Unexpectedly, Tile clocks also outperformed Cluster-based clocks, despite the latter’s biologically motivated aim of capturing co-regulated CpG modules. This suggests that correlation-driven clustering, particularly when combined with aggressive preprocessing, specifically high feature and R^2 filtering, may remove informative variance in tissues where methylation changes are spatially structured but only weakly correlated across samples.

In such contexts, age information is distributed across neighbouring CpGs rather than concentrated in tightly co-varying modules, rendering correlation-based grouping (Cluster) inefficient. By contrast, moderate preprocessing preserves absolute methylation levels, whilst local genomic coherence retains the necessary spatial signal. Together, these findings define a clear design hierarchy, CpG aggregation followed by moderate R^2 -based filtering and Elastic Net regression, that offers a transparent and reproducible alternative to increasingly opaque clock construction pipelines.

3.5.1. CpG scaling reveals tissue-specific efficiency limits

Scaling analyses revealed that the relationship between clock performance and CpG number is strongly tissue dependent. Brain and caudal fin clocks benefitted from increasing numbers of selected CpGs, whereas intestinal clocks rapidly saturated, achieving high performance with remarkably few features. This establishes a novel concept of tissue-specific CpG budgets, defined here as the minimum number of CpGs required to achieve near-maximal predictive performance, reflecting underlying differences in epigenetic complexity, cellular heterogeneity, and regulatory organisation across tissues.

Notably, the highest-performing intestine clocks required as few as two CpGs, contrasting sharply with brain and caudal fin clocks, which achieved top decile performance with approximately 20–30 CpGs. Low-dimensional ageing signatures have similarly been reported in highly proliferative mammalian tissues, particularly peripheral blood and haematopoietic lineages in humans and mice, where methylation ageing is dominated by a limited set of mitotically linked loci (Hannum et al., 2013, Stubbs et al., 2017). These parallels suggest that epigenetic ageing signatures may be

fundamentally simpler in tissues with high cell turnover, regardless of species.

Conversely, expanding the candidate CpG pool negatively impacted brain and intestine clocks, highlighting overfitting risks associated with excessive feature availability, particularly for Lasso-based models, which are sensitive to correlated predictors. Similar behaviour has been observed in human clock development, where increasing feature counts often improve training performance but reduce generalisation to independent cohorts (Higgins-Chen et al., 2022). Our results formalise this trade-off and show that optimal clocks balance feature richness against tissue-specific signal efficiency rather than maximising CpG inclusion.

3.5.2. Multi-tissue clocks require different design paradigms

One of the most striking results was the reversal of performance rankings between single-tissue and multi-tissue clocks. Strategies optimised for individual tissues, particularly Tile aggregation, systematically underperformed in multi-tissue settings, where Point-based clocks showed superior generalisation. This mirrors the design logic of canonical universal clocks, including Horvath's pan-tissue clock (Horvath, 2013) and the Mammalian Methylation Consortium clocks (Meer et al., 2018, Lu et al., 2019, Lu et al., 2023), which prioritise CpGs with conserved age-associated trajectories across tissues and species.

Across all models, improvements in R^2 were not consistently accompanied by reductions in MAE, and rank variability differed substantially between single- and multi-tissue clocks. Classification of models by cross-tissue stability (Tables S15, Figure 8C) highlighted these differences. Tile-based clocks were enriched for high-performing stable models in single tissues (e.g., brain, 36.5 % high-performing stable), whereas Cluster- and Point-based

clocks exhibited higher instability. In contrast, multi-tissue clocks were overwhelmingly low-performing and unstable (Tables S15). This instability was reflected in the distribution: 5.4% of models were high-performing stable, 8.5% were high-performing unstable, 16.5% low-performing stable, and nearly 70% low-performing unstable. These data indicate that CpGs optimised for single-tissue precision, even if highly predictive locally, rarely generalise across tissues, revealing an inherent trade-off between tissue-specific accuracy and cross-tissue robustness. Similar metric discordance has been noted in human clocks applied outside their original training context (Levine et al., 2018, Belsky et al., 2015).

Importantly, R^2 -based pre-filtering, which improved performance in brain clocks, significantly degraded multi-tissue models. This suggests that CpGs with strong tissue-specific age correlations may actively undermine generalisation when tissues are pooled. Rather than implying that universal clocks are infeasible, these findings clarify why they require fundamentally different design constraints and accept reduced tissue-specific resolution in exchange for robustness. Together, the results argue for a scale-aware framework in which clocks are explicitly designed either for tissue-specific precision or for cross-tissue applicability, but not both simultaneously.

3.5.3. Implications for clock benchmarking and reproducibility

Our analysis reveals a dense landscape of statistically distinct yet similarly performing clocks within the same parameter space. This multiplicity underscores a key limitation of current practice: selecting a single “best” clock risks overinterpretation, as minor perturbations in training data, preprocessing, or feature selection can invert performance rankings. This concern extends to widely used epigenetic clocks, where model selection is frequently based on a single training cohort or limited validation

frameworks. Comparative analyses have demonstrated substantial variability in clock behaviour across tissues, interventions, and biological contexts (Levine et al., 2018, Higgins-Chen et al., 2022). More broadly, ageing biomarkers, including DNAm-based measures, capture partially non-overlapping dimensions of biological ageing. Our results provide a systematic demonstration of this instability at scale, and support calls for multi-model benchmarking and ensemble-based approaches, particularly when clocks are deployed outside their original training context.

Although conducted in zebrafish, the principles identified here, CpG aggregation, tissue-specific feature efficiency, and metric instability under generalisation, align closely with observations in mammalian systems. Cross-species validation remains essential, but the framework presented offers a scalable and interpretable strategy for epigenetic clock construction that avoids ad hoc feature selection.

3.5.4. Conclusion

Epigenetic clocks are powerful, yet inherently sensitive biomarkers shaped by complex interactions between methylome architecture, tissue composition, and modelling choices. This study demonstrates that no single clock construction, whether Point-, Tile-, or Cluster-based, single- or multi-tissue, consistently outperforms other in terms of accuracy, precision, and generalisability. The central insight is that robustness emerges from diversity rather than optimisation. Reliance on a single “best-performing” clock is fragile, as performance rankings are highly contingent on sampling and design parameters. Instead, constructing and benchmarking ensembles of clocks and summarising predictions across models provides more stable, interpretable, and reproducible estimates of biological age. By embracing methodological pluralism rather than seeking universal optima, epigenetic

clocks can evolve from bespoke predictors into robust quantitative tools for ageing biology.

3.6. Data Availability

All sequencing data, including RRBS data, generated for this project are available under accession (under embargo until acceptance for publication), <https://www.ebi.ac.uk/ena/browser/view/PRJEB107117>. All scripts and supplementary data are available at the git repository: <https://github.com/asdfjohnny1/epigenetic-clock-determinants.git>.

3.7. Contributions

The studies were conceptualised and designed by S.I , J.C.D.C., T.C.

Data analysis and visualisation was carried out by J.C.D.C. ,N.R

Samples were collected and processed by J.C.D.C, N.O

All code was conceived and generated by J.C.D.C. , N.R.

Supervision was provided by S.I., R.R, A.A.M.

Funding was provided by S.I., R.R

This first draft of the manuscript was written by J.C.D.C. and S.I.

3.8. Ethics declaration

Competing interests: All authors declare no conflict or competing interests

Chapter 3 - Extended data

All zebrafish work was completed in accordance with institutional, local and national laws and governing bodies. Provide home office project license number here

3.9. Extended data

See Appendix Chapter 3 for supplementary figures, tables and legends.

All scripts and processed data files are available on the GitHub page : <https://github.com/asdfjohnny1/epigenetic-clock-determinants.git>.

3.10. Abbreviations

DNAm – DNA methylation

MAE – Mean Average Error

3.11. References

AKALIN, A., KORMAKSSON, M., LI, S., GARRETT-BAKELMAN, F. E., FIGUEROA, M. E., MELNICK, A. & MASON, C. E. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*, 13, R87.

APSLEY, A. T., YE, Q., CASPI, A., CHIARO, C., ETZEL, L., HASTINGS, W. J., HEIM, C. C., KOZLOSKY, J., NOLL, J. G., SCHREIER, H. M. C., SHENK, C. E., SUGDEN, K. & SHALEV, I. 2024. Cross-Tissue Comparison of Epigenetic Aging Clocks in Humans. Cold Spring Harbor Laboratory.

Chapter 3 - References

- AQUINO, E. M., BENTON, M. C., HAUPT, L. M., SUTHERLAND, H. G., GRIFFITHS, L. R. & LEA, R. A. 2018. Current Understanding of DNA Methylation and Age-related Disease. *OBM Genetics*, 02, 1-16.
- BELSKY, D. W., CASPI, A., HOUTS, R., COHEN, H. J., CORCORAN, D. L., DANESE, A., HARRINGTON, H., ISRAEL, S., LEVINE, M. E., SCHAEFER, J. D., SUGDEN, K., WILLIAMS, B., YASHIN, A. I., POULTON, R. & MOFFITT, T. E. 2015. Quantification of biological aging in young adults. *Proc Natl Acad Sci U S A*, 112, E4104-10.
- CHRUSCIEL, J. H., ORSO, R., DE MATTOS, B. P., FRIES, G. R., KRISTENSEN, C. H., GRASSI-OLIVEIRA, R. & VIOLA, T. W. 2022. A systematic review and meta-analysis of epigenetic clocks in schizophrenia. *Schizophr Res*, 246, 172-174.
- DABROWSKI, J. K., YANG, E. J., CROFTS, S. J. C., HILLARY, R. F., SIMPSON, D. J., MCCARTNEY, D. L., MARIONI, R. E., LATORRE-CRESPO, E. & CHANDRA, T. 2023. Probabilistic inference of epigenetic age acceleration from cellular dynamics. Cold Spring Harbor Laboratory.
- DEATON, A. M. & BIRD, A. 2011. CpG islands and the regulation of transcription. *Genes & Development*, 25, 1010-1022.
- ESTER, M., KRIEGEL, H.-P., SANDER, J. & XU, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. 1996/01/1 1996. 226-331.
- FAN, H., XIE, Q., ZHANG, Z., WANG, J., CHEN, X. & QIU, P. 2022. Chronological Age Prediction: Developmental Evaluation of DNA Methylation-Based Machine Learning Models. *Frontiers in Bioengineering and Biotechnology*, 9.
- FIELD, A. E., ROBERTSON, N. A., WANG, T., HAVAS, A., IDEKER, T. & ADAMS, P. D. 2018. DNA Methylation Clocks in Aging: Categories, Causes, and Consequences. *Molecular Cell*, 71, 882-895.

Chapter 3 - References

- FRIEDMAN, J., HASTIE, T. & TIBSHIRANI, R. 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*, 33, 1-22.
- HANNUM, G., GUINNEY, J., ZHAO, L., ZHANG, L., HUGHES, G., SADDA, S., KLOTZLE, B., BIBIKOVA, M., FAN, J.-B., GAO, Y., DECONDE, R., CHEN, M., RAJAPAKSE, I., FRIEND, S., IDEKER, T. & ZHANG, K. 2013. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell*, 49, 359-367.
- HERNANDEZ-MEZA, G., VON FELDEN, J., GONZALEZ-KOZLOVA, E. E., GARCIA-LEZANA, T., PEIX, J., PORTELA, A., CRAIG, A. J., SAYOLS, S., SCHWARTZ, M., LOSIC, B., MAZZAFERRO, V., ESTELLER, M., LLOVET, J. M. & VILLANUEVA, A. 2021. DNA Methylation Profiling of Human Hepatocarcinogenesis. *Hepatology*, 74, 183-199.
- HIGGINS-CHEN, A. T., THRUSH, K. L., WANG, Y., MINTEER, C. J., KUO, P.-L., WANG, M., NIIMI, P., STURM, G., LIN, J., MOORE, A. Z., BANDINELLI, S., VINKERS, C. H., VERMETTEN, E., RUTTEN, B. P. F., GEUZE, E., OKHUIJSEN-PFEIFER, C., VAN DER HORST, M. Z., SCHREITER, S., GUTWINSKI, S., LUYKX, J. J., PICARD, M., FERRUCCI, L., CRIMMINS, E. M., BOKS, M. P., HÄGG, S., HUSELIGER, T. T. & LEVINE, M. E. 2022. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nature Aging*, 2, 644-661.
- HILLARY, R. F., TREJO-BANOS, D., KOUSATHANAS, A., MCCARTNEY, D. L., HARRIS, S. E., STEVENSON, A. J., PATXOT, M., OJAVEE, S. E., ZHANG, Q., LIEWALD, D. C., RITCHIE, C. W., EVANS, K. L., TUCKER-DROB, E. M., WRAY, N. R., MCRAE, A. F., VISSCHER, P. M., DEARY, I. J., ROBINSON, M. R. & MARIONI, R. E. 2020. Multi-method genome- and epigenome-wide studies of inflammatory protein levels in healthy older adults. *Genome Med*, 12, 60.

Chapter 3 - References

- HORVATH, S. 2013. DNA methylation age of human tissues and cell types. *Genome Biology*, 14, R115.
- JIANG, S. & GUO, Y. 2020. Epigenetic Clock: DNA Methylation in Aging. *Stem Cells International*, 2020, 1-9.
- JONES, M. J., GOODMAN, S. J. & KOBOR, M. S. 2015. <scp>DNA</scp> methylation and healthy human aging. *Aging Cell*, 14, 924-932.
- KRUEGER, F. & ANDREWS, S. R. 2011. Bismark: a flexible aligner and methylation caller for Bisulphite-Seq applications. *Bioinformatics*, 27, 1571-2.
- LEHNE, B., DRONG, A. W., LOH, M., ZHANG, W., SCOTT, W. R., TAN, S.-T., AFZAL, U., SCOTT, J., JARVELIN, M.-R., ELLIOTT, P., MCCARTHY, M. I., KOONER, J. S. & CHAMBERS, J. C. 2015. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biology*, 16, 37.
- LEVINE, M. E., LU, A. T., QUACH, A., CHEN, B. H., ASSIMES, T. L., BANDINELLI, S., HOU, L., BACCARELLI, A. A., STEWART, J. D., LI, Y., WHITSEL, E. A., WILSON, J. G., REINER, A. P., AVIV, A., LOHMAN, K., LIU, Y., FERRUCCI, L. & HORVATH, S. 2018. An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, 10, 573-591.
- LI, A., KOCH, Z. & IDEKER, T. 2022. Epigenetic aging: Biological age prediction and informing a mechanistic theory of aging. *Journal of Internal Medicine*, 292, 733-744.
- LI, X., LI, W. & XU, Y. 2018. Human Age Prediction Based on DNA Methylation Using a Gradient Boosting Regressor. *Genes (Basel)*, 9.
- LÓPEZ-OTÍN, C., BLASCO, M. A., PARTRIDGE, L., SERRANO, M. & KROEMER, G. 2013. The hallmarks of aging. *Cell*, 153, 1194-217.

Chapter 3 - References

LU, A. T., FEI, Z., HAGHANI, A., ROBECK, T. R., ZOLLER, J. A., LI, C. Z., LOWE, R., YAN, Q., ZHANG, J., VU, H., ABLAEVA, J., ACOSTA-RODRIGUEZ, V. A., ADAMS, D. M., ALMUNIA, J., ALOYSIUS, A., ARDEHALI, R., ARNESON, A., BAKER, C. S., BANKS, G., BELOV, K., BENNETT, N. C., BLACK, P., BLUMSTEIN, D. T., BORS, E. K., BREEZE, C. E., BROOKE, R. T., BROWN, J. L., CARTER, G. G., CAULTON, A., CAVIN, J. M., CHAKRABARTI, L., CHATZISTAMOU, I., CHEN, H., CHENG, K., CHIAVELLINI, P., CHOI, O. W., CLARKE, S. M., COOPER, L. N., COSSETTE, M. L., DAY, J., DEYOUNG, J., DIROCCO, S., DOLD, C., EHMKE, E. E., EMMONS, C. K., EMMRICH, S., ERBAY, E., ERLACHER-REID, C., FAULKES, C. G., FERGUSON, S. H., FINNO, C. J., FLOWER, J. E., GAILLARD, J. M., GARDE, E., GERBER, L., GLADYSHEV, V. N., GORBUNOVA, V., GOYA, R. G., GRANT, M. J., GREEN, C. B., HALES, E. N., HANSON, M. B., HART, D. W., HAULENA, M., HERRICK, K., HOGAN, A. N., HOGG, C. J., HORE, T. A., HUANG, T., IZPISUA BELMONTE, J. C., JASINSKA, A. J., JONES, G., JOURDAIN, E., KASHPUR, O., KATCHER, H., KATSUMATA, E., KAZA, V., KIARIS, H., KOBOR, M. S., KORDOWITZKI, P., KOSKI, W. R., KRÜTZEN, M., KWON, S. B., LARISON, B., LEE, S. G., LEHMANN, M., LEMAITRE, J. F., LEVINE, A. J., LI, C., LI, X., LIM, A. R., LIN, D. T. S., LINDEMANN, D. M., LITTLE, T. J., MACORETTA, N., MADDOX, D., MATKIN, C. O., MATTISON, J. A., MCCLURE, M., MERGL, J., et al. 2023. Universal DNA methylation age across mammalian tissues. *Nature Aging*, 3, 1144-1166.

LU, A. T., QUACH, A., WILSON, J. G., REINER, A. P., AVIV, A., RAJ, K., HOU, L., BACCARELLI, A. A., LI, Y., STEWART, J. D., WHITSEL, E. A., ASSIMES, T. L., FERRUCCI, L. & HORVATH, S. 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, 11, 303-327.

MARTÍNEZ-ENGUITA, D., HILLERTON, T., ÅKESSON, J., KLING, D., LERM, M. & GUSTAFSSON, M. 2025. Precise and interpretable neural networks reveal epigenetic signatures of aging across youth in health and disease. *Frontiers in Aging*, 5.

Chapter 3 - References

MATHUR, A., TAURIN, S. & ALSHAMMARY, S. 2024. New insights into methods to measure biological age: a literature review. *Frontiers in Aging*, 5.

MCEWEN, L. M., O'DONNELL, K. J., MCGILL, M. G., EDGAR, R. D., JONES, M. J., MACISAAC, J. L., LIN, D. T. S., RAMADORI, K., MORIN, A., GLADISH, N., GARG, E., UNTERNAEHRER, E., POKHVISNEVA, I., KARNANI, N., KEE, M. Z. L., KLENGEL, T., ADLER, N. E., BARR, R. G., LETOURNEAU, N., GIESBRECHT, G. F., REYNOLDS, J. N., CZAMARA, D., ARMSTRONG, J. M., ESSEX, M. J., DE WEERTH, C., BEIJERS, R., TOLLENAAR, M. S., BRADLEY, B., JOVANOVIĆ, T., RESSLER, K. J., STEINER, M., ENTRINGER, S., WADHWA, P. D., BUSS, C., BUSH, N. R., BINDER, E. B., BOYCE, W. T., MEANEY, M. J., HORVATH, S. & KOBOR, M. S. 2020. The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proceedings of the National Academy of Sciences*, 117, 23329-23335.

MEER, M. V., PODOLSKIY, D. I., TYSHKOVSKIY, A. & GLADYSHEV, V. N. 2018. A whole lifespan mouse multi-tissue DNA methylation clock. *Elife*, 7.

R CORE TEAM 2022. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.

RAKYAN, V. K., DOWN, T. A., MASLAU, S., ANDREW, T., YANG, T.-P., BEYAN, H., WHITTAKER, P., MCCANN, O. T., FINER, S., VALDES, A. M., LESLIE, R. D., DELOUKAS, P. & SPECTOR, T. D. 2010. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Research*, 20, 434-439.

SIMPSON, D. J., ZHAO, Q., OLOVA, N. N., DABROWSKI, J., XIE, X., LATORRE-CRESPO, E. & CHANDRA, T. 2023. Region-based epigenetic clock design improves RRBS-based age prediction. *Aging Cell*, 22, e13866.

STUBBS, T. M., BONDER, M. J., STARK, A.-K., KRUEGER, F., VON MEYENN, F., STEGLE, O. & REIK, W. 2017. Multi-tissue DNA methylation age predictor in mouse. *Genome Biology*, 18.

Chapter 3 - References

- ZHANG, Q., VALLERGA, C. L., WALKER, R. M., LIN, T., HENDERS, A. K., MONTGOMERY, G. W., HE, J., FAN, D., FOWDAR, J., KENNEDY, M., PITCHER, T., PEARSON, J., HALLIDAY, G., KWOK, J. B., HICKIE, I., LEWIS, S., ANDERSON, T., SILBURN, P. A., MELLICK, G. D., HARRIS, S. E., REDMOND, P., MURRAY, A. D., PORTEOUS, D. J., HALEY, C. S., EVANS, K. L., MCINTOSH, A. M., YANG, J., GRATTEN, J., MARIONI, R. E., WRAY, N. R., DEARY, I. J., MCRAE, A. F. & VISSCHER, P. M. 2019. Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome Medicine*, 11.
- ZHOU, W., HINOUE, T., BARNES, B., MITCHELL, O., IQBAL, W., LEE, S. M., FOY, K. K., LEE, K. H., MOYER, E. J., VANDERARK, A., KOEMAN, J. M., DING, W., KALKAT, M., SPIX, N. J., EAGLESON, B., POSPISILIK, J. A., SZABÓ, P. E., BARTOLOMEI, M. S., VANDER SCHAAF, N. A., KANG, L., WISEMAN, A. K., JONES, P. A., KRAWCZYK, C. M., ADAMS, M., PORECHA, R., CHEN, B. H., SHEN, H. & LAIRD, P. W. 2022. DNA methylation dynamics and dysregulation delineated by high-throughput profiling in the mouse. *Cell Genom*, 2.
- ZOU, H. & HASTIE, T. 2005. Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67, 301-320.

CHAPTER FOUR

Cross-tissue comparison of ageing in somatic tissues of zebrafish

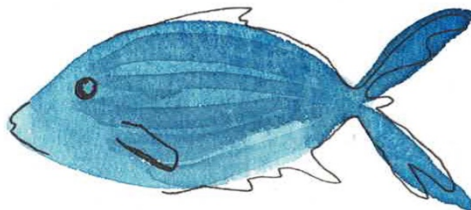
Jean-Charles de Coriolis¹, Neil A. Robertson^{2,3}, Nelly Olova², Raheleh Rahabari⁴,
Tamir Chandra^{2,3}, Alexei A. Maklakov¹, Simone Immler¹

¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, NR4 6TX, Norwich, UK

² MRC Human Genetics Unit, Institute of Genetics and Cancer. The University of Edinburgh, Edinburgh, UK

³ MAYO Clinic, Rochester, Minnesota, USA

⁴ Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK



4. Chapter 4

4.1. Abstract

Epigenetic clocks are widely used to quantify biological ageing, yet their performance and interpretation across tissues remain poorly resolved. Here, we present a systematic comparison of epigenetic ageing across zebrafish somatic tissues with contrasting proliferative and regenerative properties. Using genome-wide DNA methylation data from brain, caudal fin, and intestine, we benchmarked 3,036 epigenetic clock models spanning multiple CpG aggregation strategies, regression frameworks, and feature selection parameters. Clock performance varied strongly by tissue, with brain-derived clocks consistently outperforming fin, intestine, and multi-tissue clocks, indicating pronounced tissue-specific age predictability. CpG aggregation strategy affected performance in a tissue-dependent manner: regional aggregation improved tissue-specific clocks but reduced the performance of multi-tissue models in this dataset, highlighting a trade-off between signal stability and cross-tissue generalisability. Epigenetic age and its trajectories across chronological age differed markedly between tissues, with intestine exhibiting the highest acceleration and caudal fin the lowest, independent of modelling strategy, reflecting distinct ageing kinetics. Analysis of CpG architecture revealed both tissue-specific and overlapping somatic CpGs. While overlapping CpGs showed uniform mean methylation across tissues, they exhibited tissue-specific regulatory dynamics at the individual CpG level and were enriched near Polycomb group target genes and core transcriptional regulators. Higher-order methylation features indicated tissue-specific patterns of stochastic drift and entropy. Together, these findings support a mosaic model of epigenetic ageing, in which conserved regulatory vulnerabilities coexist with tissue-specific ageing trajectories. Methodologically, our results demonstrate that optimal

epigenetic clock design is conditional on tissue context and biological question, with important implications for the interpretation of epigenetic ageing biomarkers across systems.

4.2. Introduction

Ageing is a progressive decline in physiological integrity that leads to impaired function, increased morbidity, and elevated risk of mortality (López-Otín et al., 2013). Emerging evidence indicates that ageing is not uniform across tissues. Histological deterioration, regenerative capacity, and functional decline manifest at markedly different rates in distinct tissues (Terman and Brunk, 2006, Baker et al., 2011), suggesting that ageing is a mosaic process shaped by tissue-specific development, proliferative histories, metabolic demands, and environmental exposures (de Magalhães and Church, 2005, Tomasetti and Vogelstein, 2015, Alexandrov et al., 2015, Yamamoto et al., 2022, Stoeger et al., 2022). Despite this recognition, the molecular mechanisms that encode tissue-specific ageing trajectories remain incompletely understood.

Several theoretical frameworks predict differential ageing across tissues. Somatic mutation accumulation scales with stem cell division and replication-associated DNA damage, implying that highly proliferative tissues accumulate molecular damage more rapidly than largely post-mitotic tissues (Tomasetti and Vogelstein, 2015). Consistent with this idea, intestinal epithelium undergoes continuous renewal and exhibits elevated mutation burdens relative to post-mitotic tissues such as the brain (Martincorena et al., 2015, Blokzijl et al., 2016, Bae et al., 2018). Tissues that combine low baseline turnover with regenerative potential either accelerate epigenetic ageing through proliferative reprogramming or buffer it through activation of developmental programmes (Yun et al., 2015, Coffman et al.,

2016, Da Silva-Álvarez et al., 2020). Evolutionary frameworks, including the disposable soma theory, the expensive germline hypothesis and the developmental theory of ageing, predict that investment in somatic maintenance and repair is shaped by energetic trade-offs, reproductive priorities and developmental constraints, leading to systematic differences in ageing trajectories. (Kirkwood, 1977, Maklakov and Immler, 2016, Maklakov and Chapman, 2019).

To interrogate these principles empirically, we focused on three tissues in zebrafish (*Danio rerio*), namely brain, intestine and caudal fin. These three tissues span distinct axes of proliferative activity, mutation exposure, and regenerative potential. The zebrafish brain consists predominantly of post-mitotic neurons with limited adult neurogenesis and relatively stable chromatin architecture (Kizil et al., 2012). In contrast, the intestine are highly proliferative, with constant epithelial turnover driven by stem cell division, rendering it susceptible to replication-associated mutations and epigenetic drift (Barker, 2014, Nalapareddy et al., 2022). The caudal fin occupies an intermediate state, with largely quiescent cells under homeostasis, yet capable of rapid regeneration via progenitor activation, epigenetic reprogramming, and reactivation of developmental transcriptional programmes (Poss et al., 2002, Lee et al., 2020). By using epigenetic ageing, these tissues provide a tractable model for examining how proliferative history and regenerative potential interact with the sample's biological age.

DNA methylation-based epigenetic clocks are among the most accurate molecular biomarkers of chronological age across vertebrates (Horvath, 2013, Hannum et al., 2013). By quantifying reproducible age-associated changes at CpG sites, they allow precise estimation of chronological age and the assessment of deviations from expected age (Δ -Age or age acceleration), which are linked to mortality, disease susceptibility and environmental stress exposure (Levine et al., 2018). In zebrafish, species-specific clocks confirm

evolutionary conservation of epigenetic ageing signatures (Mayne et al., 2020). However, most clocks are trained on a single tissue, often blood or epidermis due to sampling accessibility, and consequently assume that ageing trajectories in that tissue provide a reasonable proxy for systemic ageing (Horvath and Raj, 2018, Jylhävä et al., 2017). While multi-tissue clocks exist but these are optimised across species rather than across tissues within an organism (Horvath, 2013, Higgins-Chen et al., 2022), raising the question of whether tissue-specific molecular ageing can be extrapolated to whole organisms.

Ageing-associated methylation changes are not uniform across the genome. Promoter hypomethylation of developmental and polycomb-repressed genes, a class of developmentally regulated genes maintained in a repressed chromatin state, contributes to transcriptional silencing and loss of cellular identity (Rakyan et al., 2010), while intergenic regions experience global hypomethylation. Gene body methylation, linked to active transcription, may shift with age and influence splicing or elongation (Zampieri et al., 2015). The magnitude and direction of these changes differ between tissues, with proliferative tissues accruing replication-induced errors, whereas quiescent tissues experience passive demethylation or chromatin remodelling via oxidative stress and transcription factor erosion (Field et al., 2018). Dissecting methylation dynamics across regulatory contexts is therefore crucial for understanding tissue-specific ageing. However, it is unclear whether high-turnover tissues such as intestine show greater epigenetic age acceleration than post-mitotic tissues like the brain, or whether regenerative tissues such as the caudal fin, resist or accelerate ageing. Most studies aggregate CpG signals genome-wide, potentially obscuring region-specific dynamics in promoters, gene bodies, or transposable elements.

Here, we quantify and compare epigenetic ageing across zebrafish brain, intestine and caudal fin using reduced representation bisulphite sequencing

for DNA methylation profiling. Specifically, we examine (i) whether epigenetic age and age acceleration differ systematically between tissues, (ii) whether tissue-specific methylation drift correlates with proliferative and mutational histories, and (iii) how methylation dynamics in promoter and gene bodies contribute to divergent tissue ageing trajectories. By combining evolutionary theory, epigenetic biomarkers and regulatory-context analyses, we aim to determine whether ageing is inherently mosaic and to identify the molecular signature underlying tissue-level divergence.

4.3. Material and Methods

All experimental procedures, data processing and statistical analyses used in this chapter are identical to those described in Chapter 3 [Material and Methods](#), unless stated otherwise below. The distribution of samples across ages and organs can be found in Table S16, and across technical and biological factors in Table S17.

4.3.1. Genomic feature annotation and methylation quantification

We quantified CpG methylation across four genomic contexts: gene bodies, promoters, transposable elements (TEs), and a methylation-based mutability proxy. Annotations were based on the *Danio rerio* GRCz11 genome and information on gene bodies and promoters were derived from Ensembl release 111 (Howe et al., 2013). Gene bodies spanned the full transcribed region, whilst promoters were defined as ± 1 kb from the transcription start site, accounting for strand orientation. TE coordinates were obtained from RepeatMasker (Smith et al., 2013-2015) annotations and filtered to retain non-redundant autosomal elements with Class and subfamily information. CpG

density adjusted methylation levels were additionally used as a proxy for regional mutability (Zhou et al., 2020).

4.4. Results

Following quality control and filtering, 182 samples were retained for downstream analyses, comprising 62 brain, 61 caudal fin, and 59 intestine samples. With this dataset, we constructed and evaluated a comprehensive set of 3,036 candidate epigenetic clock models across different CpG aggregation strategies, regression frameworks, and feature selection parameters. This results in 859 models for brain, 816 models for caudal fin, 824 model for intestine, and 870 models for multi-tissue models (Table S17).

4.4.1. Epigenetic clock performance is strongly shaped by tissue context and CpG aggregation strategy

Across all epigenetic clock models (Table S17), prediction performance (composite MAE + R^2 score) varied markedly by tissue of origin (LMM, $\chi^2 = 85.6$, $df = 3$, $p < 0.0001$; Figure 9A; Table S18). Brain-derived models consistently outperformed those trained on caudal fin (*Estimate* = 0.23, $z = 7.9$, $p < 0.0001$), intestine (*Estimate* = 0.20, $z = 5.6$, $p < 0.0001$), and multi-tissue models (*Estimate* = 0.10, $z = 2.6$, $p = 0.05$; Figure 9A). In contrast, caudal fin-derived models performed significantly worse than multi-tissue models (*Estimate* = -0.30, $z = -3.6$, $p = 0.002$), while differences between caudal fin and intestine (*Estimate* = -0.60, $z = -2.2$, $p = 0.1$) and between intestine and multi-tissue models (*Estimate* = -0.10, $z = -1.8$, $p = 0.2$) were not significant (Figure 9A). Together, these results demonstrate pronounced tissue-specific epigenetic age predictability, with brain clocks showing the most robust performance.

Chapter 4 - Results

Clock performance was also strongly influenced by CpG aggregation strategy (LMM, $\chi^2 = 188.5$, $df = 2$, $p < 0.0001$; Figure 9B; Table S18). Across tissues, Cluster-based CpGs performed significantly worse than Tile CpGs (*Estimate* = -0.10, $z = -9.7$, $p < 0.0001$), while no significant difference was detected between Cluster and Point-based CpGs (*Estimate* = -0.04, $z = -1.2$, $p = 0.4$; Figure 9B). The contrast between Point and Cluster CpGs was weaker and did not reach statistical significance (*Estimate* = -0.07, $z = -2.0$, $p = 0.1$).

When stratified by tissue, CpG grouping effects were highly tissue-specific (Figure S9; Table S18). In brain models, Cluster-based CpGs substantially reduced performance relative to both Tile (*Estimate* = -0.10, $z = -13.7$, $p < 0.0001$) and Point-based CpGs (*Estimate* = -0.20, $z = -3.4$, $p = 0.002$), with the Cluster–Point contrast approaching statistical significance (*Estimate* = -0.10, $z = -2.1$, $p = 0.08$; Figure S9). A similar pattern was observed in intestinal models, where clustered CpGs underperformed relative to Tile CpGs (*Estimate* = -0.20, $z = -8.5$, $p < 0.0001$), while other contrasts were not significant. In caudal fin models, no CpG grouping comparisons reached significance (Figure S9). In contrast, multi-tissue models showed the opposite trend, with Cluster-based CpGs outperforming Tile CpGs (*Estimate* = 0.10, $z = 4.4$, $p < 0.0001$), and no significant differences between other groupings. These results indicate that regional methylation features are detrimental for multi-tissue clocks but can enhance performance when modelling shared, multi-tissue ageing signals.

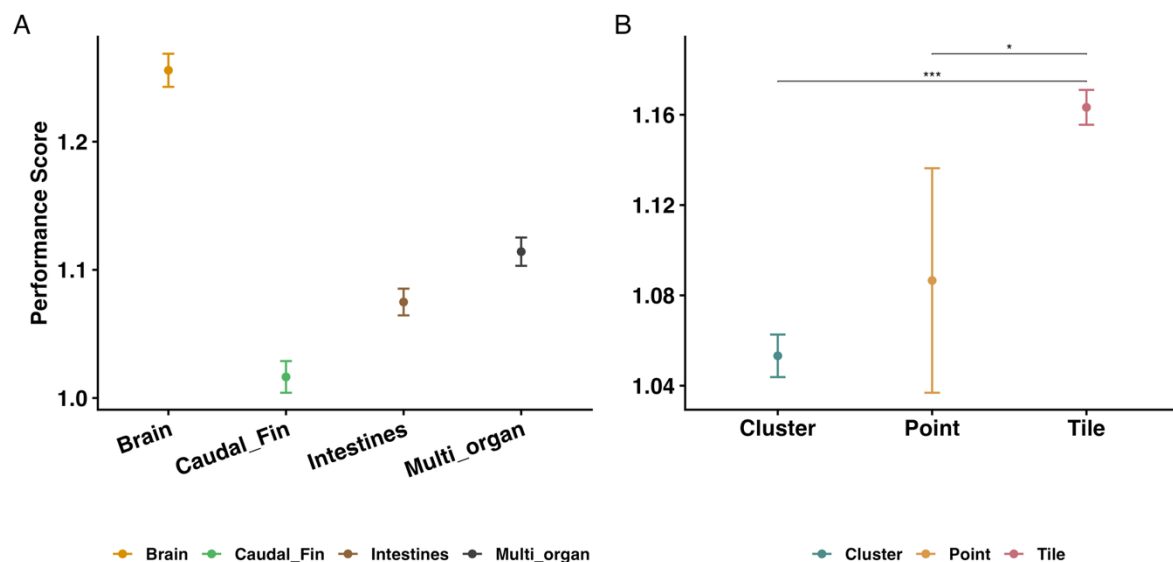


Figure 9: Tissue and CpG aggregation strongly influence clock performance score.

A) Clock performance score for epigenetic age prediction models trained within four tissue groupings: brain, caudal fin, intestine, and a multi-tissue model combining all three tissues. Performance score is defined as a composite metric integrating model fit (R^2) and prediction error (inverse mean absolute error, MAE). **B)** Clock performance score across three CpG aggregation strategies: individual CpGs, data-driven CpG clusters, and fixed genomic Tiles. Points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

Chapter 4 - Results

4.4.2. Tissue-specific differences in epigenetic age acceleration and ageing trajectories

Epigenetic age acceleration, quantified as Δ -Age (chronological age minus predicted age), differed significantly between tissues (LMM, $\chi^2 = 35.8$, $df = 3$, $p < 0.0001$), with intestine exhibiting the greatest acceleration and caudal fin the lowest (Figure 10A; Table S18). Δ -Age in intestine was significantly higher than in brain (*Estimate* = -0.90, $z = -8.4$, $p < 0.0001$) and caudal fin (*Estimate* = -1.20, $z = -10.8$, $p < 0.0001$) and also exceeded predictions from multi-tissue clocks (*Estimate* = 0.70, $z = 7.8$, $p < 0.0001$; Figure 10A). By contrast, brain showed moderately elevated Δ -Age relative to caudal fin (*Estimate* = 0.30, $z = 2.6$, $p = 0.04$), while caudal fin-derived estimates were consistently lower than those from multi-tissue clocks (*Estimate* = -0.50, $z = -5.4$, $p < 0.0001$; Figure 10A). There was a marginal brain–multi-tissue trend (*Estimate* = -0.20, $z = -2.3$, $p = 0.09$). These tissue-specific differences were independent of CpG aggregation strategy, as neither CpG grouping (LMM, $\chi^2 = 2.1$, $df = 2$, $p = 0.3$) nor CpG grouping $\times$ tissue interactions (LMM, $\chi^2 = 8.4$, $df = 6$, $p = 0.2$) significantly affected Δ -Age.

Beyond baseline offsets, Δ -Age trajectories diverged markedly with chronological age across tissues (tissue $\times$ age interaction: LMM, $\chi^2 = 1073.8$, $df = 3$, $p < 0.0001$; Figure 10B; Table S18). Brain and caudal fin exhibited significantly different age-dependent slopes (*Estimate* = 0.30, $z = 5.7$, $p < 0.0001$), as did comparisons between caudal fin and intestine (*Estimate* = -0.40, $z = -6.3$, $p < 0.0001$) and caudal fin and multi-tissue clocks (*Estimate* = -0.40, $z = -7.7$, $p < 0.0001$; Figure 10B). In contrast, slope differences between brain and intestine (*Estimate* = -0.40, $z = -0.7$, $p = 0.8$), brain and multi-tissue clocks (*Estimate* = -0.60, $z = -1.2$, $p = 0.6$), and intestine and multi-tissue clocks (*Estimate* = -0.02, $z = -0.3$, $p = 0.9$) were not statistically significant (Figure 10B).

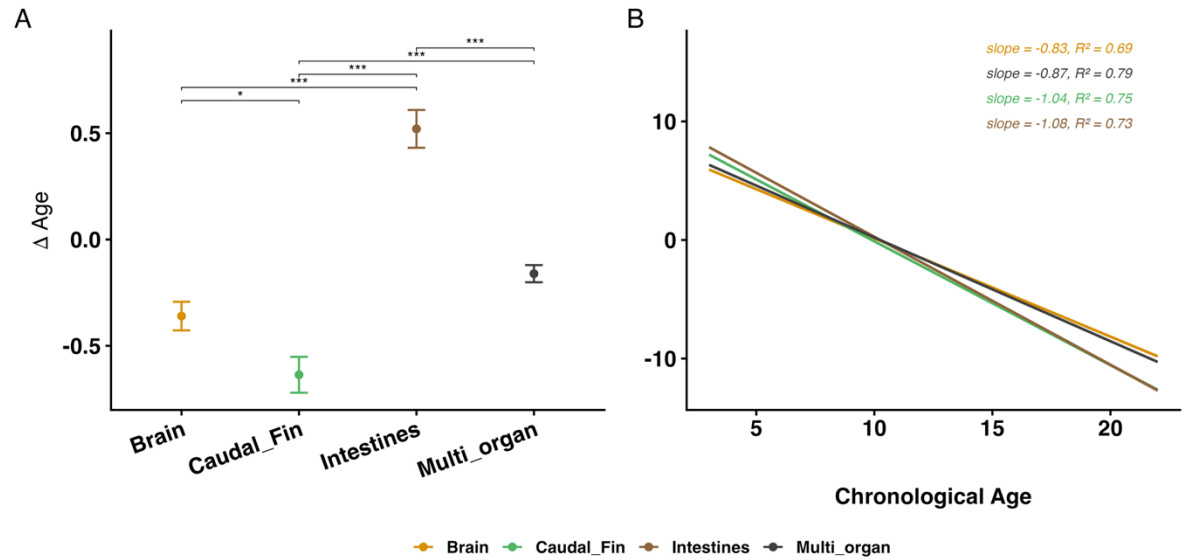


Figure 10: Tissue strongly impacts Δ -Age and the rate of epigenetic ageing.

A) Mean epigenetic age deviation (Δ -Age), defined as the difference between predicted epigenetic age and chronological age, estimated from epigenetic clock models across four tissue groupings: brain, caudal fin, intestine, and a multi-tissue model combining all three tissues. **B)** Estimated rate of epigenetic ageing, quantified as the slope of epigenetic age with respect to chronological age, for the same tissue groupings. Points indicate group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is denoted by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$)

Chapter 4 - Results

4.4.3. Tissue-specific and overlapping CpG architectures

Based on all 3,036 model iterations, two classes of CpG sets were defined. The first comprised tissue-specific CpGs, identified independently for each tissue by retaining CpGs selected in at least 10% of final models within that tissue. This yielded 18 brain-, 47 caudal fin-, 38 intestine-, and 44 multi-tissue-associated CpGs (Table S19). These CpGs represent loci with robust age-associated signal within a tissue but not necessarily shared across tissues. The second class consisted of overlapping CpGs, defined as CpGs selected across all somatic tissues (brain, intestine, and caudal fin), resulting in 30 shared loci (Table S19). These CpGs capture age-associated methylation changes that are conserved across somatic tissues and therefore reflect more general, soma-wide ageing signatures rather than tissue-restricted effects.

For tissue-specific CpGs, mean methylation differed significantly between tissues (LMM : $\chi^2 = 2,407$, $df = 3$, $p < 0.0001$; Table S20), brain CpGs exhibited higher methylation than those in caudal fin (*Estimate* = 16.6, $z = 13.7$, $p < 0.0001$), intestine (*Estimate* = 41.1, $z = 11.2$, $p < 0.0001$), and multi-tissue sets (*Estimate* = 5.8, $z = 5.3$, $p < 0.0001$). Caudal fin CpGs were more highly methylated than intestine (*Estimate* = 25.3, $z = 26.5$, $p < 0.0001$) but less than multi-tissue CpGs (*Estimate* = -10.7, $z = -14.7$, $p < 0.0001$), and CpGs in intestine showed the lowest methylation overall. (*Estimate* = -36, $z = -44.2$, $p < 0.0001$; Figure 11A). No sex-specific differences were detected (LMM: $\chi^2 = 0.2$, $df = 1$, $p = 0.6$; Figure S10A; Table S20). These tissue-specific methylation differences were also maintained across chronological age, with approximately linear trajectories within each tissue (LMM: $\chi^2 = 14.7$, $df = 27$, $p = 1$; Figure S10; Table S20).

In contrast, overlapping CpGs showed no significant differences in mean methylation between tissues (LMM: $\chi^2 = 0.9$, $df = 2$, $p = 0.6$; Figure 12A; Table

S21), sexes (LMM: $\chi^2 = 0.3$, $df = 1$, $p = 0.6$; Figure S11; Table S21), or across chronological age (LMM: $\chi^2 = 4.5$, $df = 18$, $p = 1$; Figure S11; Table S21), indicating that, at the tissue-averaged level, methylation at these CpGs appears uniform. However, examination at the individual CpG level revealed pronounced tissue-specific patterns. Heatmap-based clustering of scaled CpG methylation trajectories showed clear tissue-level separation over time, reflecting distinct regulatory dynamics within each tissue (Figure 12B). Quantitative analyses confirmed these observations: median pairwise correlations of CpG trajectories were near zero between brain and intestine ($r = 0.06$) and brain and caudal fin ($r = 0.11$), and only moderate between intestine and caudal fin ($r = 0.39$). Supporting this, permutational multivariate analysis of variance on Euclidean distances of scaled methylation profiles indicated a strong effect of tissue identity (PERMANOVA, $R^2 = 0.5$, $F = 11.9$, $p = 0.001$; Table S22). Collectively, these results suggest that although overlapping CpGs do not drive differences in mean methylation at the tissue level, they participate in highly tissue-specific regulatory programmes that are detectable only when analysed at the resolution of individual CpGs.

4.4.4. Functional enrichment of overlapping CpGs

Genes proximal to overlapping somatic CpGs were analysed for functional enrichment using g:Profiler (Raudvere et al., 2019), with zebrafish-specific annotations and all annotated zebrafish genes used as the background set. Enrichment testing was performed across gene ontology biological process, molecular function, and cellular component categories, as well as KEGG pathways, with multiple testing correction applied using the false discovery rate (FDR). Two dominant functional axes were identified. The first axis encompassed core transcriptional control and proteostasis, including RNA polymerase II elongation and CTD phosphorylation, cyclin-CDK regulatory

Chapter 4 - Results

complexes, and protein processing pathways (Figure S11). The second axis reflected neuronal structure and presynaptic signalling. Of the 30 overlapping CpGs, ten were located within gene bodies, three within promoter regions, and the remainder were unclassified (Figure S11). These CpGs also showed significant enrichment for Polycomb group target genes (23%), consistent with epigenetic ageing preferentially accumulating at regulatory hubs with broad transcriptional consequences (Figure S11). No functional annotations were found for tissue-specific CpGs.

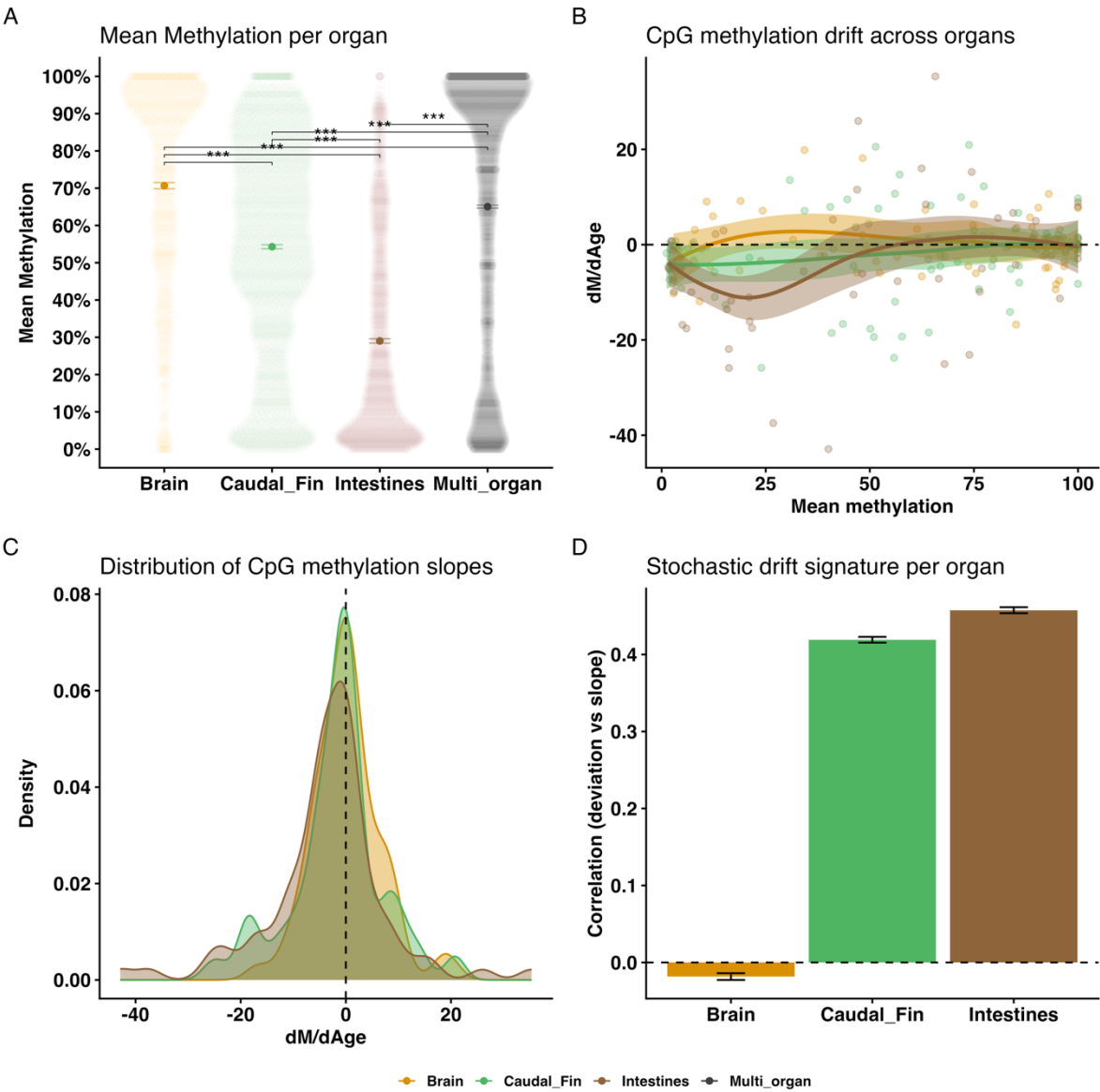


Figure 11: Tissue-specific CpGs and methylation dynamics.

CpGs shown were selected in at least 10% of epigenetic clock models within each tissue. **A)** Mean DNA methylation level per tissue, with beeswarm points representing individual. **B)** Relationship between mean CpG methylation level and age-associated methylation change ($dM/dAge$), where $dM/dAge$ denotes the slope of methylation proportion with respect to chronological age, estimated by linear regression for each CpG within each tissue. A locally estimated scatterplot smoothing (LOESS) curve is overlaid to visualise non-linear trends, with shaded bands indicating 95% confidence intervals. **C)** Density distribution of CpG-specific methylation slopes ($dM/dAge$) for each tissue, with the vertical dashed line indicating zero slope (no age-associated change). **D)** Stochastic drift signature per tissue, quantified as the mean correlation between CpG-specific deviation from the tissue mean methylation level and the corresponding $dM/dAge$ slope. Points indicate tissue-level means, and error bars represent $\pm$ SE. Statistical significance was assessed using a linear mixed-effects model (LMM) and is denoted by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

Chapter 4 - Results

4.4.5. Methylation density, stochastic drift, and entropy

For tissue-specific CpGs, methylation density differed significantly between tissues (LMM, $\chi^2 = 8.9$, $df = 2$, $p = 0.01$; Table S20), driven primarily by higher density in brain relative to intestine (*Estimate* = 3.7, $t = 3.0$, $df = 133$, $p = 0.009$; Figure 11). Differences between brain and caudal fin (*Estimate* = 1.9, $t = 1.5$, $df = 136$, $p = 0.2$) and caudal fin and intestine (*Estimate* = 1.8, $t = 1.5$, $df = 152$, $p = 0.28$) were not significant. Stochastic drift showed a near-significant tissue effect (LMM, $\chi^2 = 5.3$, $df = 2$, $p = 0.06$; Table S20), with greater drift in brain than intestine (*Estimate* = 3.3, $t = 2.6$, $df = 131$, $p = 0.03$), but no significant differences between brain and caudal fin or caudal fin and intestine (Figure 11D).

For overlapping CpGs, neither methylation density (LMM, $\chi^2 = 2.0$, $df = 2$, $p = 0.4$; Table S20) nor stochastic drift (LMM, $\chi^2 = 2.4$, $df = 2$, $p = 0.3$; Table S21) differed significantly between tissues (Figure 12C, 12D). In contrast, entropy differed strongly by tissue (LMM, $\chi^2 = 92.7$, $df = 2$, $p < 0.0001$; Table S21), with intestine exhibiting higher entropy than both brain (*Estimate* = 0.07, $t = 8.2$, $df = 126$, $p < 0.0001$) and caudal fin (*Estimate* = 0.07, $t = 8.5$, $df = 129$, $p < 0.0001$), while brain and caudal fin did not differ (*Estimate* = -0.003, $t = -0.3$, $df = 128$, $p = 0.9$; Figure 13C).

Chapter 4 - Results

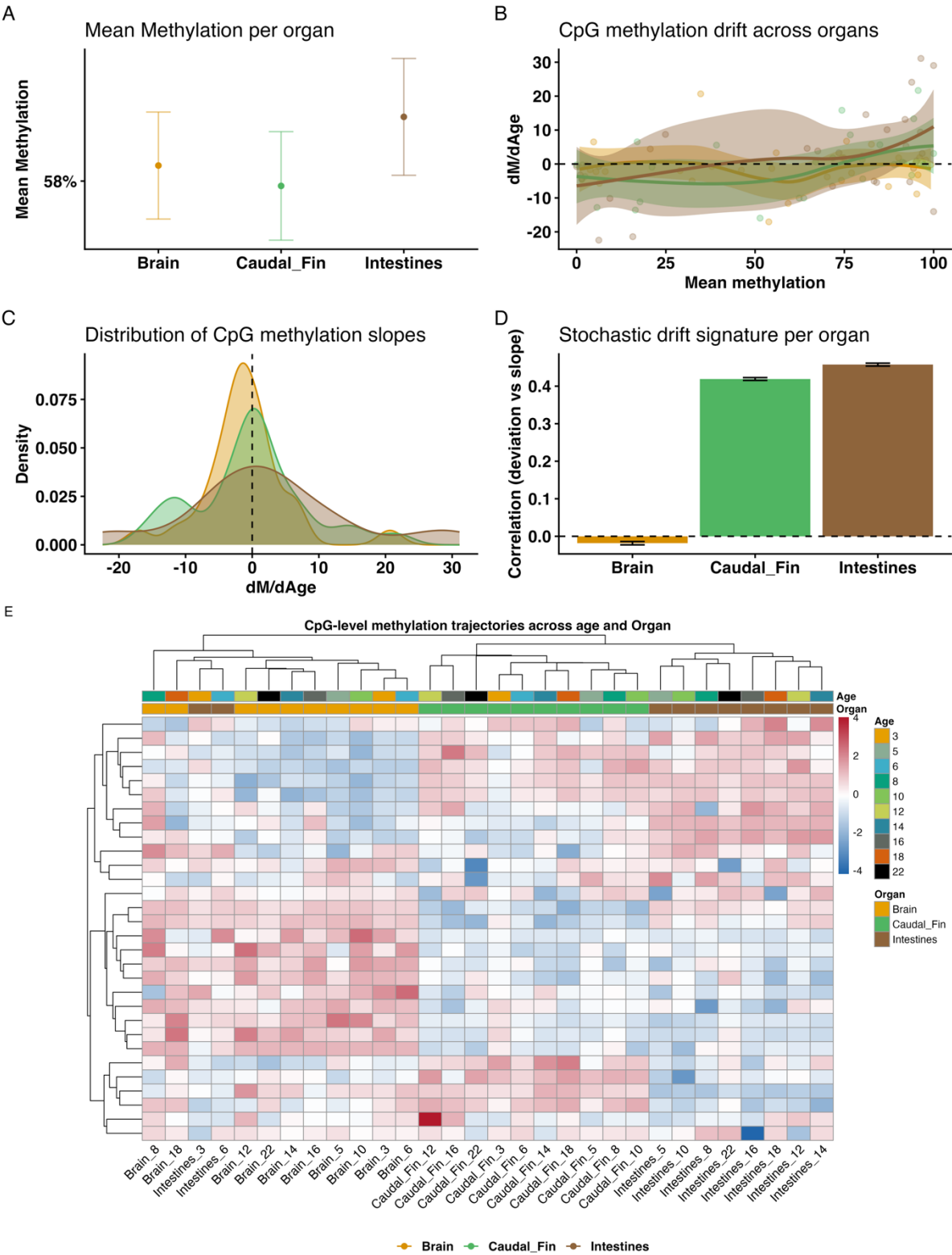


Figure 12: Shared somatic CpG methylation dynamics across tissues.

CpGs shown are those shared across all somatic tissues analysed. **A)** Mean DNA methylation level per tissue, with beeswarm points representing individual. **B)** Relationship between mean CpG methylation level and age-associated methylation change ($dM/dAge$), where $dM/dAge$ denotes the slope of methylation proportion with respect to chronological age, estimated by linear regression for each CpG within each tissue. A locally estimated scatterplot smoothing (LOESS) curve is overlaid to visualise non-linear trends, with shaded bands indicating 95% confidence intervals. **C)** Density distribution of CpG-specific methylation slopes ($dM/dAge$) for each tissue, with the vertical dashed line indicating zero slope (no age-associated change). **D)** Stochastic drift signature per tissue, quantified as the mean correlation between CpG-specific deviation from the tissue mean methylation level and the corresponding $dM/dAge$ slope. **E)** Heatmap of CpG-level methylation trajectories across chronological age and tissue, with rows representing shared CpGs and columns representing individual samples. CpGs and samples are ordered by similarity. Points indicate tissue-level means, and error bars represent $\pm$ SE. Statistical significance was assessed using a linear mixed-effects model (LMM) and is denoted by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

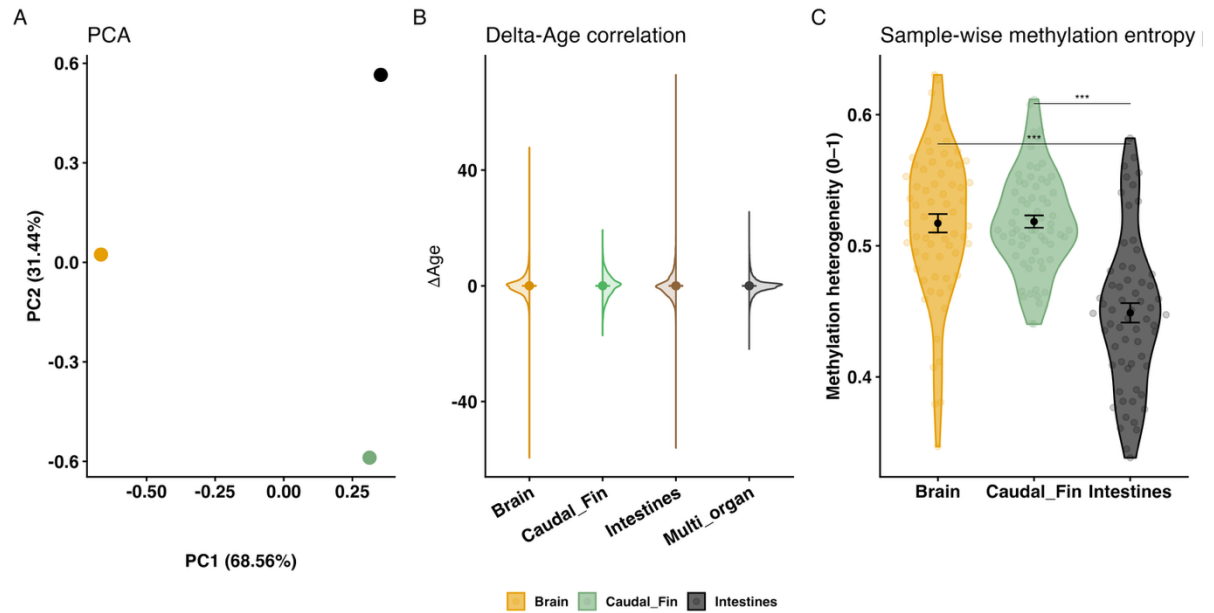


Figure 13: System-level divergence of epigenetic ageing across tissues.

A) Principal component analysis (PCA) of tissue-level DNA methylation profiles, illustrating separation of tissues in reduced-dimensional space based on genome-wide methylation patterns. **B)** Distribution of individual epigenetic age deviation (Δ -Age), defined as the difference between predicted epigenetic age and chronological age, across tissues, highlighting tissue-specific offsets in epigenetic age estimates. **C)** Sample-wise DNA methylation entropy per tissue, used as a measure of epigenetic heterogeneity, with violin plots showing the full distribution of values for each tissue. Points denote tissue-level means, and error bars represent $\pm$ SE. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

4.5. Discussion

Here we present a systematic, within-organism comparison of epigenetic ageing across zebrafish tissues with contrasting proliferative histories, regenerative capacity, and cellular turnover. By benchmarking over 3,000 independently trained epigenetic clocks alongside region-specific methylation analyses, we demonstrated that epigenetic ageing is neither uniform nor fully systemic. Instead, age-associated methylation changes emerged from a mosaic architecture, whereby tissue-specific CpG and region-level ageing trajectories coexist alongside a smaller set of CpGs exhibiting shared, cross-tissue age associations.

4.5.1. Tissue context strongly constrains epigenetic clock performance

The pronounced tissue dependence of clock performance observed here indicates that chronological age is encoded in distinct methylation architectures across tissues. Brain-derived clocks consistently outperformed fin, intestine, and multi-tissue models, suggesting that post-mitotic or low-turnover tissues provide more stable substrates for age prediction. Similar observations have been reported in mammalian systems, where brain and muscle often exhibit smoother epigenetic trajectories and reduced stochastic noise compared to proliferative tissues (Horvath, 2013, Field et al., 2018, Chapter 3). Reduced clonal turnover likely preserves coherent age-associated methylation patterns, enhancing clock precision.

In contrast, intestinal clocks performed less robustly, consistent with the high mitotic activity and continuous epithelial renewal in this tissue (Barker, 2014). Frequent cell division amplifies replication-associated methylation error and clonal heterogeneity, processes linked to both somatic mutation accumulation and epigenetic drift (Alexandrov et al., 2015, Jenkinson et al.,

2017, Blokzijl et al., 2016). Comparable reductions in clock precision have been reported for rapidly renewing mammalian tissues, including blood and gut epithelium (Hannum et al., 2013, Levine et al., 2018).

Caudal fin clocks exhibited intermediate performance. Although largely quiescent under homeostatic conditions, caudal fin tissue retains substantial regenerative capacity. Regeneration involves widespread chromatin remodelling and reactivation of developmental gene programmes (Stewart et al., 2014, Duong et al., 2024), processes that may partially decouple methylation state from chronological time and introduce additional variance into age-associated signals. Similar effects have been observed in regenerating mammalian tissues, where repair-associated reprogramming can transiently reset epigenetic age estimates (Ocampo et al., 2016, Lu et al., 2020).

The divergent behaviour of CpG aggregation strategies further reinforces the importance of tissue-specificity. Tile-based aggregation improved clock performance in the brain and the intestine but systematically degraded performance in multi-tissue models. Regional aggregation appears to enhance noise when cellular composition is relatively homogeneous, consistent with prior reports that regional methylation summaries outperform single CpGs by buffering local stochastic variation (Higgins-Chen et al., 2022, Chrusciel et al., 2022). However, when heterogeneous tissues are pooled, tiling obscures tissue-specific regulatory patterns, reducing the ability to resolve distinct tissue ageing trajectories. Consistent with this pattern, pan-tissue human clocks seem to preferentially rely on individual CpGs with conserved age associations across tissues, rather than regionally aggregated signals (Horvath, 2013, Lu et al., 2019).

Together, these results indicate that optimal clock design is conditional. Tissue-specific clocks benefit from regional aggregation, whereas multi-

tissue clocks require locus-level resolution to allow cross-tissue generalisability.

4.5.2. Epigenetic age acceleration reflects tissue-specific cellular histories

Epigenetic age acceleration differed markedly between tissues, with intestine showing the highest Δ -Age and caudal fin the lowest. These differences align with expectations based on proliferative history, stem-cell dynamics, and somatic mutation burden (Blokzijl et al., 2016, Moore et al., 2021). Importantly, the tissue differences were independent of CpG grouping strategy, indicating a higher likelihood of these results reflecting intrinsic biological properties rather than modelling artefacts.

Beyond differences in overall Δ -Age, tissues also diverged in how epigenetic ageing progressed with chronological age. In particular, caudal fin displayed pronounced age-dependent divergence in Δ -Age trajectories across chronological age, indicating differences in the rate at which epigenetic age accumulates. This pattern is consistent with repeated regenerative activation in caudal fin tissues, which may influence epigenetic ageing through cycles of chromatin remodelling, lineage re-specification, and progenitor expansion. Comparable changes in epigenetic ageing slopes have been reported in mammalian tissues following injury or regenerative interventions (Ocampo et al., 2016, Lu et al., 2020), supporting a link between cellular plasticity and ageing dynamics.

4.5.3. Shared CpGs encode conserved regulatory ageing signals

Tissue-specific CpGs displayed clear tissue-restricted methylation patterns, consistent with divergent ageing trajectories shaped by local cellular

histories. In contrast, overlapping somatic CpGs represent a conserved component of epigenetic ageing. These sites showed minimal differences in mean methylation across tissues, sex, or age when analysed jointly, suggesting they represent a conserved set of CpGs under shared regulatory constraint, rather than tissue-specific ageing signals.

The preferential enrichment of overlapping CpGs near Polycomb group target genes is consistent with the central role of Polycomb-mediated repression in maintaining stable cell identity and transcriptional fidelity across cell divisions. Polycomb repressive complexes (PRC1 and PRC2) establish and maintain repressive chromatin states at developmental and lineage-specifying loci, primarily through histone modifications (Sen et al., 2016). With ageing, Polycomb-regulated domains show progressive epigenetic drift, including loss or redistribution of repressive marks and increased susceptibility to DNA methylation gain, particularly at CpG-dense promoters (Rakyan et al., 2010, Teschendorff et al., 2013, Stubbs et al., 2017).

These features make Polycomb target loci especially prone to age-associated epigenetic change, they are transcriptionally poised, mitotically heritable, and subject to long-term regulatory constraint. As a result, modest methylation changes at these sites can have disproportionate downstream effects on gene regulatory networks, providing a plausible mechanism by which overlapping CpGs capture a conserved, systemically shared component of epigenetic ageing, and support the view that ageing reflects the progressive destabilisation of regulatory networks established during development, consistent with the developmental theory of ageing (Maklakov and Chapman, 2019).

Chapter 4 - Discussion

4.5.4. Divergent drift and entropy reveal tissue-specific ageing modes

Beyond mean methylation changes, tissues differed in higher-order methylation dynamics. Tissue-specific CpGs showed differences in methylation density and stochastic drift, with the brain exhibiting greater drift than the intestine. Although initially counterintuitive, this pattern may reflect stabilising selection in rapidly renewing tissues, where aberrant epigenetic states are purged through cell turnover, whereas long-lived cells accumulate drift over time (Bahar et al., 2006, Martinez-Jimenez et al., 2017). Similar enrichment of epigenetic drift in post-mitotic tissues has been reported in the mammalian brain (Slieker et al., 2018).

In contrast, entropy at overlapping CpGs differed strikingly between tissues, with intestine showing substantially higher entropy than brain or fin. Increased entropy reflects elevated regulatory noise and reduced constraint, hallmarks of ageing systems approaching transcriptional instability (Vaidya et al., 2023). That this effect was concentrated at shared CpGs suggests that systemic ageing manifests not as uniform methylation change, but as tissue-specific increases in regulatory variability at conserved loci.

4.5.5. Conclusion

Together, these findings support a model in which epigenetic ageing emerges from the interaction between shared regulatory vulnerabilities and tissue-specific cellular histories. This mosaic architecture reconciles classical systemic ageing frameworks (López-Otín et al., 2013) and growing evidence for tissue-specific ageing trajectories (Gladyshev et al., 2021). From a methodological perspective, our results do not challenge the utility of tissue-specific epigenetic clocks when applied consistently within a given tissue. Rather, they highlight a limitation in extrapolation. Clocks trained in one tissue cannot be assumed to report whole-organism ageing when applied

Chapter 4 - Data Availability

across tissues or individuals. Within-tissue clocks remain well suited for estimating relative or chronological age in a defined biological context, whereas multi-tissue clocks trade tissue-level specificity for robustness by capturing conserved, cross-tissue components of epigenetic ageing. In summary, epigenetic ageing across zebrafish somatic tissues is neither purely systemic nor entirely tissue-specific. By explicitly benchmarking clock architectures across tissues, this work provides a conceptual and methodological framework for the design of epigenetic biomarkers that accommodate biological heterogeneity while retaining interpretability, a principle likely to generalise across vertebrate systems.

4.6. Data Availability

All sequencing data, including RRBS data, generated for this project are available under accession (under embargo until acceptance for publication), <https://www.ebi.ac.uk/ena/browser/view/PRJEB107117>. All scripts and supplementary data are available at the git repository : <https://github.com/asdfjohnny1/cross-tissue-ageing-comparison.git>

4.7. Contributions

The studies were conceptualised and designed by S.I , J.C.D.C.

Data analysis and visualisation was carried out by J.C.D.C. ,N.R

Samples were collected and processed by J.C.D.C, N.O

All code was conceived and generated by J.C.D.C. , N.R.

Supervision was provided by S.I., R.R, A.M

Chapter 4 - Ethics declaration

Funding was provided by S.I., R.R

The first draft of the manuscript was written by J.C.D.C. and S.I.

4.8. Ethics declaration

Competing interests: All authors declare no conflict or competing interests

All zebrafish work was completed in accordance with institutional, local and national laws and governing bodies.

4.9. Extended data

Appendix Chapter 4 for supplementary methods, figures, tables and legends.

All files and processed data files are available on the GitHub page:
<https://github.com/asdfjohnny1/cross-tissue-ageing-comparison.git>

4.10. Abbreviations

Δ -Age – Delta-Age

DNA – DNA Methylation

MAE – Mean Average Error

TE – Transposable Elements

LMM – Linear Mixed Model

SE – Standard Error

4.11. References

ALEXANDROV, L. B., JONES, P. H., WEDGE, D. C., SALE, J. E., CAMPBELL, P. J., NIK-ZAINAL, S. & STRATTON, M. R. 2015. Clock-like mutational processes in human somatic cells. *Nature Genetics*, 47, 1402-1407.

BAE, T., TOMASINI, L., MARIANI, J., ZHOU, B., ROYCHOWDHURY, T., FRANJIC, D., PLETIKOS, M., PATTNI, R., CHEN, B.-J., VENTURINI, E., RILEY-GILLIS, B., SESTAN, N., URBAN, A. E., ABYZOV, A. & VACCARINO, F. M. 2018. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science*, 359, 550-555.

BAHAR, R., HARTMANN, C. H., RODRIGUEZ, K. A., DENNY, A. D., BUSUTTIL, R. A., DOLLÉ, M. E. T., CALDER, R. B., CHISHOLM, G. B., POLLOCK, B. H., KLEIN, C. A. & VIJG, J. 2006. Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature*, 441, 1011-1014.

BAKER, D. J., WIJSHAKE, T., TCHKONIA, T., LEBRASSEUR, N. K., CHILDS, B. G., VAN DE SLUIS, B., KIRKLAND, J. L. & VAN DEURSEN, J. M. 2011. Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. *Nature*, 479, 232-236.

BARKER, N. 2014. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology*, 15, 19-33.

BLOKZIJL, F., DE LIGT, J., JAGER, M., SASSELLI, V., ROERINK, S., SASAKI, N., HUCH, M., BOYMANS, S., KUIJK, E., PRINS, P., NIJMAN, I. J., MARTINCORENA, I., MOKRY, M., WIEGERINCK, C. L., MIDDENDORP, S., SATO, T., SCHWANK, G., NIEUWENHUIS, E. E. S., VERSTEGEN, M. M. A., VAN DER LAAN, L. J. W., DE

Chapter 4 - References

- JONGE, J., IJZERMANS, J. N. M., VRIES, R. G., VAN DE WETERING, M., STRATTON, M. R., CLEVERS, H., CUPPEN, E. & VAN BOXTEL, R. 2016. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature*, 538, 260-264.
- CHRUSCIEL, J. H., ORSO, R., DE MATTOS, B. P., FRIES, G. R., KRISTENSEN, C. H., GRASSI-OLIVEIRA, R. & VIOLA, T. W. 2022. A systematic review and meta-analysis of epigenetic clocks in schizophrenia. *Schizophr Res*, 246, 172-174.
- COFFMAN, J. A., RIEGER, S., ROGERS, A. N., UPDIKE, D. L. & YIN, V. P. 2016. Comparative biology of tissue repair, regeneration and aging. *npj Regenerative Medicine*, 1, 16003.
- DA SILVA-ÁLVAREZ, S., GUERRA-VARELA, J., SOBRIDO-CAMEÁN, D., QUELLE, A., BARREIRO-IGLESIAS, A., SÁNCHEZ, L. & COLLADO, M. 2020. Cell senescence contributes to tissue regeneration in zebrafish. *Aging Cell*, 19.
- DE MAGALHÃES, J. P. & CHURCH, G. M. 2005. Genomes optimize reproduction: aging as a consequence of the developmental program. *Physiology*, 20, 252-259.
- DUONG, P., RODRIGUEZ-PARKS, A., KANG, J. & MURPHY, P. J. 2024. CUT&Tag Applied to Zebrafish Adult Tail Fins Reveals a Return of Embryonic H3K4me3 Patterns During Regeneration. Springer Science and Business Media LLC.
- FIELD, A. E., ROBERTSON, N. A., WANG, T., HAVAS, A., IDEKER, T. & ADAMS, P. D. 2018. DNA Methylation Clocks in Aging: Categories, Causes, and Consequences. *Molecular Cell*, 71, 882-895.
- GLADYSHEV, V. N., KRITCHEVSKY, S. B., CLARKE, S. G., CUERVO, A. M., FIEHN, O., DE MAGALHÃES, J. P., MAU, T., MAES, M., MORITZ, R. L., NIEDERNHOFER,

Chapter 4 - References

- L. J., VAN SCHAFTINGEN, E., TRANAH, G. J., WALSH, K., YURA, Y., ZHANG, B. & CUMMINGS, S. R. 2021. Molecular damage in aging. *Nature Aging*, 1, 1096-1106.
- HANNUM, G., GUINNEY, J., ZHAO, L., ZHANG, L., HUGHES, G., SADDA, S., KLOTZLE, B., BIBIKOVA, M., FAN, J.-B., GAO, Y., DECONDE, R., CHEN, M., RAJAPAKSE, I., FRIEND, S., IDEKER, T. & ZHANG, K. 2013. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell*, 49, 359-367.
- HIGGINS-CHEN, A. T., THRUSH, K. L., WANG, Y., MINTEER, C. J., KUO, P.-L., WANG, M., NIIMI, P., STURM, G., LIN, J., MOORE, A. Z., BANDINELLI, S., VINKERS, C. H., VERMETTEN, E., RUTTEN, B. P. F., GEUZE, E., OKHUIJSEN-PFEIFER, C., VAN DER HORST, M. Z., SCHREITER, S., GUTWINSKI, S., LUYKX, J. J., PICARD, M., FERRUCCI, L., CRIMMINS, E. M., BOKS, M. P., HÄGG, S., HUSELIGER, T. T. & LEVINE, M. E. 2022. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nature Aging*, 2, 644-661.
- HORVATH, S. 2013. DNA methylation age of human tissues and cell types. *Genome Biology*, 14, R115.
- HORVATH, S. & RAJ, K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*, 19, 371-384.
- HOWE, K., CLARK, M. D., TORROJA, C. F., TORRANCE, J., BERTHELOT, C., MUFFATO, M., COLLINS, J. E., HUMPHRAY, S., MCLAREN, K., MATTHEWS, L., MCLAREN, S., SEALY, I., CACCAMO, M., CHURCHER, C., SCOTT, C., BARRETT, J. C., KOCH, R., RAUCH, G.-J., WHITE, S., CHOW, W., KILIAN, B., QUINTAIS, L. T., GUERRA-ASSUNÇÃO, J. A., ZHOU, Y., GU, Y., YEN, J., VOGEL, J.-H., EYRE, T., REDMOND, S., BANERJEE, R., CHI, J., FU, B., LANGLEY, E., MAGUIRE, S. F., LAIRD, G. K., LLOYD, D., KENYON, E., DONALDSON, S., SEHRA, H., ALMEIDA-KING, J., LOVELAND, J., TREVANION, S., JONES, M., QUAIL, M., WILLEY, D.,

Chapter 4 - References

HUNT, A., BURTON, J., SIMS, S., MCLAY, K., PLUMB, B., DAVIS, J., CLEE, C., OLIVER, K., CLARK, R., RIDDLE, C., ELLIOTT, D., THREADGOLD, G., HARDEN, G., WARE, D., BEGUM, S., MORTIMORE, B., KERRY, G., HEATH, P., PHILLIMORE, B., TRACEY, A., CORBY, N., DUNN, M., JOHNSON, C., WOOD, J., CLARK, S., PELAN, S., GRIFFITHS, G., SMITH, M., GLITHERO, R., HOWDEN, P., BARKER, N., LLOYD, C., STEVENS, C., HARLEY, J., HOLT, K., PANAGIOTIDIS, G., LOVELL, J., BEASLEY, H., HENDERSON, C., GORDON, D., AUGER, K., WRIGHT, D., COLLINS, J., RAISEN, C., DYER, L., LEUNG, K., ROBERTSON, L., AMBRIDGE, K., LEONGAMORNLEET, D., MCGUIRE, S., GILDERTHORP, R., GRIFFITHS, C., MANTHRAVADI, D., NICHOL, S., BARKER, G., et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, 496, 498-503.

JENKINSON, G., PUJADAS, E., GOUTSIAS, J. & FEINBERG, A. P. 2017. Potential energy landscapes identify the information-theoretic nature of the epigenome. *Nature Genetics*, 49, 719-729.

JYLHÄVÄ, J., PEDERSEN, N. L. & HÄGG, S. 2017. Biological Age Predictors. *EBioMedicine*, 21, 29-36.

KIRKWOOD, T. B. L. 1977. Evolution of ageing. *Nature*, 270, 301-304.

KIZIL, C., KASLIN, J., KROEHNE, V. & BRAND, M. 2012. Adult neurogenesis and brain regeneration in zebrafish. *Dev Neurobiol*, 72, 429-61.

LEE, H. J., HOU, Y., CHEN, Y., DAILEY, Z. Z., RIDDIOUGH, A., JANG, H. S., WANG, T. & JOHNSON, S. L. 2020. Regenerating zebrafish fin epigenome is characterized by stable lineage-specific DNA methylation and dynamic chromatin accessibility. *Genome Biology*, 21.

LEVINE, M. E., LU, A. T., QUACH, A., CHEN, B. H., ASSIMES, T. L., BANDINELLI, S., HOU, L., BACCARELLI, A. A., STEWART, J. D., LI, Y., WHITSEL, E. A., WILSON, J. G., REINER, A. P., AVIV, A., LOHMAN, K., LIU, Y., FERRUCCI, L. & HORVATH, S.

Chapter 4 - References

2018. An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, 10, 573-591.
- LÓPEZ-OTÍN, C., BLASCO, M. A., PARTRIDGE, L., SERRANO, M. & KROEMER, G. 2013. The hallmarks of aging. *Cell*, 153, 1194-217.
- LU, A. T., QUACH, A., WILSON, J. G., REINER, A. P., AVIV, A., RAJ, K., HOU, L., BACCARELLI, A. A., LI, Y., STEWART, J. D., WHITSEL, E. A., ASSIMES, T. L., FERRUCCI, L. & HORVATH, S. 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, 11, 303-327.
- LU, Y., BROMMER, B., TIAN, X., KRISHNAN, A., MEER, M., WANG, C., VERA, D. L., ZENG, Q., YU, D., BONKOWSKI, M. S., YANG, J. H., ZHOU, S., HOFFMANN, E. M., KARG, M. M., SCHULTZ, M. B., KANE, A. E., DAVIDSOHN, N., KOROBKINA, E., CHWALEK, K., RAJMAN, L. A., CHURCH, G. M., HOCHEDLINGER, K., GLADYSHEV, V. N., HORVATH, S., LEVINE, M. E., GREGORY-KSANDER, M. S., KSANDER, B. R., HE, Z. & SINCLAIR, D. A. 2020. Reprogramming to recover youthful epigenetic information and restore vision. *Nature*, 588, 124-129.
- MAKLAKOV, A., ALEXEI & IMMLER, S. 2016. The Expensive Germline and the Evolution of Ageing. *Current Biology*, 26, R577-R586.
- MAKLAKOV, A. A. & CHAPMAN, T. 2019. Evolution of ageing as a tangle of trade-offs: energy versus function. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20191604.
- MARTINCORENA, I., ROSHAN, A., GERSTUNG, M., ELLIS, P., VAN LOO, P., MCLAREN, S., WEDGE, D. C., FULLAM, A., ALEXANDROV, L. B., TUBIO, J. M., STEBBINGS, L., MENZIES, A., WIDAA, S., STRATTON, M. R., JONES, P. H. & CAMPBELL, P. J. 2015. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*, 348, 880-886.

Chapter 4 - References

MARTINEZ-JIMENEZ, C. P., ELING, N., CHEN, H. C., VALLEJOS, C. A., KOLODZIEJCZYK, A. A., CONNOR, F., STOJIC, L., RAYNER, T. F., STUBBINGTON, M. J. T., TEICHMANN, S. A., DE LA ROCHE, M., MARIONI, J. C. & ODOM, D. T. 2017. Aging increases cell-to-cell transcriptional variability upon immune stimulation. *Science*, 355, 1433-1436.

MAYNE, B., KORBIE, D., KENCHINGTON, L., EZZY, B., BERRY, O. & JARMAN, S. 2020. A DNA methylation age predictor for zebrafish. *Aging*, 12, 24817-24835.

MOORE, L., CAGAN, A., COORENS, T. H. H., NEVILLE, M. D. C., SANGHVI, R., SANDERS, M. A., OLIVER, T. R. W., LEONGAMORNLEET, D., ELLIS, P., NOORANI, A., MITCHELL, T. J., BUTLER, T. M., HOOKS, Y., WARREN, A. Y., JORGENSEN, M., DAWSON, K. J., MENZIES, A., O'NEILL, L., LATIMER, C., TENG, M., VAN BOXTEL, R., IACOBUZIO-DONAHUE, C. A., MARTINCORENA, I., HEER, R., CAMPBELL, P. J., FITZGERALD, R. C., STRATTON, M. R. & RAHBARI, R. 2021. The mutational landscape of human somatic and germline cells. *Nature*, 597, 381-386.

NALAPAREDDY, K., ZHENG, Y. & GEIGER, H. 2022. Aging of intestinal stem cells. *Stem Cell Reports*, 17, 734-740.

OCAMPO, A., REDDY, P., MARTINEZ-REDONDO, P., PLATERO-LUENGO, A., HATANAKA, F., HISHIDA, T., LI, M., LAM, D., KURITA, M., BEYRET, E., ARAOKA, T., VAZQUEZ-FERRER, E., DONOSO, D., ROMAN, J. L., XU, J., RODRIGUEZ ESTEBAN, C., NUÑEZ, G., NUÑEZ DELICADO, E., CAMPISTOL, J. M., GUILLEN, I., GUILLEN, P. & IZPISUA BELMONTE, J. C. 2016. In Vivo Amelioration of Age-Associated Hallmarks by Partial Reprogramming. *Cell*, 167, 1719-1733.e12.

POSS, K. D., WILSON, L. G. & KEATING, M. T. 2002. Heart regeneration in zebrafish. *Science*, 298, 2188-90.

RAKYAN, V. K., DOWN, T. A., MASLAU, S., ANDREW, T., YANG, T.-P., BEYAN, H., WHITTAKER, P., MCCANN, O. T., FINER, S., VALDES, A. M., LESLIE, R. D., DELOUKAS, P. & SPECTOR, T. D. 2010. Human aging-associated DNA

Chapter 4 - References

hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Research*, 20, 434-439.

RAUDVERE, U., KOLBERG, L., KUZMIN, I., ARAK, T., ADLER, P., PETERSON, H. & VILO, J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*, 47, W191-w198.

SEN, P, SHAH, P.P., NATIVO,R. and BERGER, S.L. 2016. Epigenetic mechanisms of longevity and ageing. *Cell*166(4),pp.822-839.,

SLIEKER, R. C., RELTON, C. L., GAUNT, T. R., SLAGBOOM, P. E. & HEIJMANS, B. T. 2018. Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception. *Epigenetics & Chromatin*, 11.

SMITH, A., HUBLEY, R. & GREEN, P. 2013-2015. RepeatMasker Open-4.0.

STEWART, S., GOMEZ, W., ALAN, ARMSTRONG, E., BENJAMIN, HENNER, A. & STANKUNAS, K. 2014. Sequential and Opposing Activities of Wnt and BMP Coordinate Zebrafish Bone Regeneration. *Cell Reports*, 6, 482-498.

STOEGER, T., GRANT, R. A., MCQUATTIE-PIMENTEL, A. C., ANEKALLA, K. R., LIU, S. S., TEJEDOR-NAVARRO, H., SINGER, B. D., ABDALA-VALENCIA, H., SCHWAKE, M., TETREAULT, M.-P., PERLMAN, H., BALCH, W. E., CHANDEL, N. S., RIDGE, K. M., SZNAJDER, J. I., MORIMOTO, R. I., MISHARIN, A. V., BUDINGER, G. R. S. & NUNES AMARAL, L. A. 2022. Aging is associated with a systemic length-associated transcriptome imbalance. *Nature Aging*, 2, 1191-1206.

STUBBS, T. M., BONDER, M. J., STARK, A.-K., KRUEGER, F., VON MEYENN, F., STEGLE, O. & REIK, W. 2017. Multi-tissue DNA methylation age predictor in mouse. *Genome Biology*, 18.

TERMAN, A. & BRUNK, U. T. 2006. Oxidative stress, accumulation of biological 'garbage', and aging. *Antioxid Redox Signal*, 8, 197-204.

Chapter 4 - References

TESCHENDORFF, A. E., WEST, J. & BECK, S. 2013. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Human Molecular Genetics*, 22, R7-R15.

TOMASETTI, C. & VOGELSTEIN, B. 2015. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science*, 347, 78-81.

VAIDYA, H., JEONG, H. S., KEITH, K., MAEGAWA, S., CALENDO, G., MADZO, J., JELINEK, J. & ISSA, J.-P. J. 2023. DNA methylation entropy as a measure of stem cell replication and aging. *Genome Biology*, 24.

YAMAMOTO, R., CHUNG, R., VAZQUEZ, J. M., SHENG, H., STEINBERG, P. L., IOANNIDIS, N. M. & SUDMANT, P. H. 2022. Tissue-specific impacts of aging and genetics on gene expression patterns in humans. *Nature Communications*, 13.

YUN, M. H., DAVAAPIL, H. & BROCKES, J. P. 2015. Recurrent turnover of senescent cells during regeneration of a complex structure. *Elife*, 4.

ZAMPIERI, M., CICCARONE, F., CALABRESE, R., FRANCESCHI, C., BÜRKLE, A. & CAIAFA, P. 2015. Reconfiguration of DNA methylation in aging. *Mech Ageing Dev*, 151, 60-70.

ZHOU, Y., HE, F., PU, W., GU, X., WANG, J. & SU, Z. 2020. The Impact of DNA Methylation Dynamics on the Mutation Rate During Human Germline Development. *G3 Genes/Genomes/Genetics*, 10, 3337-3346.

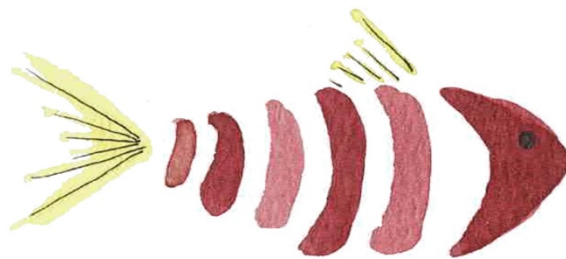
CHAPTER FIVE

Germline ablation affects epigenetic somatic ageing in zebrafish

Jean-Charles de Coriolis¹, Raheleh Rahabari², Alexei Maklakov, Simone Immler¹

¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, NR4 6TX, Norwich, UK

² Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 ISA, UK



5. Chapter 5

5.1. Abstract

Ageing is characterised by a progressive loss of molecular and regulatory homeostasis, yet how interactions between the germline and somatic tissues shape somatic ageing dynamics remains poorly understood. While evolutionary theory predicts that germline status can influence somatic maintenance, direct longitudinal molecular evidence linking germline–soma interactions to ageing trajectories in vertebrates is limited. Here, we investigate how germline ablation alters somatic epigenetic ageing using longitudinal DNA methylation profiling in zebrafish. Using reduced representation bisulphite sequencing, we quantified age-associated DNA methylation changes across individual lifespans in germline-carrying (GLC) and germline-free (GLF) fish. Tissue-specific epigenetic clocks based on fin clip methylation were used to estimate epigenetic age and its deviation from chronological age, enabling direct comparison of ageing trajectories between groups. GLF and GLC fish exhibited significantly different epigenetic ageing dynamics, with GLF fish showing reduced epigenetic age early in life relative to GLC fish. In both groups, epigenetic ageing decelerated with advancing chronological age; however, GLF fish displayed altered late-life dynamics characterised by a weaker age-related deceleration. At the CpG level, age-associated methylation changes were concentrated near genes involved in developmental regulation, chromatin organisation, and transcriptional control, rather than core metabolic or structural housekeeping functions. In late life, GLF fish showed directional divergence away from intermediate methylation states, consistent with reduced epigenetic stability, whereas GLC fish exhibited comparatively stabilised methylation patterns. These molecular differences were accompanied by behavioural divergence in late life, with GLF fish displaying burst-like, high-

intensity locomotor activity relative to the more sustained movement observed in GLC fish. Together, these findings demonstrate that the germline shapes not only the magnitude but also the temporal organisation and stability of somatic epigenetic ageing. Germline ablation does not uniformly slow ageing; instead, it confers early-life molecular youth at the cost of altered late-life regulatory stability. By integrating longitudinal epigenetic and behavioural analyses, this work identifies the germline as a general modulator of somatic ageing and shows that its influence on somatic ageing dynamics operates across the vertebrate lifespan.

5.2. Introduction

Ageing is characterised by a progressive loss of molecular and cellular homeostasis that emerges from the cumulative effects of DNA damage, replication error, metabolic stress and declining repair capacity across tissues (López-Otín et al., 2013). Endogenous processes such as oxidative phosphorylation generate reactive oxygen species that induce base modifications and DNA strand breaks, while DNA replication itself introduces mismatches and replication stress even under normal physiological conditions (Jackson and Bartek, 2009). Although cells deploy multiple repair pathways, including base-excision repair, nucleotide-excision repair, mismatch repair and homologous recombination, the efficiency, fidelity and coordination of these systems decline with age. This leads to the accumulation of somatic mutations and associated disruption of gene regulatory control, manifested as reduced coordination of transcriptional and epigenetic programmes, increased variation in regulatory states, and diminished fidelity of stress-response activation (Lodato et al., 2018, Vijg and Suh, 2013). Importantly, these processes do not affect all cell types equally.

Chapter 5 - Introduction

Long-lived somatic tissues experience progressive challenges to regulatory homeostasis with age, whereas germline perturbation can elicit organism-wide protective responses, highlighting distinct regulatory roles of germline and soma (Ermolaeva et al., 2013).

The asymmetry between germline and soma has long been central to evolutionary theories of ageing. The disposable soma theory proposes that limited energetic resources are preferentially allocated to reproduction at the expense of somatic maintenance, resulting in gradual somatic deterioration over time (Kirkwood, 1977). The expensive germline hypothesis refines this view by arguing that germ cells actively sequester molecular maintenance capacity, including DNA repair, chromatin regulation and proteostasis machinery, thereby indirectly accelerating somatic ageing (Maklakov and Immler, 2016). In contrast, the developmental theory of ageing emphasises the persistence or misregulation of developmental programmes beyond their adaptive window, proposing that late-life decline reflects regulatory imbalance rather than the simple accumulation of stochastic damage (Maklakov and Chapman, 2019). While these frameworks differ in emphasis, they converge on the view that germline-soma interactions are central to shaping ageing, even though they conceptualise this influence through distinct evolutionary and mechanistic lenses.

Experimental studies have provided strong empirical support for a regulatory role of the germline in shaping somatic ageing. In *Caenorhabditis elegans*, selective ablation of germ cells extends lifespan and enhances stress resistance, whereas removal of the entire gonad, including somatic gonadal tissue, fails to produce comparable effects, demonstrating that longevity arises from germline-specific signalling rather than from loss of reproduction per se (Arantes-Oliveira et al., 2002, Hsin and Kenyon, 1999). Similar dissociations have been reported in *Drosophila melanogaster*, where germline proliferation negatively regulates lifespan through insulin-like

signalling pathways independently of egg production (Yamawaki et al., 2008). Together, these findings highlight the systemic regulatory effects that germ cells exert on somatic physiology, rather than acting solely as sinks for energetic resources.

At the molecular level, germline-derived signals modulate conserved nutrient-sensing and stress-response pathways, including insulin/IGF-1 signalling (Templeman and Murphy, 2018) and mTOR (Saxton and Sabatini, 2017), which in turn regulate autophagy, cell-cycle progression, proteostasis and DNA repair. In *C. elegans*, DNA damage restricted to germ cells is sufficient to induce organism-wide stress resistance via innate immune signalling, illustrating direct communication between germline integrity and somatic maintenance (Ermolaeva et al., 2013). Vertebrate studies similarly suggest that reproductive or germline-associated signalling can influence somatic transcriptional programmes linked to DNA repair, stress response, and metabolism, particularly under conditions of physiological stress (Chen et al., 2020, Mitchell et al., 2024). Notably, these effects are often context- and time-dependent, suggesting that germline signalling shapes the dynamics of ageing rather than imposing a uniform ageing rate.

From this perspective, ageing-related molecular change is increasingly understood as a process of coordinated regulatory alteration rather than diffuse epigenomic erosion. DNA methylation (DNAm), a key epigenetic modification involved in transcriptional regulation and cellular identity, exhibits reproducible age-associated changes across tissues and species (Horvath, 2013, Horvath and Raj, 2018). These changes in methylation are not randomly distributed across the genome; instead, they are frequently clustered across neighbouring CpG sites and enriched near genes involved in development, differentiation and transcriptional regulation (Bell et al., 2019). Such patterns support the view that ageing reflects the progressive destabilisation of regulatory networks established during development,

consistent with the developmental theory of ageing (Maklakov and Chapman, 2019).

DNAm-based epigenetic clocks provide a quantitative framework for capturing these coordinated age-associated changes. By integrating methylation status across selected CpG sites which correlate with chronological age, epigenetic clocks summarise complex regulatory dynamics into a single predictive measure (Horvath, 2013). Deviations between predicted and chronological age, termed Δ -Age (predicted – chronological age) or epigenetic age acceleration, are widely interpreted as indicators of altered biological ageing (Horvath and Raj, 2018). However, Δ -Age does not reflect a single underlying process. One source of Δ -Age is a change in the rate at which age-associated methylation accumulates, such that methylation at clock-associated CpG sites changes more steeply or more shallowly across adulthood than expected (Horvath, 2013, Belsky et al., 2015). Alternatively, Δ -Age can arise from shifts in the timing of age-related regulatory transitions, for example, when coordinated methylation changes emerge abruptly in late life (Snir et al., 2019, Levine et al., 2018). Δ -Age can also arise from increased stochastic divergence in methylation states at individual CpG sites, where sites that are normally maintained within a narrow regulatory range progressively drift away from intermediate or coordinated methylation levels, reducing the coherence of clock signals without necessarily altering mean methylation (Snir et al., 2019, Bell et al., 2019). Consequently, Δ -Age reflects not only differences in the magnitude of age-associated change but also alterations in the temporal organisation and regulatory stability of methylation dynamics. Interpreting Δ -Age, therefore, requires more than a snapshot at a single time point. Instead, it requires longitudinal analyses that track within-individual methylation trajectories over time, rather than reliance on cross-sectional comparisons alone.

Chapter 5 - Introduction

Despite substantial theoretical and experimental interest, direct longitudinal analyses of epigenetic ageing remains rare, particularly in vertebrates. While DNA methylation has been longitudinally measured within individuals, and epigenetic clocks have been applied to repeated measures, primarily in human blood, these studies are observational, and not designed to interrogate experimental models, such as germline-soma interactions (Marioni et al., 2015, Snir et al., 2019). Consequently, most existing work relies on cross-sectional epigenetic measurements or indirect inference from lifespan, limiting insight into regulatory stability, late-life divergence, and stochastic methylation behaviour (Bell et al., 2019, Horvath and Raj, 2018). Moreover, few studies explicitly examine how modelling choices in epigenetic clock construction influence sensitivity to biologically meaningful perturbations, including changes in germline status.

In this chapter, we investigated how germline status shapes the longitudinal dynamics of somatic epigenetic ageing. We conducted a longitudinal study in male zebrafish (*Danio rerio*), comparing germline-carrying (GLC) and germline-free (GLF) individuals from 4 to 22 months of age, to determine whether germline ablation alters the timing, coordination and stability of epigenetic ageing trajectories. Repeated DNA methylation measurements from fin clips were used to construct tissue-specific epigenetic clocks under multiple CpG aggregation strategies, enabling within individual tracking of epigenetic age over time. We additionally examined CpG-level methylation dynamics to assess whether germline status influences stochastic methylation drift at regulatory loci. Finally, late-life behavioural phenotypes were quantified to test whether molecular differences in epigenetic ageing are accompanied by functional divergence.

5.3. Material and methods

5.3.1. Fish Lines and Maintenance

Experimental zebrafish (*Danio rerio*) used in this study were derived from a laboratory-maintained *dnd1* mutant line generated and bred in the Controlled Environment Facility at the University of East Anglia, UK. This line was originally established on an AB genetic background obtained from the Zebrafish International Resource Centre (ZIRC; University of Oregon, Eugene, USA) and has since been maintained under a strict outbreeding programme to preserve genetic diversity.

Fish were housed in 3L tanks under a 12: 12 light : dark photoperiod at 28°C. Adult fish were fed three times daily with live *Artemia* (ZM System) and dried food (Sparos, Adult zebrafish diet, nutrition in aquaculture). Juvenile fish were fed four times daily to support growth and maintain consistent nutritional conditions across developmental stages.

5.3.2. Germline-free fish

We used a CRISPR/Cas9 knockout strain of the germline-specific gene dead-end (*dnd1*). *dnd1* is essential for primordial germ cell (PGC) specification and survival in zebrafish and is expressed predominantly in PGCs during early embryogenesis (Weidinger et al., 2003). Loss of *dnd1* disrupts PGC maintenance and migration, resulting in complete germ cell ablation (Weidinger et al., 2003; Slanchev et al., 2005). In zebrafish, the absence of germ cells biases sex determination towards male development, producing sterile males with somatic testes, male secondary sexual characteristics, and typical male courtship behaviour despite an inability to fertilise eggs (Slanchev et al., 2005). Fish for this study were generated by crossing *dnd1* heterozygous carriers. Homozygous *dnd1* (GLF) males were compared with

Chapter 5 - Material and methods

full-sibling homozygous wild-type (GLC) males, ensuring a shared genetic background between experimental groups.

5.3.3. Sample Collection

A total of 27 fish were followed, yielding 96 samples after attrition/missing timepoints cover a period of 22 months at six-month intervals (4, 10, 16, 22), resulting in four distinct time points. Fin clips were obtained from the same fish at each time point (Table S23).

5.3.4. DNA extractions

Genomic DNA was extracted from all fin clips using a modified Qiagen DNeasy Blood and Tissue Kit protocol (Qiagen, 69506). In short, samples were digested in proteinase K at 56 °C for 1-2h with intermittent vortexing until complete lysis was achieved. Lysates were processed through silica spin columns according to the manufacturer's instructions, and DNA was eluted in 30 µL EB buffer. DNA concentrations and purity were evaluated using a Qbit fluorometer and a Nanodrop spectrophotometer prior to library preparation.

5.3.5. Library preparation and sequencing

A total of 96 RRBS libraries were prepared at the University of East Anglia, using the TECAN Ovation® RRBS methyl-Seq System Kit (TECAN, 9522-A01), following an optimised version of the manufacturer's protocol (see detailed procedure at <https://www.protocols.io/view/ovation-rrbs-methyl-seq-library-preparation-protocol-kxygx3kydg8j/v1>). Briefly, genomic DNA was digested with MspI, ligated to methylated adapters, end-repaired, and subjected to bisulphite conversion. Libraries were then purified via magnetic beads, PCR amplified, and size selected. Fragment size distribution and library quality

Chapter 5 - Material and methods

were assessed using an Agilent TapeStation, and final quantification was performed prior to equimolar pooling.

Libraries were sequenced on an Illumina NovaSeq X platform using a 10B flow cell over three lanes, producing 2 x 150 bp paired-end reads, generating approximately 2.5 billion read pairs per lane.

5.3.6. Data trimming, alignment and methylation Extraction

Raw reads were processed using the NuGEN technologies pipeline (<https://www.nugen.com/products/nugen-nuget-rrbs>). Adapter and low-quality bases were trimmed with Trim Galore v 0.6.5 (https://Delta-Age.bioinformatics.babraham.ac.uk/projects/trim_galore/), and diversity adaptor artefacts were removed using a NuGEN-provided Python script (<https://github.com/nugentechnologies/NuMetRRBS/trimRRBSdiversityAdaptCustomers.py>). Cleaned reads were aligned to the zebrafish reference genome (GRCz11; https://Delta-Age.ncbi.nlm.nih.gov/datasets/genome/GCF_000002035.6/) using Bismark v 0.25.1 (Krueger and Andrews, 2011), and duplicate reads were identified and marked with the NuDup tool in NuGEN (<https://github.com/tecangenomics/nudup>).

Methylation calling was performed using the Bismark methylation extractor (Krueger and Andrews, 2011), generating cytosine methylation reports in “.cov” format. Methylation metrics (methylated and unmethylated counts per CpG) were used as input for downstream analyses.

5.3.7. Methylation data pre-processing

Methylation data were imported into R v 4.3.1 (R Core Team, 2022) using the methylKit package v 1.3 (Akalin et al., 2012). Coverage files were read using

the `methRead()` function, applying a minimum per-CpG coverage threshold of five reads (*mincov* =5). CpG sites were retained only if present in at least 90% of samples within each tissue (*min_samples* <- *as.interger(ceiling(n_samples * 0.9))*)).

Methylation percentage was calculated as $\text{methylation percentage} = \frac{\text{methylated reads}}{\text{total reads}} \times 100$. The resulting tissue-specific methylation matrices formed the basis for subsequent clock training across all analytical configurations.

5.3.8. Epigenetic clock model training and selection

A comprehensive epigenetic clock modelling framework was applied, encompassing multiple CpG aggregation strategies, regression approaches, feature filtering schemes and parameter settings, yielding a total of 1440 distinct model configurations (Table S23). For each configuration, models were trained and evaluated using repeated random splits of individuals into training and held-out validation sets, with optimisation performed through a nested iterative procedure to ensure robust CpG selection and unbiased performance assessment. Importantly, because individuals were sampled repeatedly over time, training–validation splits were defined at the individual level, such that all time points from a given individual were assigned exclusively to either the training or validation set.

Epigenetic clock construction, optimisation and validation followed the pipeline described in detail in Chapter 3, including nested resampling, performance-based CpG selection and final model refitting using optimised feature sets. Predicted ages and model diagnostics were retained for downstream analyses. All modelling in this chapter was restricted to fin tissue, enabling direct comparison of ageing trajectories under different germline conditions without introducing tissue-specific confounding effects.

Chapter 5 - Material and methods

5.3.9. Behaviour tracking and speed analysis

At 24 months of age, for each assay, all focal individuals from a given tank were transferred together into a circular arena (235mm diameter) containing 2L of system water. Behaviour was recorded for 16 minutes using an overhead GoPro HERO4 camera. The first minute of recording was excluded to allow for acclimation.

Video data were processed using *id.Tracker.ai* v 5.1.3 (Romero-Ferrero et al., 2019) to extract per-frame x-y coordinates for each individual. From these trajectories, eight standardised locomotor and spatial metrics were derived following Kermen et al. (2020).

- (i) Total distance (mm) travelled was calculated as the cumulative Euclidean displacement between consecutive frames and converted to millimetres using a geometry-derived calibration factor of 3.
- (ii) Mean speed (mm s^{-1}) was calculated as total distance divided by assay duration.
- (iii) Maximum speed was defined as the highest speed value exceeding 1 mm s^{-1} .
- (iv) Mean Acceleration was computed from frame-to-frame displacement at 30 Hz.
- (v) Max acceleration was defined as the highest acceleration value exceeding 1 mm s^{-1} .
- (vi) Path straightness was defined as the ratio between the straight-line distance from start to end position and the total distance travelled, with values close to 1 indicating direct swimming.
- (vii) Thigmotaxis was quantified as the proportion of frames in which the fish centroid was located within 20 mm of the arena wall.
- (viii) Angular velocity ($^{\circ} \text{ s}^{-1}$) was calculated as the mean absolute angular change between successive movement vectors, converted from radians and adjusted to the frame rate.

5.4. Results

In total, 27 fish were sampled longitudinally, forming the input dataset for all downstream analyses. A total of 94 samples passed initial sequencing and alignment quality control and were retained as candidates for epigenetic clock construction. Subsequent filtering steps, including CpG coverage thresholds, CpG completeness requirements, and feature selection, were applied within each modelling configuration, such that the number of retained CpGs and resulting sample-feature matrices varied across models.

In total, 1440 distinct epigenetic clock configurations were trained, spanning multiple CpG aggregation strategies, regression frameworks, and filtering parameterisations (Table S23). Because filtering was configuration-specific, no single CpG set or feature count applied across all models. Following model training and validation, 958 models met predefined performance and quality criteria and were retained for downstream analyses. Unless stated otherwise, all results below refer to this filtered model set, with comparisons interpreted across configurations rather than assuming a shared underlying feature space.

5.4.1. CpG aggregation strategy shapes epigenetic clock performance

We first assessed how CpG aggregation strategy influenced epigenetic clock performance across all model iterations. CpG grouping had a significant effect on model performance (linear mixed-effects model, $X^2 = 9.8$, $df = 2$, $p = 0.01$; Table S24). Tile-based aggregation performed significantly worse than Cluster-based aggregation (*Estimate* = -0.05, $t = -2.8$, $p = 0.006$), whereas Cluster-based and Point-based models did not differ in performance (*Estimate* = 0.004, $t = 0.1$, $p = 0.9$; Figure 14A, table S24).

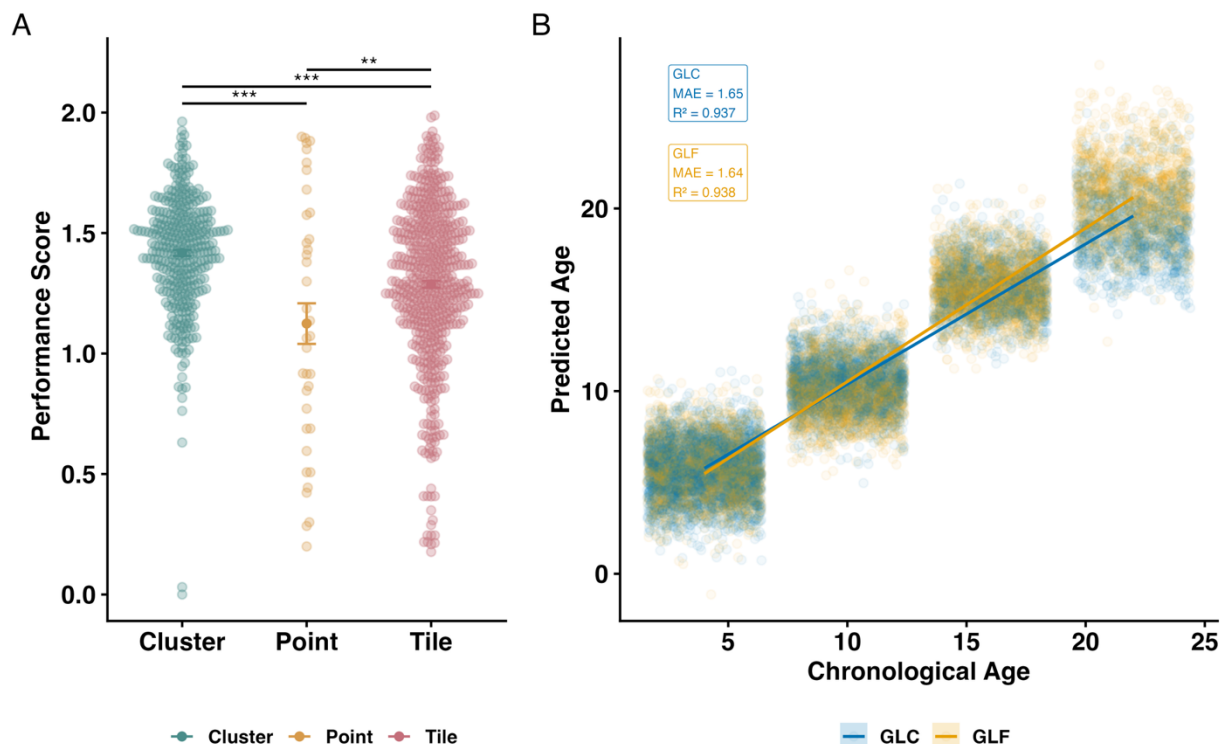


Figure 14: Epigenetic clock performance across CpG aggregation strategies.

A) Epigenetic clock performance scores (combining scaled R^2 + inverse Mean Average Error) by CpG aggregation method. Raw data is presented in beeswarm. **B)** Predicted epigenetic age plotted against chronological age, stratified by treatment. Points represent individual predictions, solid lines show linear regression fits with 95% confidence intervals and inset annotations report mean absolute error and coefficient of determination (R^2) per treatment. Points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

5.4.2. Identification and methylation dynamics of robust CpG sites

To examine the contribution of individual CpG sites to clock predictions, we identified CpGs that were repeatedly selected across 958 independently trained models. 64 CpG sites appeared in at least 10% of model runs and were therefore classified as robustly selected sites, using a selection-frequency criterion analogous to stability selection approaches for penalised regressions (Meinshausen and Bühlmann, 2010; Haury et al., 2012). Across all samples, these CpGs did not differ in overall mean methylation between GLC and GLF fish (linear mixed model, $X^2 = 2.5$, $df = 1$, $p = 0.1$; Figure 15A). However, there was a near-significant interaction between germline status and age ($X^2 = 7.0$, $df = 3$, $p = 0.07$; Table S24). Despite the marginal interaction term, post hoc analysis revealed that GLF fish at 22 months of age exhibited significantly reduced methylation at these sites compared to age-matched GLC fish (EMM, *Estimate* = -3.2, $z = -3.1$, $p = 0.002$; Figure 15B;).

Hierarchical clustering of methylation profiles further revealed a strong age-by-germline interaction, with late-life GLF samples forming a distinct cluster (Figure 15E). Permutational multivariate analysis of variance on Euclidean distances of scaled methylation profiles confirmed a significant effect of individual tissues across time ($R^2 = 0.7$, $F = 2.3$, $p = 0.02$, Table S4).

Collectively, these results indicate that robust CpG sites do not drive differences in global methylation levels but instead exhibit age-specific variation that becomes more prominent late in life.

Chapter 5 - Results

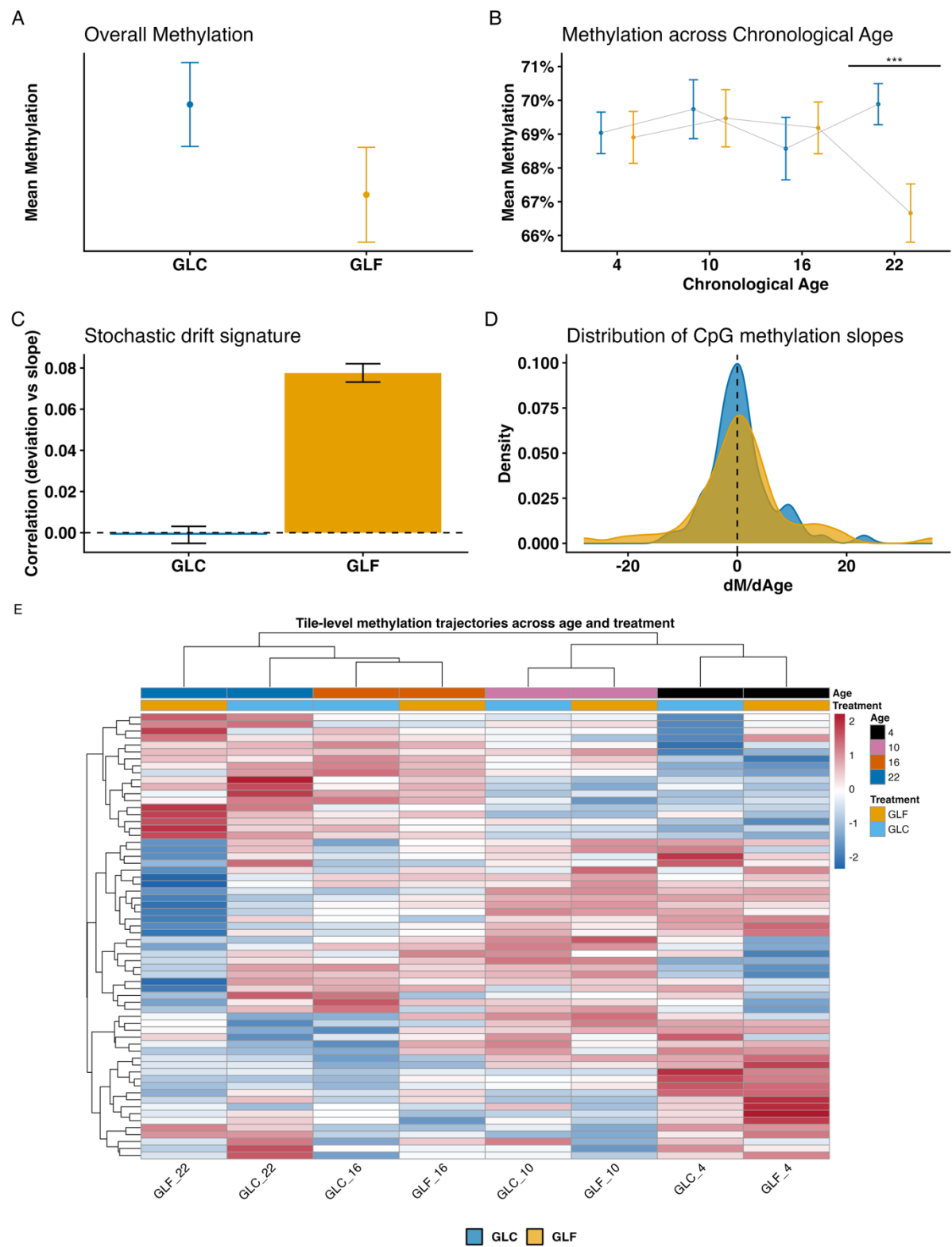


Figure 15: Age-dependent treatment effect in CpG methylation.

A) Mean methylation trajectories of CpGs shared across models, summarised by treatment. **B)** Age-dependent methylation dynamics of the same CpGs across chronological age, with locally weighted regression (LOESS) fits and 95% confidence intervals, highlighting treatment-specific deviations in methylation trajectories. **C)** Stochastic drift signature quantified as the correlation between individual CpG methylation deviations and their age-associated slopes; bars represent mean correlation $\pm$ SE across bootstrap iterations, with the dashed line indicating no association. **D)** Distribution of CpG methylation slopes (dM/dAge) per treatment, illustrating heterogeneity in age-associated methylation change. **(E)** Heatmap of age-associated methylation patterns across selected CpGs, ordered by chronological age, summarising coordinated methylation structure underlying epigenetic ageing. Points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

5.4.3. Functional context of robust CpG sites

To place these CpG sites in a functional context, each robust CpG was assigned to its nearest gene and enrichment analysis was performed. Genes associated with robust CpGs were enriched for biological processes related to tissue morphogenesis, sensory and visual development, including retinal and lens formation, and transcriptional regulation (Figure 16A). These enrichments highlight developmental regulatory networks as prominent targets of age-associated methylation change.

16% of robust CpGs overlapped with Polycomb group target regions (Figure S12), indicating that Polycomb-associated regions were overrepresented among robust CpGs, suggesting that ageing-associated methylation changes preferentially affect regulatory circuits established during development.

Chapter 5 - Results

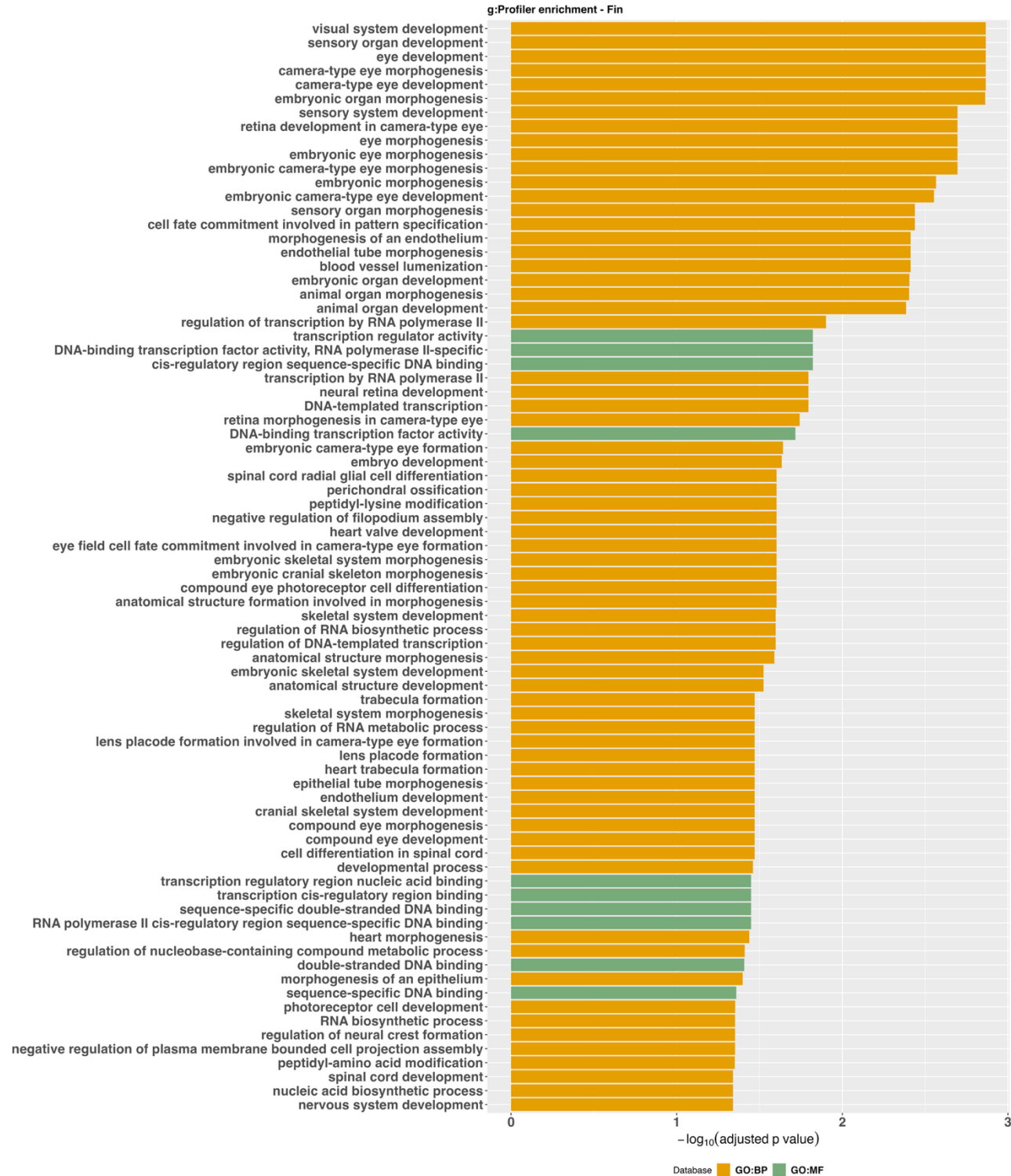


Figure 16: System-level divergence of CpG methylation in caudal fin.

A) Functional enrichment of genes associated with the top-performing CpGs, as assessed by g:Profiler. Bars indicate the $-\log_{10}$ adjusted p-value for significantly enriched Gene Ontology terms in Biological Processes (BP) and Molecular Function (MF), coloured by source database.

5.4.4. Germline ablation alters somatic epigenetic ageing trajectories

We next examined whether germline ablation influenced epigenetic ageing at the level of clock predictions. Δ -Age differed significantly between GLF and GLC fish ($X^2= 122.5$, $df = 1$, $p < 0.0001$), with GLC fish exhibiting higher overall epigenetic age relative to GLF fish (*Estimate* =1.4, $t = 4.0$, $p < 0.0001$; Figure 17A, Table S24).

Across chronological age, however, Δ -Age declined significantly in both groups ($X^2= 314.9$, $df = 1$, $p < 0.0001$), indicating reduced epigenetic ageing over time in both strains (Figure 17B, Table S24). Interestingly, this decline was significantly weaker in GLF fish than in GLC fish (*Estimate* =7.4, $t = 18.2$, $p < 0.0001$), demonstrating that germline ablation alters the temporal structure of epigenetic ageing rather than producing a simple shift in mean epigenetic age.

This late-life reversal towards relative epigenetic ageing in GLF fish parallels methylation changes at robust CpG sites, with reduced methylation coinciding with increased Δ -Age at 22 months (Figure S1A, S1B).

Chapter 5 - Results

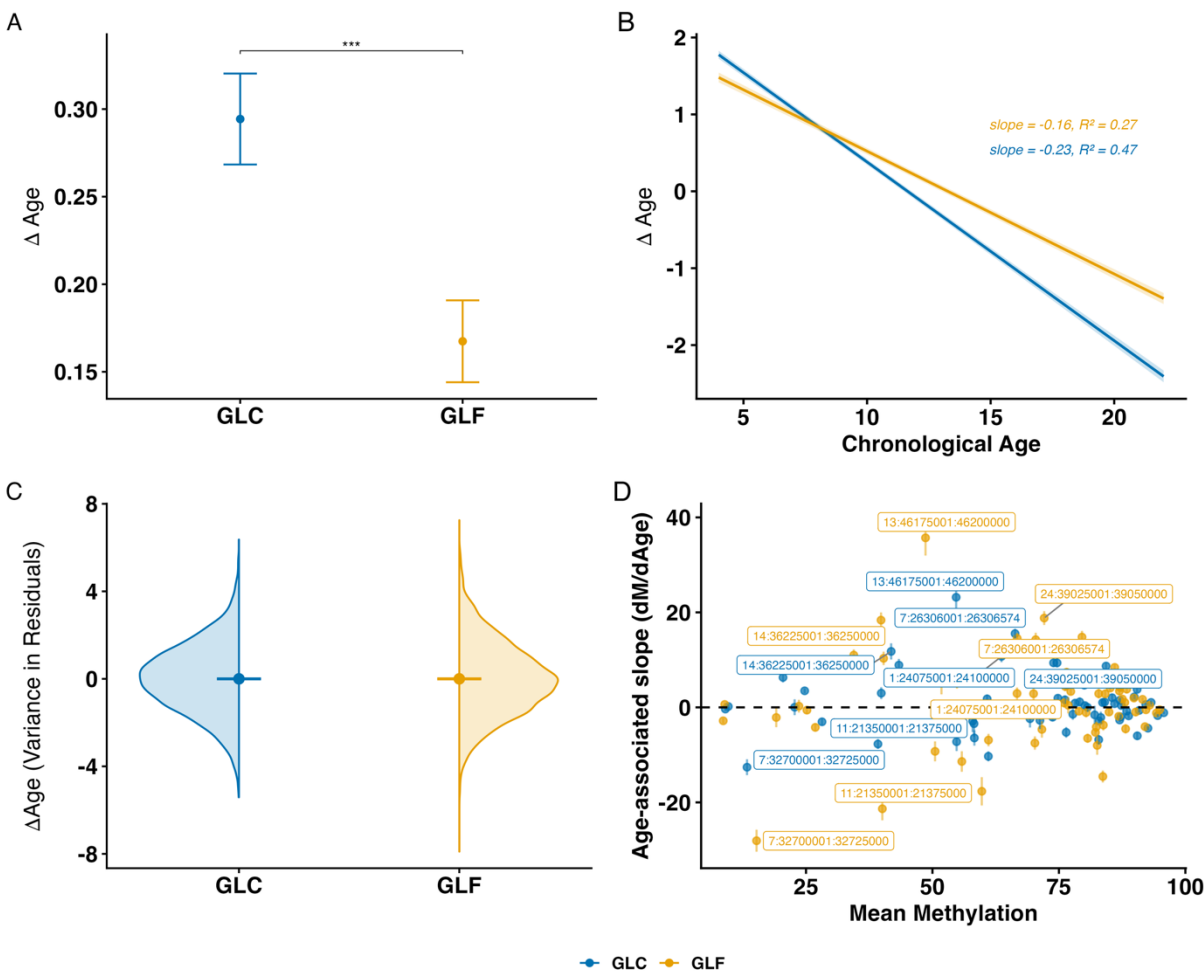


Figure 17: Germline ablation significantly alters Delta age trajectories and residual structure.

A) Mean epigenetic age deviation (Δ -Age), defined as the difference between predicted epigenetic age and chronological age, stratified between GLC and GLF strains. **B)** Δ -Age as a function of chronological age, stratified by treatment. Solid lines represent linear model fits with 95% confidence intervals; annotated slopes indicate treatment-specific ageing rates. **C)** Distribution of Δ -Age residual variance per treatment. Split violins show the distribution of individual residuals, overlaid with mean $\pm$ SE, highlighting treatment-associated differences in stochastic deviation from the fitted epigenetic ageing trajectory. **D)** Relationship between mean methylation and age-associated slope (dM/dA-ge) for individual CpGs, highlighting top-ranked CpGs driving age-related changes. Dashed horizontal line denotes no change with age; labels indicate CpGs with the steepest slopes. Points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

5.4.5. Germline ablation influences stochastic methylation drift

To assess whether germline status affects the stability of methylation regulation, we analysed longitudinal methylation drift at robust CpG sites. Although the overall effect of germline ablation on drift did not reach statistical significance (linear mixed model, $X^2 = 1.5$, $df = 1$, $p = 0.2$; Figure 15C; Table S24), quantitatively distinct longitudinal patterns were observed between groups. In GLC fish, CpG sites with extreme initial methylation values showed a tendency to revert towards intermediate methylation levels with age, consistent with stabilising regulatory control. In contrast, GLF fish exhibited the opposite pattern, with CpGs displaying progressive divergence away from intermediate methylation states over time. This pattern suggests reduced regulatory constraint and increased stochastic drift in the absence of the germline.

5.4.6. Late-life behavioural correlates of epigenetic ageing

Behavioural analysis of 24-month-old fish revealed that GLC fish were associated with reduced maximal acceleration (*Estimate* = -138,672, $t = -2.4$, $p = 0.02$), maximal speed (*Estimate* = -15,149, $t = -2.6$, $p = 0.02$), and total distance travelled (*Estimate* = -155,309, $t = -2.1$, $p = 0.05$). In contrast, these GLC fish spent more time moving overall (*Estimate* = 0.3, $t = 2.1$, $p = 0.04$; Figure 18; Table S25). This pattern indicates that relative epigenetic ageing acceleration in GLF fish is associated with burst-like, high-intensity locomotor activity, whereas GLC fish maintain more sustained, moderate movement patterns in late life.

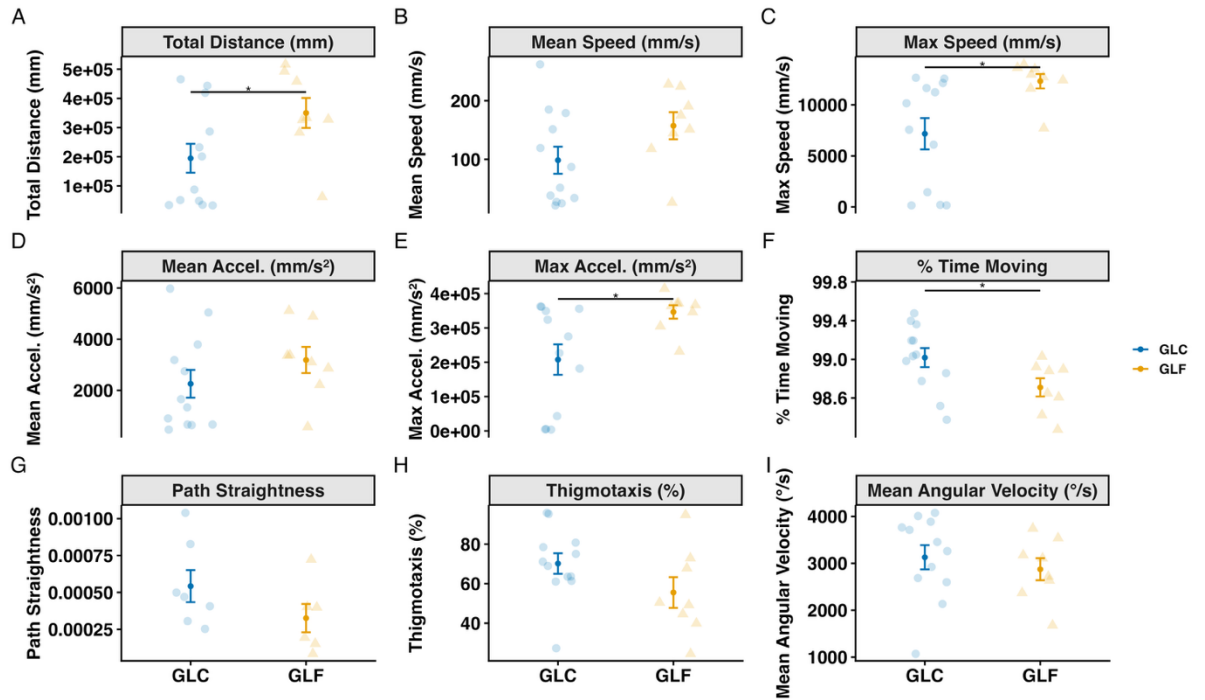


Figure 18: Behavioural phenotypes of GLC and GLF zebrafish.

Panels show quantitative assessment of locomotor and exploratory behaviours across nine metrics: **A)** total distance travelled, **B)** mean swimming speed, **C)** maximum swimming speed, **D)** mean acceleration, **E)** maximum acceleration, **F)** proportion of time moving, **G)** path straightness, **H)** thigmotaxis, and **I)** mean angular velocity. Individual points represent single animals; points denote group means (GLC and GLF), error bars represent standard errors; with each metric plotted in a separate panel with independent y-axis scaling. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

5.5. Discussion

This study provides molecular evidence that the germline directly affects somatic ageing. By developing a longitudinal epigenetic clock based on repeated fin clips from individual zebrafish, this study shows that germline ablation alters not only the level but also the temporal organisation and stability of somatic epigenetic ageing, with measurable consequences for late-life behaviour. These findings extend previous work based largely on lifespan or cross-sectional molecular readouts (Chen et al 2020; Chapter 2) and provide insight into how germline signalling influences ageing across the life course.

5.5.1. Germline ablation reshapes somatic epigenetic ageing trajectories

At the level of epigenetic age estimates, GLF fish exhibited a lower overall epigenetic age than GLC controls, consistent with a substantial body of work showing that germline removal or suppression can enhance somatic maintenance and stress resistance across diverse taxa (Hsin and Kenyon, 1999, Ermolaeva et al., 2013, Chen et al., 2020, Zhou et al., 2024). Comparable reductions in molecular ageing readouts have also been observed following interventions that alter reproductive or growth signalling, consistent with the involvement of conserved nutrient-sensing pathways in germline–soma communication (Templeman and Murphy, 2018, Saxton and Sabatini, 2017).

However, longitudinal evaluation revealed a more complex temporal structure. In both GLF and GLC fish, epigenetic age relative to chronological age declined with advancing age, indicating that epigenetic ageing trajectories are non-linear rather than reflecting uniform monotonic acceleration across the lifespan. Comparable non-linear behaviour has been reported in longitudinal human cohorts and experimental systems, where epigenetic clocks often show reduced slope, deceleration, or plateauing at

advanced ages despite continued chronological ageing (Marioni et al., 2019, Horvath, 2013, Snir et al., 2019). These observations suggest that epigenetic clocks are sensitive to age-dependent changes in regulatory dynamics and stability, rather than simply integrating cumulative molecular damage over time.

Importantly, the age-associated deceleration of epigenetic ageing was markedly weaker in GLF fish, such that their somatic epigenetic age exceeded that of GLC individuals at late chronological ages. Comparable stage-dependent effects have been reported in experimental studies of longevity and genome maintenance interventions, in which early-life molecular or physiological benefits are not sustained into advanced age. For example, mouse models with enhanced genome surveillance or tumour suppressor activity show improved early life stress resistance and cancer protection but develop premature ageing phenotypes and tissue dysfunction later in life (Tyner et al., 2002, Campisi, 2005). Similarly, constitutive activation of DNA damage response pathways improves early genome stability while promoting senescence-associated decline and epigenetic instability at later ages (Dabin et al., 2016, Baker et al., 2011). Perturbations of conserved growth and reproductive signalling pathways also support this pattern. Reduced IGF-1 signalling confers early life stress resistance and lifespan extension in some contexts, but does not uniformly preserve late life tissue function and can exacerbate age-associated decline under specific conditions (Holzenberger et al., 2003). In vertebrate systems, several interventions that bias resources or regulatory control towards early life genome protection can therefore impose long term costs on tissue maintenance and regulatory stability, highlighting a trade-off between early life robustness and late life homeostasis rather than uniform slowing of ageing across the lifespan. Indeed, persistent activation of DNA damage response pathways, including p53-mediated signalling, has been shown to

enhance early life genome integrity while promoting cellular senescence and epigenetic instability later in life (Lane, 1992, Vousden and Lane, 2007, Dabin et al., 2016).

This pattern is difficult to reconcile with a purely energetic interpretation of the disposable soma theory, which predicts sustained improvement in somatic maintenance following reduced reproductive investment (Kirkwood, 1977). Instead, the observed trajectories are more consistent with models in which the germline exerts active regulatory control over somatic repair and stress response pathways, shaping ageing trajectories in a stage-dependent manner (Maklakov and Chapman, 2019, Maklakov and Immler, 2016). In this framework, germline-carrying individuals exhibit elevated somatic epigenetic age during periods of active reproduction but undergo a late-life deceleration or reorganisation of epigenetic ageing as reproductive investment declines. Germline-free individuals, by contrast, show early-life molecular youth but lack this late-life reorganisation, resulting in weaker slowing of epigenetic ageing at advanced ages.

Together, these findings support the idea that germline ablation shifts somatic maintenance in a way that is beneficial early in life but compromises long-term epigenetic stability, rather than uniformly slowing ageing throughout life.

5.5.2. Developmental gene regulation as a substrate of epigenetic ageing

The CpG sites that were robustly selected across clock models were enriched near genes involved in tissue morphogenesis, sensory and visual development, and transcriptional regulation, indicating that age-associated methylation change in this system preferentially targets developmental regulatory networks rather than core metabolic or structural housekeeping genes. Such enrichment is a recurring feature of epigenetic clocks across species: in both humans and mice, age-informative CpGs are repeatedly

located near developmental transcription factors and regulators of cell fate, supporting the view that epigenetic ageing reflects altered regulation of developmental programmes rather than uniform accumulation of molecular damage (Horvath, 2013; Stubbs et al., 2017).

Within this broader developmental signature, genes associated with sensory and visual system development were particularly prominent. This is consistent with evidence that neuronal and sensory regulatory pathways are frequent targets of age-associated DNA methylation change, reflecting the high degree of epigenetic constraint imposed during development and the sustained sensitivity of these systems to regulatory disruption later in life (Horvath, 2013; Slieker et al., 2018).

Finally, the pronounced late-life hypomethylation at these robustly selected CpG sites in GLF fish suggests reduced stabilisation of developmental gene regulation in the absence of the germline. Loss of such stabilisation is expected to increase regulatory noise at loci that require precise methylation control, providing a plausible molecular basis for the late-life divergence in epigenetic ageing trajectories observed in GLF individuals.

5.5.3. Stochastic methylation drift and loss of homeostatic control

Although the overall effect of germline ablation on stochastic methylation drift did not reach statistical significance, the qualitative divergence in drift behaviour between groups is biologically informative. In GLC fish, CpG sites showed a tendency to revert towards intermediate methylation states over time. Similar convergence towards intermediate methylation levels has been reported in ageing mammalian tissues and has been interpreted as the outcome of active but imperfect epigenetic maintenance rather than purely random drift (Fraga et al., 2005; Issa, 2014). Mechanistically, such behaviour is consistent with the probabilistic nature of maintenance methylation during

DNA replication, whereby copying errors accumulate gradually but remain constrained by ongoing maintenance activity (Bestor, 2000, Jones and Liang, 2009). In addition, DNA repair-associated chromatin remodelling can locally reshape methylation states, contributing to bounded convergence rather than unstructured divergence (Dabin et al., 2016, Cuzzo et al., 2017). Together, these processes provide a plausible basis for the stabilising tendency observed in GLC fish, in which extreme methylation states are progressively constrained despite increasing cellular stress. In contrast, GLF fish exhibited divergence away from intermediate methylation states, a pattern consistent with weakened epigenetic homeostasis and increased stochastic methylation drift (Fraga et al., 2005, Gensous et al., 2017). Increased methylation drift of this kind has been associated with elevated transcriptional noise, erosion of cellular identity and reduced regulatory precision in ageing tissues (Gensous et al., 2017, Viñuela et al., 2017). Experimental and observational studies further indicate that such drift can decouple regulatory programmes from their developmental constraints, increasing vulnerability to functional decline and stress sensitivity in late life (Fraga et al., 2005, Gensous et al., 2017). While the present data do not demonstrate a statistically significant increase in drift, the directionality observed in GLF fish is consistent with these previously described patterns.

One plausible contributor to the altered methylation dynamics observed in GLF fish is sustained activation of DNA damage response signalling. Germline ablation has been shown to enhance genome surveillance and DNA damage responses in somatic tissues, particularly under genotoxic stress, indicating increased engagement of repair and checkpoint pathways (Ermolaeva et al., 2013, Chen et al., 2020). While such activation can improve early life genome protection, experimental evidence indicates that prolonged engagement of DNA damage response pathways can disrupt

chromatin organisation and compromise long-term epigenetic stability (Dabin et al., 2016, Panier et al., 2024).

Within this context, the p53 signalling network provides a coherent mechanistic link. p53 integrates DNA damage signalling with cell cycle arrest, apoptosis and senescence, and plays a central role in enforcing genome integrity (Lane, 1992, Vousden and Lane, 2007). Experimental models with elevated or constitutive p53 activity show enhanced tumour suppression and genome protection early in life, but develop premature tissue dysfunction, increased senescence and age-associated epigenetic alterations at later stages (Tyner et al., 2002, Campisi, 2005, Campisi and d'Adda di Fagagna, 2007). Consistent with this framework, we have previously reported increased expression of p53-associated genes in GLF somatic tissues (Chapter 2), supporting a link between germline ablation, heightened genome surveillance and the late-life epigenetic destabilisation observed here. Together, these findings support a model in which germline ablation alters the balance between genome protection and long-term epigenetic stability, resulting in early life molecular benefits that are accompanied by increased susceptibility to late-life regulatory drift.

5.5.4. Molecular ageing and behavioural outcomes

The association between altered epigenetic ageing trajectories and late-life locomotor behaviour provides an important functional context for the molecular findings. GLF fish exhibited increased burst-like, high-intensity movement, whereas GLC fish maintained more sustained, moderate activity. Comparable shifts in locomotor patterning have been reported in ageing vertebrates and are often interpreted as reflecting changes in neuromuscular control, sensory processing, or stress responsiveness rather than simple changes in overall activity levels (Kishi et al., 2009, Bronikowski et al., 2022).

The behavioural phenotype observed in GLF fish can therefore be interpreted in at least two non-exclusive ways. One possibility is that, despite exhibiting accelerated epigenetic ageing in late life, GLF fish retain enhanced short-term energetic capacity or neuromuscular responsiveness, resulting in elevated peak acceleration and swimming speed. Under this interpretation, DNA methylation-based age reflects cumulative regulatory or metabolic alteration rather than an immediate loss of functional capacity. Such decoupling between molecular ageing markers and short-term performance has been documented in vertebrates, where ageing or stress exposure can coexist with preserved or even exaggerated burst performance, particularly in tasks relying on anaerobic output or rapid motor recruitment (Bronikowski et al., 2022). This interpretation is consistent with the view that epigenetic clocks capture integrated regulatory state rather than moment-to-moment physiological competence (Horvath and Raj, 2018).

An alternative, non-exclusive interpretation is that the burst-dominated movement pattern reflects altered locomotor regulation, in which activity is expressed as short, intense bouts rather than sustained swimming. Age-associated changes in swimming behaviour and locomotor performance have been documented in zebrafish, including reduced endurance, altered activity patterns, and increased fragmentation of movement with age (Gerhard et al., 2002, Kishi et al., 2003). In addition, acute stressors are well established to elicit burst-like or erratic locomotion in zebrafish, reflecting heightened arousal rather than coordinated exploratory behaviour (Stewart et al., 2012, Egan et al., 2009). From this perspective, increased peak acceleration may arise from dysregulated motor control or elevated stress responsiveness rather than improved functional performance per se.

Together, these interpretations indicate that increased burst activity does not necessarily imply improved locomotor health. Instead, it may reflect a

shift in how locomotor output is deployed, with preserved or exaggerated short-term responses occurring alongside broader age-associated regulatory changes. Discriminating between these possibilities would require direct assessment of endurance, neuromuscular coordination, and stress responsiveness, which lies beyond the scope of the present study.

5.5.5. CpG organisation and the structure of ageing signals

A central methodological finding of this study is that Cluster-based CpG aggregation outperformed tiled aggregation in longitudinal epigenetic clock models. This contrasts with our previous observations in the determinant clock framework, expanded on further in Chapter 3, and underscores an important conceptual point: the genomic architecture of age-associated methylation change is context dependent. In longitudinal settings, where within-individual trajectories and late-life dynamics are of interest, biologically coherent CpG clusters are likely to better capture coordinated regulatory change than arbitrary genomic windows.

Clustered CpG behaviour is consistent with the organisation of DNA methylation within regulatory domains such as promoters, enhancers and chromatin interaction units, where neighbouring CpGs often exhibit correlated dynamics driven by shared transcriptional and chromatin states. In contrast, Tiled aggregation combines CpGs that may belong to distinct regulatory elements and be governed by different transcriptional or chromatin states. Averaging across such heterogeneous sites reduced sensitivity to coordinated methylation change, thereby weakening the signal relevant for modelling longitudinal epigenetic ageing. This observation reinforces the view that epigenetic ageing is not uniformly distributed across the genome but structured around regulatory units that are selectively sensitive to cumulative cellular stress and altered maintenance regimes.

Chapter 5 - Discussion

More broadly, it highlights the importance of benchmarking multiple clock architectures when testing mechanistic hypotheses, rather than assuming that an optimal design is transferable across experimental contexts.

5.5.6. Conclusion

These results provide direct evidence that the germline influences not only the rate but also the temporal structure and regulatory stability of epigenetic somatic ageing. By following individual trajectories over time, this study shows that germline ablation is associated with reduced epigenetic age early in life but altered late-life dynamics characterised by weakened slowing and increased regulatory instability. Similar early-life enhancements of somatic maintenance and stress resistance following germline removal or suppression have been demonstrated in invertebrate and vertebrate systems (Hsin and Kenyon, 1999, Ermolaeva et al., 2013, Chen et al., 2020). Our findings extend these observations by showing that such benefits are accompanied by altered late-life epigenetic organisation, consistent with a trade-off between early-life somatic robustness and long-term regulatory stability.

In summary, these findings support a model in which the germline acts as an active regulator of somatic ageing. In this framework, germline ablation rebalances genome surveillance, stress response and maintenance pathways in a manner that confers early life molecular benefits but compromises long-term epigenetic stability. Such stage-dependent effects reconcile observations of lifespan extension or early life molecular youth with late life dysregulation and align with evolutionary theories that emphasise trade-offs between robustness, reproduction and ageing, including the expensive germline hypothesis and developmental theory of ageing (Maklakov and Immler, 2016, Maklakov and Chapman, 2019).

5.6. Data Availability

All sequencing data, including RRBS data, generated for this project are available under accession (under embargo until acceptance for publication), <https://www.ebi.ac.uk/ena/browser/view/PRJEB107118>. All scripts and supplementary data are available at the git repository : <https://github.com/asdfjohnny1/germline-ablation-epigenetic-ageing.git>

5.7. Contributions

The studies were conceptualised and designed by S.I , J.C.D.C.

Data analysis and visualisation was carried out by J.C.D.C. ,

Samples were collected and processed by J.C.D.C

All code was conceived and generated by J.C.D.C.

Supervision was provided by S.I., A.M

Funding was provided by S.I., R.R

The first draft of the manuscript was written by J.C.D.C. and S.I.

5.8. Ethics declaration

Competing interests: All authors declare no conflict or competing interests.

All zebrafish work was completed in accordance with institutional, local and national laws and governing bodies.

5.9. Extended data

Appendix Chapter 4 for supplementary methods, figures, tables and legends.

All files and processed data files are available on the GitHub page : <https://github.com/asdfjohnny1/germline-ablation-epigenetic-ageing.git>

5.10. Abbreviations

GLF - Germline-free fish

GLC Germline-carrying fish

DNAm - DNA methylation

CpG - 5'-C-phosphate-G-3'

Δ -Age - Delta age

5.11. References

AKALIN, A., KORMAKSSON, M., LI, S., GARRETT-BAKELMAN, F. E., FIGUEROA, M. E., MELNICK, A. & MASON, C. E. 2012. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*, 13, R87.

ARANTES-OLIVEIRA, N., APFELD, J., DILLIN, A. & KENYON, C. 2002. Regulation of life-span by germline stem cells in *Caenorhabditis elegans*. *Science*, 295, 502-5.

BAKER, D. J., WIJSHAKE, T., TCHKONIA, T., LEBRASSEUR, N. K., CHILDS, B. G., VAN DE SLUIS, B., KIRKLAND, J. L. & VAN DEURSEN, J. M. 2011. Clearance of

Chapter 5 - References

p16Ink4a-positive senescent cells delays ageing-associated disorders. *Nature*, 479, 232-236.

BELL, C. G., LOWE, R., ADAMS, P. D., BACCARELLI, A. A., BECK, S., BELL, J. T., CHRISTENSEN, B. C., GLADYSHEV, V. N., HEIJMANS, B. T., HORVATH, S., IDEKER, T., ISSA, J.-P. J., KELSEY, K. T., MARIONI, R. E., REIK, W., RELTON, C. L., SCHALKWYK, L. C., TESCHENDORFF, A. E., WAGNER, W., ZHANG, K. & RAKYAN, V. K. 2019. DNA methylation aging clocks: challenges and recommendations. *Genome Biology*, 20.

BELSKY, D. W., CASPI, A., HOUTS, R., COHEN, H. J., CORCORAN, D. L., DANESE, A., HARRINGTON, H., ISRAEL, S., LEVINE, M. E., SCHAEFER, J. D., SUGDEN, K., WILLIAMS, B., YASHIN, A. I., POULTON, R. & MOFFITT, T. E. 2015. Quantification of biological aging in young adults. *Proc Natl Acad Sci U S A*, 112, E4104-10.

BENAYOUN, B. A., POLLINA, E. A. & BRUNET, A. 2015. Epigenetic regulation of ageing: linking environmental inputs to genomic stability. *Nature Reviews Molecular Cell Biology*, 16, 593-610.

BESTOR, T. H. 2000. The DNA methyltransferases of mammals. *Hum Mol Genet*, 9, 2395-402.

BRONIKOWSKI, A. M., MEISEL, R. P., BIGA, P. R., WALTERS, J. R., MANK, J. E., LARSCHAN, E., WILKINSON, G. S., VALENZUELA, N., CONARD, A. M., DE MAGALHÃES, J. P., DUAN, J. E., ELIAS, A. E., GAMBLE, T., GRAZE, R. M., GRIBBLE, K. E., KREILING, J. A. & RIDDLE, N. C. 2022. Sex-specific aging in animals: Perspective and future directions. *Aging Cell*, 21, e13542.

CAMPISI, J. 2005. Senescent Cells, Tumor Suppression, and Organismal Aging: Good Citizens, Bad Neighbors. *Cell*, 120, 513-522.

CAMPISI, J. & D'ADDA DI FAGAGNA, F. 2007. Cellular senescence: when bad things happen to good cells. *Nat Rev Mol Cell Biol*, 8, 729-40.

Chapter 5 - References

CHEN, H.-Y., JOLLY, C., BUBLYS, K., MARCU, D. & IMMLER, S. 2020. Trade-off between somatic and germline repair in a vertebrate supports the expensive germ line hypothesis. *Proceedings of the National Academy of Sciences*, 117, 8973-8979.

CUOZZO, C., PORCELLINI, A., ANGRISANO, T., MORANO, A., LEE, B., PARDO, A. D., MESSINA, S., IULIANO, R., FUSCO, A., SANTILLO, M. R., MULLER, M. T., CHIARIOTTI, L., GOTTESMAN, M. E. & AVVEDIMENTO, E. V. 2017. Correction: DNA Damage, Homology-Directed Repair, and DNA Methylation. *PLoS Genet.* United States.

DABIN, J., FORTUNY, A. & POLO, E., SOPHIE 2016. Epigenome Maintenance in Response to DNA Damage. *Molecular Cell*, 62, 712-727.

DAY, K., WAITE, L. L., THALACKER-MERCER, A., WEST, A., BAMMAN, M. M., BROOKS, J. D., MYERS, R. M. & ABSHER, D. 2013. Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome Biology*, 14, R102.

EGAN, R. J., BERGNER, C. L., HART, P. C., CACHAT, J. M., CANAVELLO, P. R., ELEGANTE, M. F., ELKHAYAT, S. I., BARTELS, B. K., TIEN, A. K., TIEN, D. H., MOHNOT, S., BEESON, E., GLASGOW, E., AMRI, H., ZUKOWSKA, Z. & KALUEFF, A. V. 2009. Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish. *Behav Brain Res*, 205, 38-44.

ERMOLAEVA, M. A., SEGREF, A., DAKHOVNIK, A., OU, H.-L., SCHNEIDER, J. I., UTERMÖHLEN, O., HOPPE, T. & SCHUMACHER, B. 2013. DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance. *Nature*, 501, 416-420.

FRAGA, M. F., BALLESTAR, E., PAZ, M. F., ROPERO, S., SETIEN, F., BALLESTAR, M. L., HEINE-SUÑER, D., CIGUDOSA, J. C., URIOSTE, M., BENITEZ, J., BOIX-CHORNET, M., SANCHEZ-AGUILERA, A., LING, C., CARLSSON, E., POULSEN, P.,

Chapter 5 - References

- VAAG, A., STEPHAN, Z., SPECTOR, T. D., WU, Y.-Z., PLASS, C. & ESTELLER, M. 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences*, 102, 10604-10609.
- GENSOUS, N., BACALINI, M. G., PIRAZZINI, C., MARASCO, E., GIULIANI, C., RAVAIOLI, F., MENGOZZI, G., BERTARELLI, C., PALMAS, M. G., FRANCESCHI, C. & GARAGNANI, P. 2017. The epigenetic landscape of age-related diseases: the geroscience perspective. *Biogerontology*, 18, 549-559.
- GERHARD, G. S., KAUFFMAN, E. J., WANG, X., STEWART, R., MOORE, J. L., KSALES, C. J., DEMIDENKO, E. & CHENG, K. C. 2002. Life spans and senescent phenotypes in two strains of Zebrafish (*Danio rerio*). *Exp Gerontol*, 37, 1055-68.
- HAURY A-C, GESTRAUD P, VERT J-P . 2011. The Influence of Feature Selection Methods on Accuracy, Stability and Interpretability of Molecular Signatures. *PLoS ONE* 6(12): e28210. <https://doi.org/10.1371/journal.pone.0028210>
- HOLZENBERGER, M., DUPONT, J., DUCOS, B., LENEUE, P., GÉLOËN, A., EVEN, P. C., CERVERA, P. & LE BOUC, Y. 2003. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature*, 421, 182-7.
- HORVATH, S. 2013. DNA methylation age of human tissues and cell types. *Genome Biology*, 14, R115.
- HORVATH, S. & RAJ, K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*, 19, 371-384.
- HSIN, H. & KENYON, C. 1999. Signals from the reproductive system regulate the lifespan of *C. elegans*. *Nature*, 399, 362-366.
- ISSA, J. P. 2014. Aging and epigenetic drift: a vicious cycle. *J Clin Invest*, 124, 24-9.

Chapter 5 - References

- JACKSON, S. P. & BARTEK, J. 2009. The DNA-damage response in human biology and disease. *Nature*, 461, 1071-1078.
- JONES, P. A. & LIANG, G. 2009. Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet*, 10, 805-11.
- KIRKWOOD, T. B. L. 1977. Evolution of ageing. *Nature*, 270, 301-304.
- KISHI, S., SLACK, B. E., UCHIYAMA, J. & ZHDANOVA, I. V. 2009. Zebrafish as a genetic model in biological and behavioral gerontology: where development meets aging in vertebrates--a mini-review. *Gerontology*, 55, 430-41.
- KISHI, S., UCHIYAMA, J., BAUGHMAN, A. M., GOTO, T., LIN, M. C. & TSAI, S. B. 2003. The zebrafish as a vertebrate model of functional aging and very gradual senescence. *Exp Gerontol*, 38, 777-86.
- KRUEGER, F. & ANDREWS, S. R. 2011. Bismark: a flexible aligner and methylation caller for Bisulphite-Seq applications. *Bioinformatics*, 27, 1571-2.
- LANE, D. P. 1992. p53, guardian of the genome. *Nature*, 358, 15-16.
- LEVINE, M. E., LU, A. T., QUACH, A., CHEN, B. H., ASSIMES, T. L., BANDINELLI, S., HOU, L., BACCARELLI, A. A., STEWART, J. D., LI, Y., WHITSEL, E. A., WILSON, J. G., REINER, A. P., AVIV, A., LOHMAN, K., LIU, Y., FERRUCCI, L. & HORVATH, S. 2018. An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, 10, 573-591.
- LODATO, M. A., RODIN, R. E., BOHRSON, C. L., COULTER, M. E., BARTON, A. R., KWON, M., SHERMAN, M. A., VITZTHUM, C. M., LUQUETTE, L. J., YANDAVA, C. N., YANG, P., CHITTENDEN, T. W., HATEM, N. E., RYU, S. C., WOODWORTH, M. B., PARK, P. J. & WALSH, C. A. 2018. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science*, 359, 555-559.

Chapter 5 - References

- LÓPEZ-OTÍN, C., BLASCO, M. A., PARTRIDGE, L., SERRANO, M. & KROEMER, G. 2013. The hallmarks of aging. *Cell*, 153, 1194-217.
- MAKLAKOV, A., ALEXEI & IMMLER, S. 2016. The Expensive Germline and the Evolution of Ageing. *Current Biology*, 26, R577-R586.
- MAKLAKOV, A. A. & CHAPMAN, T. 2019. Evolution of ageing as a tangle of trade-offs: energy versus function. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20191604.
- MARIONI, R. E., SUDERMAN, M., CHEN, B. H., HORVATH, S., BANDINELLI, S., MORRIS, T., BECK, S., FERRUCCI, L., PEDERSEN, N. L., RELTON, C. L., DEARY, I. J. & HÄGG, S. 2019. Tracking the Epigenetic Clock Across the Human Life Course: A Meta-analysis of Longitudinal Cohort Data. *J Gerontol A Biol Sci Med Sci*, 74, 57-61.
- MARIONI RE, HARRIS SE, ZHANG Q, MCRAEAF, HAGENAARS SP, HILL WD, DAVIS G, RITCHIE CW, GALE CR, STARR JM, GOATE AM, PORTEOUS DJ, YANG J, EVVANS KL, DEARY IJ, WRAY NR, VISSCHER PM. GWAS on family history of Alzheimer's disease. *Transl Psychiatry*. 2018 May 18;8(1):99. doi: 10.1038/s41398-018-0150-6. Erratum in: *Transl Psychiatry*. 2019 Jun 6;9(1):161. doi: 10.1038/s41398-019-0498-2. PMID: 29777097; PMCID: PMC5959890.
- MEINSHAUSEN, N., & BÜHLMANN, P. 2010. Stability selection. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 72(4), 417-473.
- MITCHELL, S. E., SIMPSON, M., COULET, L., GOUEDARD, S., HAMBLY, C., MORIMOTO, J., ALLISON, D. B. & SPEAKMAN, J. R. 2024. Reproduction has immediate effects on female mortality, but no discernible lasting physiological impacts: A test of the disposable soma theory. *Proceedings of the National Academy of Sciences*, 121.

Chapter 5 - References

- PANIER, S., WANG, S. & SCHUMACHER, B. 2024. Genome Instability and DNA Repair in Somatic and Reproductive Aging. *Annual Review of Pathology: Mechanisms of Disease*, 19, 261-290.
- R CORE TEAM 2022. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.
- ROMERO-FERRERO, F., BERGOMI, M. G., HINZ, R. C., HERAS, F. J. H. & DE POLAVIEJA, G. G. 2019. idtracker.ai: tracking all individuals in small or large collectives of unmarked animals. *Nat Methods*, 16, 179-182.
- SAXTON, R. A. & SABATINI, D. M. 2017. mTOR Signaling in Growth, Metabolism, and Disease. *Cell*, 168, 960-976.
- SIMPSON, D. J. & CHANDRA, T. 2021. Epigenetic age prediction. *Aging Cell*, 20.
- SLANCHEV K, STEBLER J, DE LA CUEVA-MÉNDEZ G, RAZ E. Development without germ cells: the role of the germ line in zebrafish sex differentiation. *Proc Natl Acad Sci U S A*. 2005 Mar 15;102(11):4074-9. doi: 10.1073/pnas.0407475102. Epub 2005 Feb 23. PMID: 15728735; PMCID: PMC549510.
- SLIEKER, R. C., RELTON, C. L., GAUNT, T. R., SLAGBOOM, P. E. & HEIJMANS, B. T. 2018. Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception. *Epigenetics & Chromatin*, 11.
- SNIR, S., FARRELL, C. & PELLEGRINI, M. 2019. Human epigenetic ageing is logarithmic with time across the entire lifespan. *Epigenetics*, 14, 912-926.
- STEWART, A. M., GAIKWAD, S., KYZAR, E. & KALUEFF, A. V. 2012. Understanding spatio-temporal strategies of adult zebrafish exploration in the open field test. *Brain Res*, 1451, 44-52.

Chapter 5 - References

STUBBS, T. M., BONDER, M. J., STARK, A.-K., KRUEGER, F., VON MEYENN, F., STEGLE, O. & REIK, W. 2017. Multi-tissue DNA methylation age predictor in mouse. *Genome Biology*, 18.

TEMPLEMAN, N. M. & MURPHY, C. T. 2018. Regulation of reproduction and longevity by nutrient-sensing pathways. *J Cell Biol*, 217, 93-106.

TYNER, S. D., VENKATACHALAM, S., CHOI, J., JONES, S., GHEBRANIOUS, N., IGELMANN, H., LU, X., SORON, G., COOPER, B., BRAYTON, C., PARK, S. H., THOMPSON, T., KARSENTY, G., BRADLEY, A. & DONEHOWER, L. A. 2002. p53 mutant mice that display early ageing-associated phenotypes. *Nature*, 415, 45-53.

VIJG, J. & SUH, Y. 2013. Genome instability and aging. *Annu Rev Physiol*, 75, 645-68.

VIÑUELA, A., BROWN, A. A., BUIL, A., TSAI, P.-C., DAVIES, M. N., BELL, J. T., DERMITZAKIS, E. T., SPECTOR, T. D. & SMALL, K. S. 2017. Age-dependent changes in mean and variance of gene expression across tissues in a twin cohort. *Human Molecular Genetics*, 27, 732-741.

VOUSDEN, K. H. & LANE, D. P. 2007. p53 in health and disease. *Nat Rev Mol Cell Biol*, 8, 275-83.

GILBERT WEIDINGER, JÜRGEN STEBLER, KRASIMIR SLANCHEV, KARIN DUMSTREI, CLARE WISE, ROBIN LOVELL-BADGE, CHRISTINE THISSE, BERNARD THISSE, EREZ RAZ, 2003,

dead end, a Novel Vertebrate Germ Plasm Component, Is Required for Zebrafish Primordial Germ Cell Migration and Survival, *Current Biology*, Volume 13, Issue 16, Pages 1429-1434, ISSN 0960-9822,

YAMAWAKI, T. M., ARANTES-OLIVEIRA, N., BERMAN, J. R., ZHANG, P. & KENYON, C. 2008. Distinct Activities of the Germline and Somatic

Chapter 5 - References

Reproductive Tissues in the Regulation of *Caenorhabditis elegans*' Longevity. *Genetics*, 178, 513-526.

ZHOU, L., JIANG, L., LI, L., MA, C., XIA, P., DING, W. & LIU, Y. 2024. A germline-to-soma signal triggers an age-related decline of mitochondrial stress response. *Nature Communications*, 15.

CHAPTER SIX

Rapamycin reconfigures ageing phenotypes in a genotype- and sex-dependent manner

de Coriolis, Jean-Charles ¹; Carlsson, Hanne ¹; Rahabari, Raheleh ²; Maklakov, Alexei ¹; Immler, Simone ¹

¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, NR4 6TX, Norwich, UK

² Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK



6. Chapter 6

6.1. Abstract

Rapamycin, a pharmacological inhibitor of mechanistic target of rapamycin (mTOR) signalling, robustly extends lifespan in multiple mammalian models and improves selected healthspan-associated traits in dogs and humans. However, how rapamycin influences molecular ageing trajectories and non-survival phenotypes across different exposure regimes and genetic contexts remains incompletely resolved. In this chapter, we examined the effects of rapamycin on epigenetic ageing, reproduction, behaviour, and ovarian lesion susceptibility using two complementary treatment paradigms in zebrafish (*Danio rerio*): a short-term late-life exposure initiated in aged adults and a longitudinal exposure spanning early to late life. Epigenetic ageing was quantified using DNA methylation-based age predictors derived from reduced representation bisulphite sequencing, while reproductive output, locomotor behaviour, and macroscopic ovarian lesions were assessed using standardised phenotypic and histopathological measures.

To assess whether genome maintenance capacity modulates responses to mTOR inhibition, experiments were conducted in both wild-type fish and a mismatch repair-deficient strain lacking functional *msh2*, which exhibits elevated genome instability and increased tumour susceptibility. Short-term rapamycin exposure in late life resulted in increased DNA methylation-based estimates of epigenetic age alongside reduced male reproductive output, indicating an acute perturbation of epigenetic and reproductive regulation rather than a shift towards a younger molecular state. By contrast, longitudinal exposure across adulthood altered behavioural ageing trajectories and increased the prevalence of macroscopic ovarian lesions, with effects most pronounced in mismatch repair-deficient individuals.

These regime-specific outcomes demonstrate that the effects of rapamycin on ageing-related phenotypes depend on the timing and duration of exposure as well as underlying genome maintenance capacity. Collectively, these findings emphasise that rapamycin functions as a context-dependent geromodulator rather than a universally beneficial anti-ageing intervention.

6.2. Introduction

Ageing is characterised by a progressive decline in physiological function accompanied by an increased risk of morbidity and mortality. Rather than constituting a discrete disease, ageing represents the primary risk factor for most chronic conditions, including cancer, cardiovascular disease, and neurodegeneration (Kennedy et al., 2014). The biological study of ageing, or biogerontology, proposed a set of conserved biological processes, commonly referred to as the hallmarks of ageing (López-Otín et al., 2013), and include genomic instability, loss of proteostasis, cellular senescence, and altered intercellular communication, among others (López-Otín et al., 2013). This framework provides a conceptual basis for identifying interventions improving ageing trajectories, preserving function, and/or reduce susceptibility to age-associated disease.

Among the most reproducible interventions known to influence ageing is dietary restriction, which extends lifespan and, in some contexts, healthspan across a wide range of taxa, from yeast to primates (Fontana et al., 2010; Partridge et al., 2018). Genetic and pharmacological manipulation of these pathways, including insulin/IGF signalling and the mTOR pathway, recapitulates many dietary restriction-associated effects in model organisms such as yeast (*Saccharomyces cerevisiae*), nematodes (*Caenorhabditis elegans*), fruit flies (*Drosophila melanogaster*), and mice (*Mus musculus*), positioning nutrient-sensing networks as central integrators of growth, metabolism, reproduction, and somatic maintenance rather than as a single

ageing mechanism (Ingram and Roth, 2015; Partridge et al., 2018). Within this framework, mTOR plays a key role in coupling nutrient and growth factor signals to anabolic processes. While essential for development and reproductive investment, sustained mTOR activity in adulthood has been associated with tissue dysfunction and increased susceptibility to age-related pathologies, including cancer and neurodegeneration (Blagosklonny, 2010; Weichhart, 2018; Papadopoli et al., 2019). These observations align mTOR signalling with evolutionary theories of ageing that emphasise trade-offs between early-life growth and reproduction and late-life somatic maintenance, including disposable soma and developmental ageing frameworks (Kirkwood, 1977; Maklakov and Chapman, 2019).

At the molecular level, mTOR functions as a serine/threonine kinase within two distinct multiprotein complexes. mTOR complex 1 (mTORC1) promotes anabolic processes, including protein synthesis, lipid biosynthesis, and cell growth, while suppressing autophagy, whereas mTOR complex 2 (mTORC2) regulates cytoskeletal organisation, cell survival, and insulin signalling (Saxton and Sabatini, 2017). Chronic activation of mTOR, particularly mTORC1, has been linked to several hallmarks of ageing, including impaired proteostasis, dysregulated nutrient sensing, and reduced autophagic capacity (Johnson et al., 2013). Consequently, pharmacological inhibition of mTOR signalling has emerged as a prominent strategy for influencing ageing-related processes.

Several pharmacological strategies have been used to attenuate mTOR signalling, differing in their specificity and mechanism of action. Some clinically used agents, including metformin and acarbose, influence mTOR activity indirectly by altering upstream metabolic and hormonal inputs such as glucose handling, insulin signalling, and cellular energy balance; as a result, their effects on ageing-related phenotypes are pleiotropic and cannot be attributed specifically to mTOR modulation (Harrison et al., 2014; Barzilai

et al., 2016). By contrast, rapamycin and related rapalogues inhibit mTOR signalling more directly through FKBP12-dependent suppression of mTORC1, with comparatively limited effects on mTORC2 (Saxton and Sabatini, 2017). Rapamycin robustly extends lifespan across multiple taxa, including yeast, nematodes, *Drosophila*, and mice, even when treatment is initiated late in life (Harrison et al., 2009; Bjedov et al., 2010; Miller et al., 2011). In addition to longevity, rapamycin modulates selected healthspan-associated traits, including immune function, cardiac performance, and metabolic homeostasis; however, these effects are strongly dependent on dose, treatment duration, sex, and genetic background (Kaeberlein, 2014). Evidence from domestic dogs and older human cohorts similarly indicates that rapamycin responses are context-dependent rather than uniformly beneficial (Mannick et al., 2014; Urfer et al., 2017). Despite its efficacy, the translation of sustained mTOR inhibition to long-term clinical use has been limited by adverse effects, including metabolic dysregulation and immunosuppression, which vary across tissues and life stages (Papadopoli et al., 2019; Roark and Iffland, 2025).

While lifespan and healthspan provide integrative measures of ageing, molecular biomarkers are required to assess whether these interventions directly modulate underlying biological ageing processes. Among these, epigenetic alterations, particularly changes in DNA methylation, have emerged as robust and evolutionarily conserved markers of biological ageing (Bell et al., 2019, Horvath and Raj, 2018). DNA methylation-based epigenetic clocks provide quantitative estimates of biological age that correlate with functional decline, disease risk, and mortality, and are increasingly used to evaluate the molecular impact of ageing interventions. In mice, rapamycin treatment slows epigenetic ageing in liver and blood (Petkovich et al., 2017, Horvath et al., 2019), and partial rejuvenation in specific tissues following mTOR inhibition (Shindyapina et al., 2022). However, these

effects are not universal: limited or absent effects on epigenetic age have been reported in brain and skeletal muscle, and responses vary by sex, genetic background, and the age at which treatment is initiated (Petkovich et al., 2017; Horvath et al., 2019; Miller et al., 2014). Together, these findings highlight the need for systematic assessments that integrate molecular ageing markers with functional and pathological outcomes.

Here, we use zebrafish as a vertebrate model system to investigate the effects of rapamycin on ageing phenotypes, including epigenetic ageing by reduced representation bisulphite sequencing, behaviour, reproductive output, and ovarian lesion prevalence. We compare two experimental procedures: a short-term late-life exposure, in which rapamycin was administered for four months beginning in aged adults after the completion of somatic growth and reproductive maturation, and a lifelong exposure, in which treatment was initiated in early adulthood and maintained continuously until natural death. This design allows separation of acute, late-life responses to mTOR inhibition from cumulative effects arising from sustained pathway suppression across the adult lifespan. Importantly our experimental design incorporates two genetic backgrounds: a wild-type strain (ABWT) and a strain lacking the mismatch repair (MMR) gene *msh2*. *Msh2* is a core component of the *MutS* complex and plays an essential role in correcting DNA replication errors and maintaining genomic stability; its loss leads to microsatellite instability, elevated mutation rates, and a strong predisposition to cancer (Baretti and Le, 2018, Lynch et al., 2009). The *msh2*-deficient zebrafish strain therefore provides a clinically relevant model of compromised genome maintenance and elevated susceptibility to genomic instability-associated pathology. By combining pharmacological inhibition of mTOR with genetic disruption of DNA repair, this study examines how rapamycin affects genome stability, tumour risk, and epigenetic ageing, allowing us to test whether mTOR inhibition mitigates, exacerbates, or

reshapes ageing trajectories under conditions of heightened genomic instability.

6.3. Material and methods

We performed two separate experiments to test the effects of rapamycin: Experiment 1 – rapamycin was fed to fish late in life for a short period of 4 months to test possible rejuvenation. Experiment 2 – fish were fed rapamycin throughout life to test for long-term effects of rapamycin.

6.3.1. Shared procedures

The following procedures were common to both the Aged Exposure and Lifelong Exposure experiments unless otherwise stated.

6.3.1.1. Rapamycin feed preparation

Rapamycin (LC Laboratories, USA) was incorporated into Sparos Adult Zebrafish Diet dry feed. Two treatment concentrations were prepared:

- (i) Low dose: 24.5 mg rapamycin kg⁻¹ body weight
- (ii) High dose: 245 mg rapamycin kg⁻¹ body weight

Feeds were prepared every six weeks and stored at -20 °C in airtight containers to preserve stability. Aliquots were thawed weekly and protected from light. Body weight dosing was calculated by regularly weighing the fish to maintain correct dosage of Rapamycin.

Chapter 6 - Material and methods

6.3.1.2. Behavioural tracking and analysis

For each assay, all focal individuals from a given tank were transferred together into a circular arena (235mm diameter) containing 2 L of system water. Behaviour was recorded for 21 minutes using an overhead GoPro HERO4 camera. The first minute of recording was excluded to allow for acclimation.

Video data were processed using *id.Tracker.ai* v 5.1.3 (Romero-Ferrero et al., 2019) to extract per-frame x-y coordinates for each individual. From these trajectories, eight standardised locomotor and spatial metrics were derived following Kermen et al. (2020), including total distance (mm), mean speed (mm s^{-1}), maximum speed (mm s^{-1}), mean acceleration, max acceleration, path straightness, thigmotaxis, and mean angular velocity. Full details on how these were calculated can be found in Chapter 5 [Material and methods](#).

6.3.1.3. Statistical analyses

All statistical analyses were performed in R v4.3.1 (R Core Team, 2022). Continuous variables, including methylation levels, epigenetic age estimates and behavioural metrics, were summarised as means $\pm$ standard error unless stated otherwise.

Treatment effects were evaluated using linear models or linear mixed-effects models from the lme4 package (Bates et al., 2015), with fish identity and/or experimental tank included as random effects where repeated measures or hierarchical structure was present.

Proportional data (e.g. probability of spawning, proportion of time spent in the upper half of the tank, ovarian lesion incidence) were analysed using generalised linear models or generalised linear mixed models with binomial error distribution and logit link function.

Chapter 6 - Material and methods

Age-dependent trajectories of behavioural or epigenetic variables were modelled using generalised additive models (mgcv package) to capture non-linear relationships with chronological age (Wood, 2017).

Significance of fixed effects (treatment, sex, genotype and their interactions) was evaluated using Type III ANOVA or likelihood ratio tests. Pairwise contrasts were performed using estimated marginal means, with p-values corrected for multiple testing using false discovery rate (FDR) correction (Lenth, 2021).

All statistical tests were two-sided, and significance was defined as $\text{FDR-}p_{adj} < 0.05$.

6.3.2. Experiment 1: Ageing Experiment: Short-term Rapamycin exposure in late life

6.3.2.1. Experimental design and procedures

Wild-type AB zebrafish were obtained from the European Zebrafish Resource Centre (EZRC; Karlsruhe Institute of Technology, Germany) and maintained under standardised laboratory conditions in 3.5 L tanks at 28 °C on a 14 h:10 h light:dark cycle, with five females and five males per tank. Fish were selected at 18 months of age and randomly assigned to control, low-dose, or high-dose rapamycin treatment groups, as described in Section 6.3.1 (Shared Procedures).

Rapamycin was administered via supplemented feed, with fish receiving daily doses corresponding to the low- and high-dose concentrations defined in Section 6.3.1 (Shared Procedures). Treatment was maintained for four months to assess short-term late-life responses to mTOR inhibition.

Chapter 6 - Material and methods

Reproductive performance was assessed after two and four months of treatment (approximately 20 and 22 months of age). Each experimental fish was paired with a non-experimental, untreated wild-type partner of the opposite sex, approximately one year younger, to minimise age-related fertility constraints in the mating partner. Each focal fish underwent two mating attempts separated by 24 h using standard zebrafish breeding protocols. Tanks were monitored at 30-minute intervals for spawning activity, and eggs were collected and counted manually. Reproductive success was quantified as both the probability of spawning and the total number of eggs produced per spawning event (Table S26).

Group swimming behaviour was assessed at baseline, and after two and four months of treatment, using the behavioural assay procedures described in Section 6.3.1 (Shared Procedures).

6.3.2.2. Epigenetic clock by RRBS

Fin clips were collected from female and male fish for DNA methylation analysis. Caudal fin clips were taken after two months of rapamycin treatment, and dorsal fin clips were taken after four months of treatment.

Eight females and eight males were sample per treatment group ($n = 16$ per group, 96 samples total; Table S27). Fin tissue was immediately snap-frozen in liquid nitrogen and stored at -80°C until DNA extraction.

Genomic DNA was extracted from individual fin clips using a modified Qiagen DNeasy Blood and Tissue Kit protocol (Qiagen, 69506). Briefly, tissue was lysed in ATL buffer containing proteinase K and incubated at 56°C for 1–2 hours with intermittent vortexing until complete digestion.

Lysates were processed through silica membrane spin columns according to the manufacturer's protocol, washed, and DNA was eluted in 100 μL of EB

Chapter 6 - Material and methods

buffer. DNA concentration and purity were assessed using a NanoDrop spectrophotometer and a Qubit fluorometer prior to downstream processing.

A total of 96 samples were processed for Reduced Representation Bisulphite Sequencing (RRBS) at the University of East Anglia using the TECAN Ovation® RRBS Methyl-Seq System. Libraries were prepared following a modified version of the manufacturer's protocol (<https://www.protocols.io/view/ovation-rrbs-methyl-seq-library-preparation-protocol-kxygx3kydg8j/v1>).

Briefly, genomic DNA was digested using MspI, followed by adapter ligation, end repair and size selection. Bisulphite conversion was then performed, followed by PCR amplification and final purification.

Library quality and fragment size distribution were assessed using an Agilent TapeStation system. Libraries were quantified prior to pooling for sequencing.

Sequencing was performed at the Wellcome Sanger Institute (Cambridge, UK) on an Illumina NovaSeq X platform using a 10B flow cell across three lanes, generating 150 bp paired-end reads.

6.3.2.3. Data trimming, alignment and methylation extraction for epigenetic clock analysis

Raw reads were processed using the NuGEN technologies pipeline (<https://www.nugen.com/products/nugen-nuget-rrbs>). Adapter and low-quality bases were trimmed with Trim Galore v 0.6.5 (https://Delta-Age.bioinformatics.babraham.ac.uk/projects/trim_galore/), and diversity adaptor artefacts were removed using a NuGEN-provided Python script (<https://github.com/nugentechnologies/NuMetRRBS/trimRRBSdiversityAda>

[ptCustomers.py](#)). Cleaned reads were aligned to the zebrafish reference genome (GRCz11; https://Δ-Age.ncbi.nlm.nih.gov/datasets/genome/GCF_000002035.6/) using Bismark v 0.25.1 (Krueger and Andrews, 2011), and duplicate reads were identified and marked with the NuDup tool in NuGEN (<https://github.com/tecangenomics/nudup>).

Methylation calling was performed using the Bismark methylation extractor (Krueger and Andrews, 2011), generating cytosine methylation reports in “.cov” format. Methylation metrics (methylated and unmethylated counts per CpG) were used as input for downstream analyses.

6.3.2.4. Methylation data preprocessing and epigenetic clock prediction

RRBS methylation data were imported as Bismark “.cov” files into R and processed using the methylKit package. CpGs with coverage below 5× were excluded, and cytosines overlapping known single-nucleotide polymorphisms were removed. For each retained CpG, methylation levels were calculated as the ratio of methylated reads to total coverage.

For the rapamycin RRBS dataset, epigenetic age prediction was performed using the final fin clip clock models identified in Chapter 3. In that analysis, a total of 816 final fin clip models had been selected. For each of these models, a CpG-level methylation matrix was reconstructed from the rapamycin RRBS data, retaining only the CpG sites included in the corresponding final model. CpG identifiers were intersected between the rapamycin dataset and each model, and features were ordered to exactly match the coefficient structure of the original model. Where CpG sites were missing from individual samples, values were imputed using the median methylation level across samples to preserve matrix completeness. Each reconstructed matrix was

Chapter 6 - Material and methods

then used to predict epigenetic age for the rapamycin-treated samples using the corresponding final model. No additional normalisation or transformation was applied prior to model-based prediction.

Epigenetic age estimates were generated using the model coefficients. Delta age (Δ -Age = Predicted – Chronological age) was calculated for each sample as the difference between predicted epigenetic age and chronological age, providing a measure of accelerated or decelerated ageing relative to chronological time.

6.3.3. Experiment 2: Lifelong Exposure experiment (chronic rapamycin treatment in ABWT and msh2 background)

6.3.3.1. Experimental design and procedures

ABWT fish were obtained from the Zebrafish International Resource Centre (ZIRC, University of Oregon, USA). The msh2 mutant line (allele Sa22101, Tübingen background) was obtained from EZRC. Homozygous mutants were outcrossed to ABWT to generate a heterozygous ABWT–Sa22101 line, which was maintained by continued outcrossing to ABWT. Heterozygous fish were used throughout the experiment.

Experimental tanks contained 13 females and 3 age-matched males of the same genotype and treatment. Male mortality was mitigated by replacement with fish of identical age, genotype, and treatment. Fish were randomly assigned to control, low-dose, or high-dose rapamycin treatment groups at three months of age.

Fish received feed supplemented with 10mg per kg body weight, per tank of rapamycin daily, corresponding to the low and high-dose rapamycin concentrations described in Section 6.3.1 (Shared Procedures).

Chapter 6 - Material and methods

Group swimming behaviour was measured longitudinally at six time points between 5 and 19 months of age (5,8,11,13,16,19). Anxiety-like behaviour was assessed using the Novel Tank Diving Test at 9 and 20 months. Behavioural tracking and metric extraction followed the protocol in Section 6.3.1 (Shared Procedures).

6.3.3.2. Novel Tank Diving Test

Anxiety-like behaviour and vertical exploration were assessed using the Novel Tank Diving Test (NTDT) at 9 and 20 months of age. At each time point, a randomly selected subset of female fish from each tank was tested individually.

Each fish was transferred into a rectangular novel tank (dimensions: 420 × 84 × 130 mm) filled with system water and recorded for 6 minutes using a side-mounted camera. The first minute was excluded from analysis to account for handling stress and initial acclimation.

Videos were analysed using IDtrackerAI to extract x-y coordinates. Vertical position was quantified by dividing the tank into equal top and bottom halves, with boundaries defined at 396 pixels for 9-month assays and 260 pixels for 20-month assays, based on camera geometry and tank dimensions. The proportion of time spent in the upper half of the tank was calculated for each individual as a measure of exploratory and anxiety-related behaviour.

In addition, freezing behaviour was quantified as described above, based on frame-to-frame displacement in the y-coordinate, with movements of less than 2 pixels scored as freezing events. All behavioural metrics were calculated across the full assay duration and used for statistical comparisons.

6.3.3.3. Tumour-like ovarian mass (gross pathology)

At the end of the lifelong exposure experiment, fish were euthanised using Aquacalm (Syndel) in accordance with institutional and UK Home Office guidelines. Animals were dissected immediately, and ovaries were examined under a stereomicroscope.

Ovarian pathology was assessed based on gross morphology. Macroscopic ovarian lesions were operationally defined as solid, abnormal tissue masses clearly distinct from normal ovarian architecture. Small age-associated foci commonly observed in older females were not classified as lesions and were excluded from scoring.

Lesion burden was recorded as a binary outcome (presence or absence) and, where applicable, further categorised by number (e.g. multiple lesions present). No histological or cellular analyses were performed; therefore, classifications reflect gross morphology rather than confirmed neoplastic status.

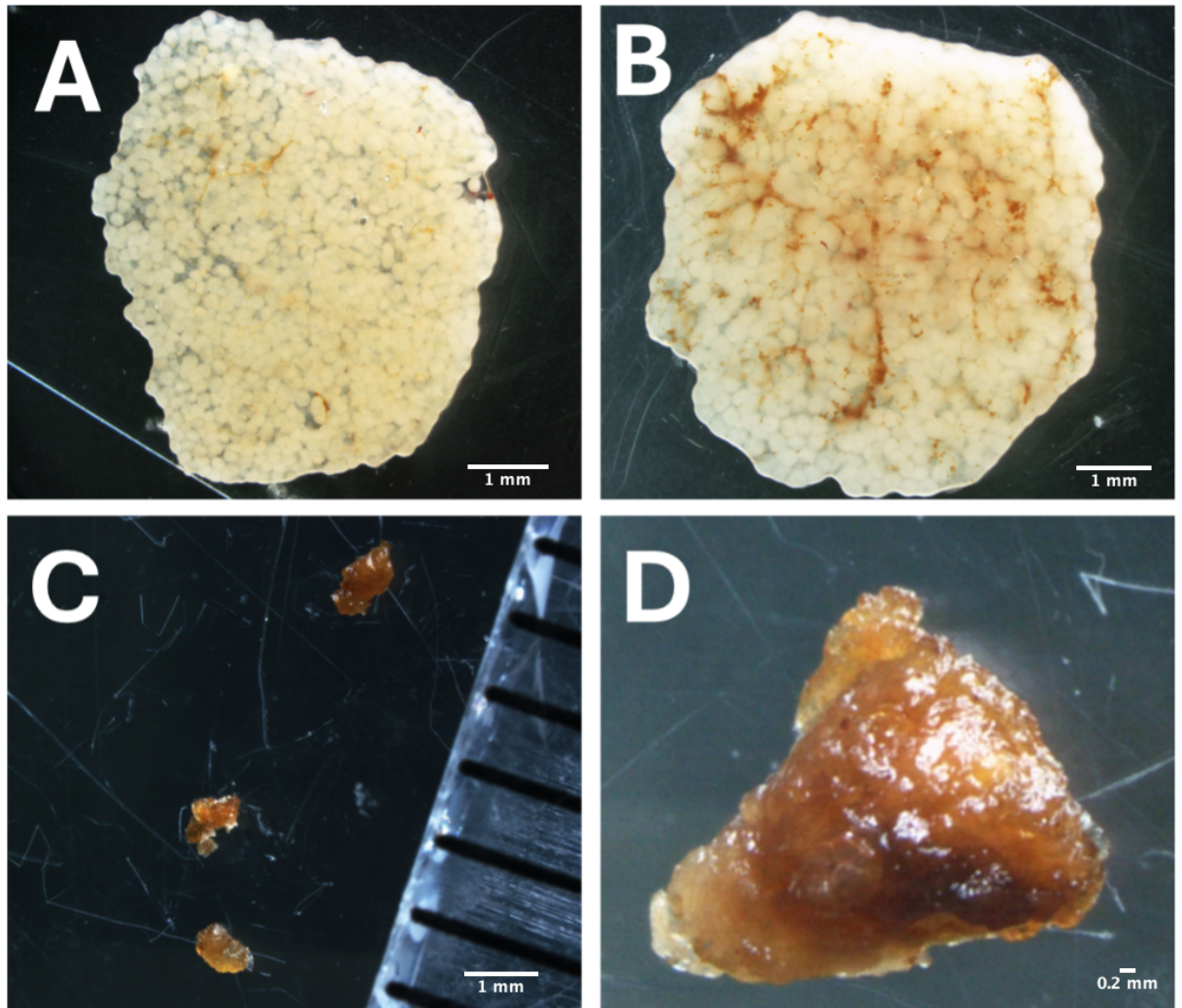


Figure 19 : Representative ovarian morphologies used to gross lesion classification

A) Normal ovarian ovaries., **B)** Ovaries with small, dark age-associated foci, not classified as lesions for the purposes of this study. **C)** Multiple small macroscopic lesions. **D)** Ovary with a single large macroscopic lesion.

Chapter 6 - Results

6.3.4. Rapamycin blood concentration measurements

Rapamycin blood concentrations were measured to confirm systemic exposure following dietary administration. Measurements were performed in female ABWT fish at 10 months of age, 3 hours after administration of a single low- or high-dose rapamycin-containing feed (Figure S13). Fish were euthanised using Aquacalm (Syndel) in accordance with institutional and UK Home Office regulations.

Whole blood was collected via incision posterior to the dorsal gill plate using a capillary tube. Due to the limited blood volume obtainable from individual fish, samples from 3–5 individuals within the same treatment group were pooled to obtain approximately 10 μL of blood per sample. Pooled blood was transferred into an Eppendorf tube containing 2 μL of 1% K_2EDTA in H_2O as anticoagulant.

Cells were lysed and proteins precipitated by addition of 10 μL precipitation solution (acetonitrile: ZnSO_4 , 4:1 v/v), followed by vortexing and centrifugation at $2,680 \times g$ for 6 minutes. The resulting supernatant was filtered through a 2 μm filter and analysed by liquid chromatography–mass spectrometry (LC–MS).

Calibration standards were generated by spiking blood from untreated fish with known concentrations of rapamycin and processed identically to experimental samples. Measured concentrations therefore reflect group-level circulating rapamycin levels rather than individual variability.

6.4. Results

Data presented in this chapter derive from two independent rapamycin exposure experiments with partially overlapping but non-identical sampling structures. In the short-term late-life exposure experiment, reproductive

assays were conducted on 114 fish, behavioural analyses were performed across 12 tanks containing 10 fish per tank, and DNA methylation analyses for epigenetic clock evaluation were based on 95 RRBS samples passing quality control. In the lifelong rapamycin exposure experiment, behavioural ageing trajectories were assessed using 54 tanks containing 16 fish per tank, ovarian lesion prevalence was evaluated following dissection of 234 fish, and novel tank diving tests were conducted using 5 fish per tank across 18 tanks. Sample sizes therefore vary by assay and experimental regime and are reported explicitly for each analysis. To verify effective drug delivery, circulating rapamycin levels were measured in blood samples from treated fish. Rapamycin concentrations were higher in the high-dose treatment group than in the low-dose group, confirming systemic uptake following administration (Figure S13). The weight of fish was also measured at the end of treatment and there was no difference in weight between treatments (Figure S13).

6.4.1. Experiment 1: Male reproductive success is reduced by high-dose rapamycin

Given the sensitivity of reproductive function to mTOR signalling, breeding success was first assessed after two months of rapamycin exposure to capture early, sex-specific effects in late life. A Treatment $\times$ Sex interaction was suggestive but did not reach the conventional significance threshold (binomial GLMM: $\chi^2 = 5.3$, $df = 2$, $p = 0.06$; Table S26). In males, breeding success was reduced under high-dose rapamycin relative to controls (*Estimate* = -2.4, $z = -1.9$, $p = 0.04$), while the low-dose group showed a similar but non-significant trend (*Estimate* = -1.4, $z = -1.7$, $p = 0.08$). No statistically significant treatment effects were detected in females (Figure 20).

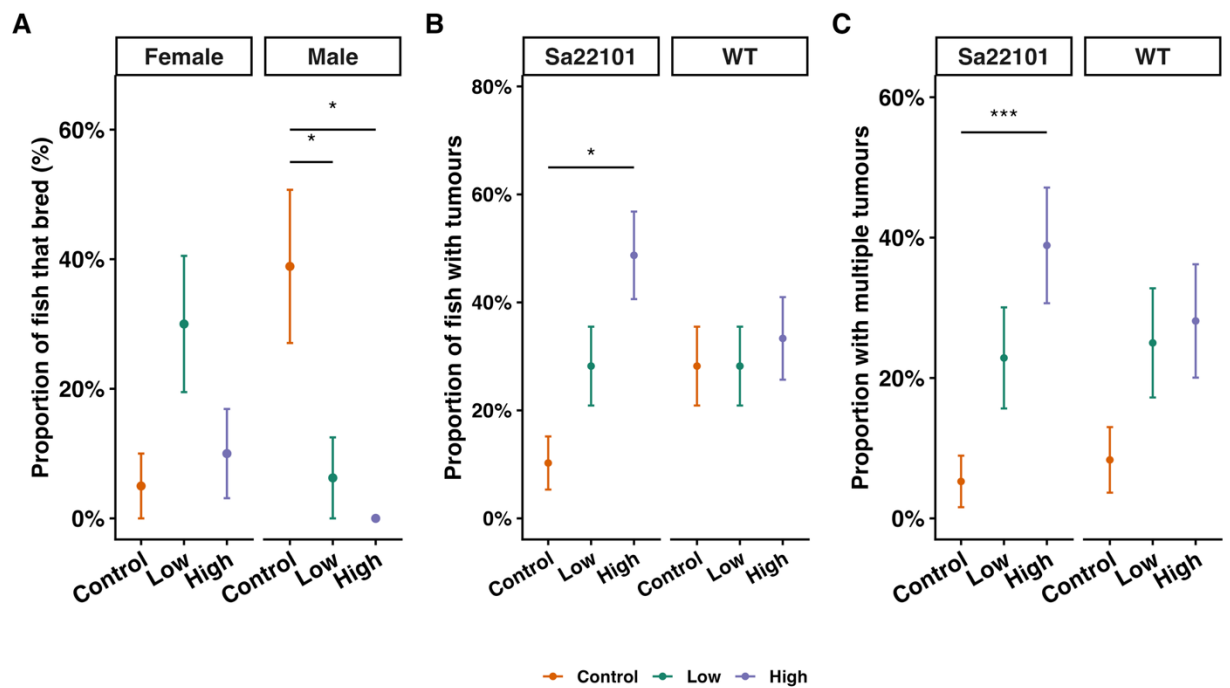


Figure 20: Rapamycin exposure alters circulating concentrations, reproductive performance and ovarian lesion incidences.

(A) Proportion of fish that bred in Experiment 1, short-term late-life experiment. Males exposed to rapamycin bred less frequently than controls. **(B)** Proportion of female fish with at least one ovarian lesion in the lifelong experiment. MMR strain individual exposed to rapamycin were more likely to develop ovarian lesions. **(C)** Proportion of female fish with multiple ovarian lesions in Experiment 2 : the lifelong experiment. MMR strain individuals exposed to rapamycin showed increased incidence of multiple ovarian lesions. Points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

6.4.2. Experiment 1: Low and High-dose rapamycin produce distinct locomotor behaviours in late-life

Rapamycin had significant effects on locomotor behaviour in a dose-dependent manner (Figure 21). Relative to controls, low-dose fish exhibited reduced total movement and decreased thigmotaxis (Table 1). In contrast, high-dose fish displayed increased peak performance metrics, including maximum speed and acceleration, spent a greater proportion of time moving, and showed increased thigmotaxis alongside reduced mean angular velocity (Tables 1, S28). Collectively, these results indicate that low-dose rapamycin attenuated exploratory activity, whereas high-dose treatment was associated with elevated locomotor activity and increased wall-associated swimming behaviour.

Chapter 6 - Results

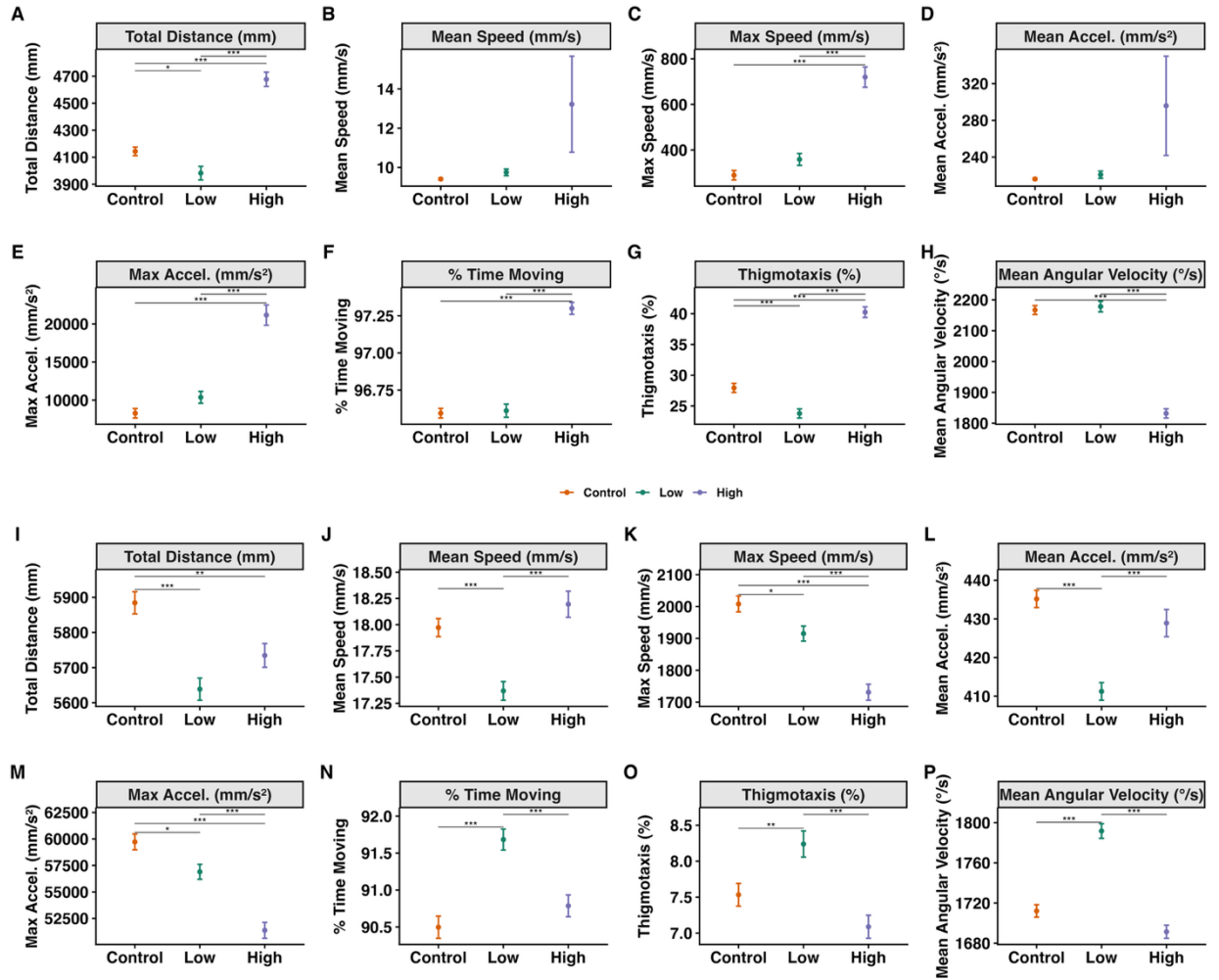


Figure 21: Behaviour metrics in Aged and Lifelong experiments

Panels show behavioural measurements for Aged (**A-H**) and Lifelong (**I-P**) experiments. Each pair of panels represents the same behavioural metric across the two experiments. The metrics are: (**A,I**) Total distance travelled (mm), (**B,J**) Mean speed (mm/s), (**C,K**) Maximum speed (mm/s), (**D,L**) Mean acceleration (mm/s²), (**E,M**) Max acceleration (mm/s²), (**F,N**) Percentage of time moving (%), (**G, O**) Thigmotaxis (%), (**H,P**) Mean angular velocity. Points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

Table 1 : Linear Mixed model showing dose-dependent effects of rapamycin on locomotor behaviour

Estimates (E) are calculated as the value of the first group minus the second group. Positive estimates indicate higher values in the first group listed.

Behaviour metric	Contrast	Estimate (E)	z-value	p-value
Total distance (mm)	Control - Low	193	3.5	0.001
	Low - High	-503	-9	<0.0001
Thigmotaxis	Control - High	-12	-11.5	<0.0001
	Control - Low	3.9	3.6	<0.0001
	Low - High	-16	-15.2	<0.0001
Maximum Speed	Control - High	-423	-9.4	0.0001
	Low - High	-281	-5.5	<0.0001
Maximum Acceleration	Control - High	-12.7	-9.4	<0.0001
	Low - High	-0.7	-7.6	<0.0001
Time moving (%)	Control - High	-0.7	-13	<0.0001
	Low - High	-0.7	-11.5	<0.0001
Mean Angular velocity	Control - High	327	16	<0.0001
	Low - High	341	16	< 0.0001

Chapter 6 - Results

6.4.3. Experiment 1: Rapamycin accelerates epigenetic ageing with dose and sex-specific effects

All 95 samples passing initial sequencing and alignment quality control were retained to predict epigenetic age. In total, 816 distinct predictions were made, spanning multiple CpG aggregation strategies, regression frameworks, and filtering parameterisations, and used for downstream analyses. Unless stated otherwise, all results below refer to this filtered model set, with comparisons interpreted across configurations rather than assuming a shared underlying feature space.

Exposure to rapamycin increased epigenetic age relative to controls (LMM: $\chi^2 = 553.3$, $df = 2$, $p < 0.0001$; Table S29), and fish exposed to both Low and High doses differed from controls (Control–Low: *Estimate* = -0.34, $z = -25$, $p < 0.0001$; Control–High: *Estimate* = -0.10, $z = -7$, $p < 0.0001$). The Low dose resulted in a more pronounced increase than the High dose (Low–High: *Estimate* = 0.23, $z = 17$, $p < 0.001$; Figure 22).

Non-Linear age-dependent patterns were observed across treatment groups in generalised additive models (Figure 22; Table S29), indicating that rapamycin affected the relationship between chronological age and the methylation-based age estimate over the course of the experiment.

Treatment effects also differed by sex (LMM: $\chi^2 = 1093.9$, $df = 2$, $p < 0.0001$; Table S29). In females, predicted epigenetic age was highest under the High dose treatment (Control–High: *Estimate* = -0.17, $z = -8.9$, $p < 0.0001$; Low–High: *Estimate* = -0.14, $z = -7$, $p < 0.0001$); whereas in males the largest increase was observed under the Low dose (Control–Low: *Estimate* = -0.64, $z = -33$, $p < 0.0001$; Low–High: *Estimate* = -0.60, $z = -32$, $p < 0.0001$).

These results indicate that epigenetic age acceleration was greatest in low-dose males and high-dose females and was accompanied by dose-

Chapter 6 - Results

dependent departures from typical hypomethylation trajectories, demonstrating that rapamycin modulates both the rate and trajectory of age-associated methylation change in a treatment- and sex-specific manner.

Chapter 6 - Results

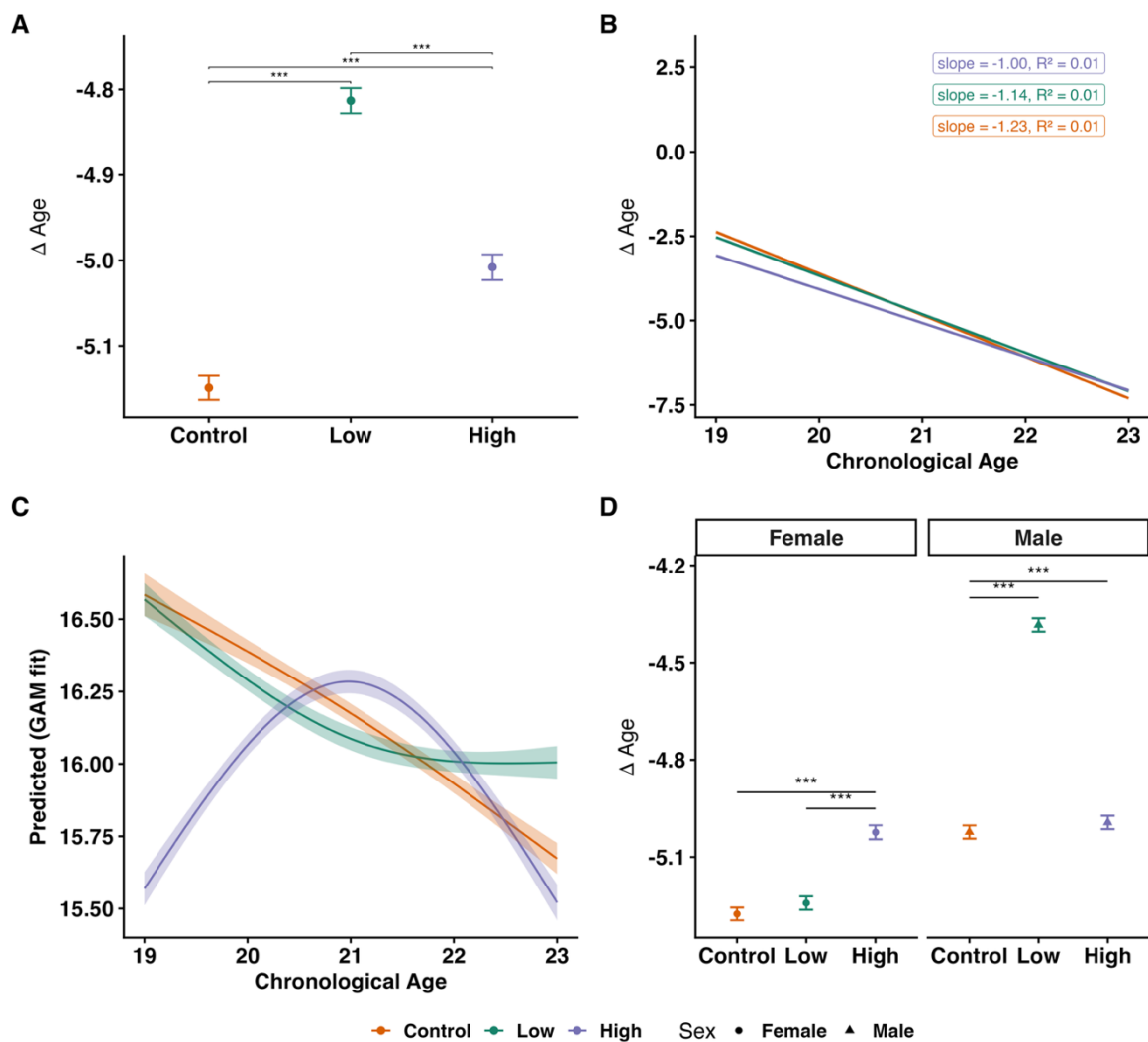


Figure 22 Rapamycin drastically effects epigenetic ageing in the Aged experiment

A) Mean epigenetic age deviation (Δ -Age), defined as the difference between predicted epigenetic age and chronological age. Low-dose fish showed significantly higher Δ -Age age than controls, and High-dose dish were also significantly higher than controls. **B)** Δ -Age trends across chronological time. Both High and Low-dose treatments showed a reduced slope compared to controls. **C)** Generalised Additive Model (GAM) of epigenetic age trajectories across chronological age. Low-dose fish showed plateauing Δ -Age age after 4 months of rapamycin supplementation, while High-dose fish exhibited early hypermethylation after 2 months followed by a subsequent decline after 4 months. Shaded areas represent 95% confidence intervals of the smooth trend. **D)** Δ -Age separated by Sex. Females exhibited epigenetic acceleration under High-dose rapamycin, whereas males showed accelerated epigenetic age under Low-dose rapamycin, indicating sex-specific treatment effect. Points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$).

Chapter 6 - Results

6.4.4. Experiment 2: Chronic rapamycin increases the prevalence of macroscopic ovarian lesions in MMR-deficient fish

Chronic rapamycin treatment was associated with an increased prevalence of macroscopic ovarian lesions in a genotype-dependent manner (Figure 20; Table S30). In *msh2* heterozygous (Sa22101) fish, rapamycin increased the probability of observing at least one macroscopic ovarian lesion (GLMM: $\chi^2 = 4.7$, $df = 2$, $p = 0.03$), with the high-dose treatment differing significantly from control (*Estimate* = 1.5, *SE* = 0.7, $p = 0.03$). Rapamycin exposure in this genotype was also associated with an increased occurrence of multiple macroscopic lesions (GLMM Low: *Estimate* = 1.2, *SE* = 0.005, $p < 0.0001$; High: *Estimate* = 2.0, *SE* = 0.0005, $p < 0.0001$).

In ABWT fish, no statistically significant treatment effects were detected for overall lesion prevalence. However, high-dose rapamycin showed a non-significant trend towards an increased occurrence of multiple macroscopic lesions (GLMM: *Estimate* = 1.7, *SE* = 0.08, $p = 0.08$).

In Experiment 2, reproductive output was not significantly affected by rapamycin treatment or genotype, with no differences observed in the proportion of fish laying eggs or in egg number across groups (Figure S13).

6.4.5. Experiment 2: High and low-dose rapamycin differentially affect locomotor ageing trajectories

Across the fish's lifespan, rapamycin was associated with dose-dependent differences in locomotor metrics (Figure 21; Table S31). High-dose fish had higher mean speed than controls (LMM Control–High: *Estimate* = -0.9, $z = -2.9$, $p = 0.009$), however, their maximum speed was also lower relative to controls (LMM Control–High: *Estimate* = 3.14, $z = 12.1$, $p < 0.0001$) and Low-dose fish (Low–High: *Estimate* = 9.47, $z = 3.7$, $p = 0.0006$). Low-dose fish, compared to

Chapter 6 - Results

controls, spent a greater proportion of time moving (LMM Control–Low: *Estimate* = 0.3, $z = 3.6$, $p = 0.0008$), and showed higher mean angular velocity (LMM Control–Low: *Estimate* = -59.8, $z = -8.1$, $p < 0.0001$). Thigmotaxis also showed a modest increase in Low-dose fish relative to control (Control - Low: *Estimate* = 0.4, $z = 1.9$, $p = 0.09$),

These results indicate that rapamycin differentially modifies behavioural ageing trajectories: High-dose treatment is associated with sustained but blunted peak performance, whereas Low-dose treatment increases overall activity and edge-preference behaviours, consistent with heightened exploratory or stress-related locomotion.

6.4.6. Experiment 2: Rapamycin modifies stress-related behaviours in a strain-dependent manner

In the NTDT, fish top-zone occupancy decreased with age, with treatment effects different by strain (GLMM: $\chi^2 = 32,995$, $df = 2$, $p < 0.0001$). In the *msh2* fish, top-zone occupancy was lower at 20 months than at 9 months under both doses (Low: *Estimate* = -1, $z = -178$, $p < 0.0001$; High: *Estimate* = -0.9, $z = -160$, $p < 0.0001$). In ABWT, a Treatment $\times$ Age interaction was detected (GLMM: $\chi^2 = 11,573$, $df = 2$, $p < 0.0001$), but no within-treatment contrasts reached significance (Figure 23; Table S32).

Freezing behaviour also increased with age (GLMM: $\chi^2 = 88,470$, $df = 1$, $p < 0.0001$) and showed a Treatment $\times$ Strain effect (GLMM: $\chi^2 = 45,426$, $df = 2$, $p < 0.0001$). Notably, in ABWT, rapamycin reduced freezing behaviour (GLMM: $\chi^2 = 46,168$, $df = 2$, $p < 0.0001$), with High-dose fish freezing less than controls (Control–High: *Estimate* = 1.7, $z = 3.9$, $p = 0.0004$), with similar trend appearing when compared to Low-dose fish (Low–High: *Estimate* = 0.9, $z = 2$, $p = 0.09$; Figure 23).

Chapter 6 - Results

These results indicate that rapamycin increases anxiety-like behaviours in the MMR-deficient genotype, and reduces freezing behaviour in ABWT, demonstrating opposing, strain-specific effects on stress-related locomotor phenotypes.

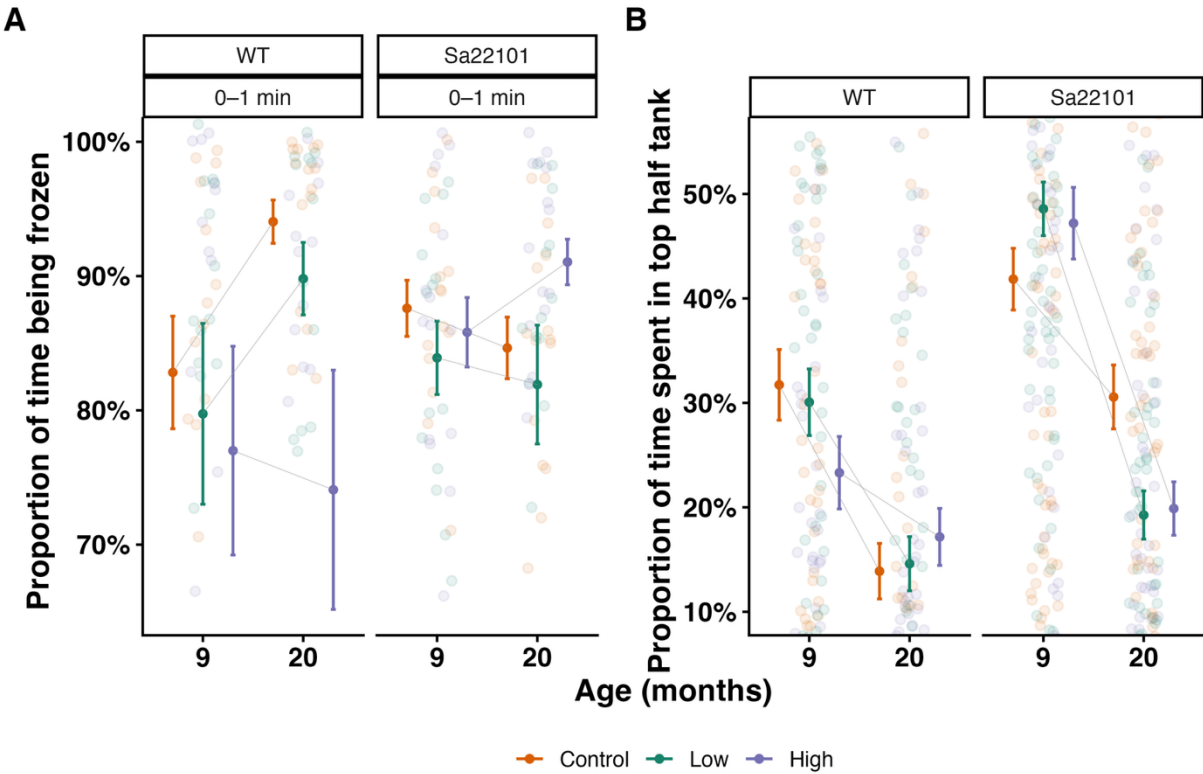


Figure 23: Effects of Rapamycin on anxiety-related behaviour in Novel Tank Diving Test

A) Proportion of time spent frozen during the first minute of the assay, **B)** Proportion of time spent in the top half of the tank across the full recording. Data are shown across ages and split by strain (WT and *msh2* heterozygous). Raw data are shown in the background, and points denote group means, and error bars represent standard errors. Statistical significance was assessed using a linear mixed-effects model (LMM) and is indicated by asterisks (* $p < 0.05$; ** $p < 0.001$; *** $p < 0.0001$). Rapamycin reduced freezing behaviour in WT fish, whereas older fish spent less time in the top half of the tank under rapamycin treatment, indicating an age and strain-specific effects on exploratory and anxiety behaviour.

6.5. Discussion

Our results provide clear evidence that rapamycin does not produce uniformly beneficial effects across ageing-related endpoints in zebrafish. Instead, rapamycin exposure was associated with increased epigenetic age estimates, dose- and strain-dependent alterations in behavioural trajectories, reduced male reproductive success, and an increased prevalence of ovarian lesions in a mismatch repair-compromised background. Collectively, these findings indicate that chronic mTOR inhibition can incur biological costs, particularly under conditions of reduced genome maintenance capacity, thereby defining limits to the beneficial effects of rapamycin reported in other systems.

6.5.1. Rapamycin and epigenetic ageing

Rapamycin increased DNA methylation-based estimates of epigenetic age in the short-term late-life experiment, with effects that were non-linear, sex-specific, and dose dependent. In this context, increased epigenetic age reflects shifts in age-associated methylation patterns interpreted by clock models as accelerated relative to chronological age, rather than uniform advancement of all ageing processes. These findings contrast with reports of reduced epigenetic ageing under rapamycin treatment in mouse liver and blood (Petkovich et al., 2017; Horvath et al., 2019). Such discrepancies are not unexpected, as epigenetic clock responses to intervention depend strongly on tissue context, clock architecture, species, and treatment regime, as highlighted in comparative analyses of epigenetic ageing (Horvath and Raj, 2018; Chapters 2 and 3). Importantly, interventions need not uniformly decelerate epigenetic age; instead, they can reshape ageing trajectories, producing non-linear or tissue-specific methylation patterns without consistent reductions in predicted biological age (Horvath and Raj, 2018;

Shindyapina et al., 2022). Consistent with this context dependence, rapamycin produced little to no change in blood-based epigenetic age estimates in rhesus macaques (*Macaca mulatta*), despite clear physiological responses to treatment (Horvath et al., 2021). Viewed alongside these studies, the dose-dependent and non-linear methylation trajectories observed here fall within the broad variability reported for epigenetic clock responses across tissues, species, and experimental design

Epigenetic clocks capture coordinated DNA methylation changes associated with biological ageing rather than mortality risk per se (Horvath and Raj, 2018). In the late-life exposure experiment, increased epigenetic age estimates coincided with altered locomotor behaviour and reduced male reproductive success, while in the lifelong exposure experiment rapamycin increased the prevalence of macroscopic ovarian lesions in the mismatch repair-deficient background. Although molecular and phenotypic measures were not assessed in the same individuals or time windows, these results indicate that rapamycin-induced shifts in epigenetic age can occur alongside changes in organismal phenotypes. In addition to altering epigenetic age estimates, rapamycin modified the relationship between chronological age and methylation-based age, producing dose-dependent non-linear trajectories rather than a single monotonic pattern. Consistent with the known sensitivity of teleost DNA methylation to environmental and nutritional inputs (Piferrer and Anastasiadi, 2023), these findings indicate that rapamycin perturbs age-associated methylation dynamics through context-dependent modulation rather than uniform deceleration or acceleration.

6.5.2. Behavioural ageing under mTOR inhibition

Rapamycin modified behavioural ageing trajectories in a dose- and strain-dependent manner under both short-term late-life and lifelong exposure. In the late-life experiment, low-dose treatment reduced overall locomotor activity and exploratory behaviour, whereas high-dose treatment increased peak locomotor performance, time spent moving, and thigmotaxis. Under lifelong exposure, rapamycin altered age-associated behavioural trajectories without preserving youthful behavioural profiles, with effects most pronounced in the mismatch repair-deficient background.

Thigmotaxis, freezing, and vertical exploration are commonly used in zebrafish to index anxiety-like or stress-related behavioural states and are sensitive to both internal physiological condition and environmental context (Stewart et al., 2014; Haghani et al., 2019; Johnson et al., 2023). In this study, rapamycin-treated fish showed increased thigmotaxis and altered freezing and exploratory patterns, particularly under high-dose and lifelong exposure, indicating heightened wall-associated swimming and changes in stress-responsive behaviour rather than delayed behavioural ageing. Importantly, these responses were not uniform: low-dose treatment tended to suppress exploratory activity, whereas high-dose treatment increased activity and wall association, and effects were further modulated by genetic background. This divergence demonstrates that rapamycin altered the structure of behavioural ageing trajectories rather than producing a consistent anxiolytic or youth-preserving effect.

Such context dependence is consistent with findings from mammalian systems. In mice, rapamycin has been reported to impair learning and memory under certain dosing regimens and treatment windows, while in other contexts modest improvements or neutral effects have been observed, depending on sex, age at initiation, and the behavioural domain assessed

(Ehninger et al., 2014; Kaeberlein, 2014). For example, chronic rapamycin treatment has been associated with deficits in spatial learning and task-dependent increases in anxiety-like behavioural measures in some studies, despite extending lifespan. Viewed in this comparative context, the present results indicate that rapamycin modifies behavioural ageing trajectories without conferring a generalised behavioural benefit, reinforcing the conclusion that behavioural responses to mTOR inhibition are contingent on dose, timing, genetic background, and the specific behavioural endpoint measured, rather than being uniformly geroprotective.

6.5.3. Reproductive suppression and ovarian lesions in a genome-unstable context

Rapamycin reduced male reproductive success following short-term late-life exposure, while female reproductive output was unaffected. Sex-dependent effects of mTOR inhibition on reproduction have been documented across vertebrate models. In rodents, chronic rapamycin treatment disrupts mTORC1 signalling in the testes of mice (*Mus musculus*) and rats (*Rattus norvegicus*), leading to arrested spermatogenesis, reduced sperm production, and testicular atrophy, consistent with impaired male reproductive investment under sustained mTOR suppression (Oliveira et al., 2017; Liu et al., 2017; Zhu et al., 2019). By contrast, effects on female reproductive output under comparable dosing regimens are less well characterised. Nonetheless, mTOR signalling is known to regulate ovarian follicle activation and follicular lifespan in mammalian systems, particularly mice, where modulation of mTOR activity influences the balance between follicle recruitment and preservation (Dou et al., 2017). These observations suggest differential sensitivity of male and female reproductive processes to mTOR inhibition. In several species, including mammals and fish, reduced

male reproductive investment has been interpreted within a life-history framework as a trade-off favouring somatic maintenance under resource limitation or stress (Lind et al., 2017). In the present study, however, reduced male reproductive success was not accompanied by improvements in other ageing-related phenotypes, indicating that reproductive suppression alone does not provide evidence for enhanced somatic maintenance under the conditions tested.

The most pronounced pathological effect observed was an increased prevalence of macroscopic ovarian lesions following lifelong rapamycin exposure in the *msh2* heterozygous background. MSH2 is a core component of the mismatch repair pathway, and partial loss of MSH2 function elevates mutation burden and cancer susceptibility in humans and model organisms (Lynch et al., 2009; Baretti and Le, 2018). In this context, rapamycin significantly increased the occurrence and multiplicity of ovarian lesions in *msh2* heterozygous fish, whereas effects in wild-type animals were comparatively modest. Despite this, rapamycin did not significantly alter egg-laying behaviour under the conditions tested, indicating that the observed pathological and behavioural effects were not driven by gross changes in reproductive output. Lesion classification was based on gross morphology; although histological validation was not performed, scoring criteria were applied consistently across genotypes and treatments.

The present data do not indicate that rapamycin is directly mutagenic. Rather, the restriction of increased lesion prevalence to the mismatch repair-compromised background, together with the observed dose dependence, supports an interaction between chronic mTOR inhibition and reduced genome maintenance capacity. This interpretation does not require a single mechanistic pathway but is consistent with a scenario in which impaired genome surveillance lowers tolerance to long-term perturbation of growth and survival signalling. Under this framework, genome maintenance

capacity emerges as a critical modifier of the long-term pathological consequences of rapamycin exposure.

6.5.4. Comparative context of rapamycin effects on ageing

Rapamycin robustly extends lifespan in mice, including when treatment is initiated late in life (Harrison et al., 2009, Miller et al., 2011), and improves selected healthspan-related outcomes in dogs and humans, particularly immune and cardiovascular function (Mannick et al., 2014, Urfer et al., 2017). The present findings do not contradict these results. Rather, they reflect differences in experimental endpoints, exposure regimes, and biological context between survival-focused studies and the current work.

Across taxa, rapamycin acts as a context-dependent geromodulator rather than a uniformly age-slowing intervention. In invertebrates, pharmacological inhibition of TOR signalling extends lifespan when administered during adulthood and activates conserved downstream processes, including autophagy and stress resistance (Bjedov et al., 2010, Kapahi et al., 2004). However, the magnitude and qualitative nature of these benefits depend on exposure timing, dose, and physiological state, supporting the view that TOR integrates nutrient availability, growth, and repair priorities rather than regulating ageing through a single linear pathway (Kapahi et al., 2004, Johnson et al., 2013).

In mammals, the strongest evidence for lifespan extension derives from long-term, survival-optimised dietary regimens. In the NIA Interventions Testing Program, (Harrison et al., 2009) administered microencapsulated rapamycin at approximately 14 ppm in diet, beginning at ~600 days of age, and observed increases in both median and maximal lifespan in genetically heterogeneous mice. Subsequent studies demonstrated clear dose- and sex-

dependence using dietary concentrations of 4.7, 14, and 42 ppm, with larger and more consistent lifespan gains at higher doses and in females, accompanied by corresponding differences in circulating rapamycin levels (Miller et al., 2011, Miller et al., 2014).

The pharmacological context is critical because rapamycin does not act exclusively through mTORC1 under all exposure conditions. Chronic administration can impair mTORC2 signalling in vivo, resulting in insulin resistance and altered metabolic homeostasis (Lamming et al., 2014, Lamming et al., 2012). Accordingly, rapamycin's net phenotype reflects a balance between beneficial reduced translation and enhanced autophagy, and costs that emerge when exposure perturbs parallel regulatory systems (Saxton and Sabatini, 2017).

Divergence is already evident within the mammalian literature when endpoints shift from survival to DNA methylation-based ageing measures. In mouse liver, dietary rapamycin slows epigenetic ageing signatures in intervention methylome analyses (Wang et al., 2017). In cultured human keratinocytes, rapamycin treatment similarly delays epigenetic ageing trajectories (Horvath et al., 2019). In contrast, rapamycin does not significantly alter blood-based epigenetic age estimates in common marmosets (Horvath et al., 2021). Even within mice, different tissues and clock architectures yield discordant responses, indicating that epigenetic clocks capture context-specific regulatory state rather than providing a uniform proxy for lifespan extension (Horvath and Raj, 2018).

This framework clarifies the interpretation of the present results. Epigenetic ageing was assessed exclusively under short-term late-life exposure, where rapamycin increased DNA methylation-based age estimates and reduced male reproductive output. In light of the broader literature, this pattern can be interpreted as an acute perturbation of epigenetic and reproductive

regulation in advanced age, rather than evidence that rapamycin fails to confer benefits under survival-optimised conditions. Short-term late-life exposure is more likely to reveal transient regulatory strain or compensatory responses than to reproduce the molecular outcomes observed under chronic dietary regimens designed to maximise lifespan.

By contrast, lifelong exposure in the present study altered behavioural ageing trajectories and increased macroscopic ovarian lesions, particularly in a mismatch repair-deficient background. These findings aligns with evidence that rapamycin does not universally replicate dietary restriction-associated benefits under compromised genome maintenance and may exacerbate vulnerabilities when DNA repair capacity is limiting (Birkisdóttir et al., 2021, Gladyshev et al., 2021). Anchoring these effects in interactions between exposure regime and repair competence provides a mechanistic explanation for the observed ovarian lesion phenotype without invoking species-specific exceptionalism.

The rapamycin literature supports a model in which mTOR inhibition yields lifespan benefits under a relatively narrow set of dosing and exposure conditions optimised for survival, while alternative regimes expose trade-offs across molecular, reproductive, behavioural, and oncogenic domains. The present findings align with this model, demonstrating that rapamycin's effects on ageing are conditional rather than universally beneficial, and that divergence from lifespan studies reflects differences in dose, timing, duration, and phenotypic resolution rather than inconsistency in rapamycin biology.

6.5.5. Conclusions

In summary, this study shows that rapamycin does not exert a uniform, directionally beneficial effect on ageing-related phenotypes in zebrafish. Instead, pharmacological mTOR inhibition reshaped molecular, behavioural, and pathological ageing trajectories in a manner that depended on dose, sex, and genetic context. Increased epigenetic age estimates, altered behavioural trajectories, reduced male reproductive success, and elevated prevalence for ovarian lesions in a mismatch repair-compromised background can be associated with adverse shifts in specific molecular and phenotypic endpoints under certain conditions. In this setting, epigenetic clocks acted as descriptive biomarkers that tracked broader treatment-associated changes in biological state rather than serving as direct proxies for lifespan.

These findings emphasise that mTOR inhibition is not universally beneficial and that genome maintenance capacity represents a critical axis influencing its long-term consequences. Future studies incorporating tissue-specific molecular profiling, alternative dosing strategies, and explicit consideration of genome stability will be necessary to define the contexts in which rapamycin may confer benefit without incurring adverse effects.

6.6. Data Availability

All sequencing data, including RRBS data, generated for this project are available under accession (under embargo until acceptance for publication), <https://www.ebi.ac.uk/ena/browser/view/PRJEB107120>. All scripts and supplementary data are available at the git repository: <https://github.com/asdfjohnny1/rapamycin-somatic-epigenetic-age.git>

6.7. Contributions

The studies were conceptualised and designed by S.I , A.M. , H.C

Data analysis and visualisation was carried out by J.C.D.C. ,

Samples were collected and processed by H.C

All code was conceived and generated by J.C.D.C.

Supervision was provided by S.I., R.R, A.M

Funding was provided by S.I., R.R

The first draft of the manuscript was written by J.C.D.C. and S.I.

6.8. Ethics declaration

Competing interests: All authors declare no conflict or competing interests

All zebrafish work was completed in accordance with institutional, local and national laws and governing bodies.

6.9. Extended data

Appendix Chapter 4 for supplementary methods, figures, tables and legends.

All files and processed data files are available on the GitHub page :
<https://github.com/asdfjohnny1/rapamycin-somatic-epigenetic-age.git>

6.10. Abbreviations

mTORC1 - mTOR complex 1

Chapter 6 - References

mTORC2 - mTOR complex 2

ABWT - Wild-type strain

MMR - Mismatch repair

CpG - 5'-C-phosphate-G-3'

LMM - Linear mixed models

GLMM - Generalised linear mixed models

6.11. References

BARETTI, M. & LE, D. T. 2018. DNA mismatch repair in cancer. *Pharmacol Ther*, 189, 45-62.

BARZILAI, N., CRANDALL, J. P., KRITCHEVSKY, S. B. & ESPELAND, M. A. 2016. Metformin as a Tool to Target Aging. *Cell Metab*, 23, 1060-1065.

BATES, D., MÄCHLER, M., BOLKER, B., & WALKER, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*.

BELL, C. G., LOWE, R., ADAMS, P. D., BACCARELLI, A. A., BECK, S., BELL, J. T., CHRISTENSEN, B. C., GLADYSHEV, V. N., HEIJMANS, B. T., HORVATH, S., IDEKER, T., ISSA, J.-P. J., KELSEY, K. T., MARIONI, R. E., REIK, W., RELTON, C. L., SCHALKWYK, L. C., TESCHENDORFF, A. E., WAGNER, W., ZHANG, K. & RAKYAN, V. K. 2019. DNA methylation aging clocks: challenges and recommendations. *Genome Biology*, 20.

BIRKISDÓTTIR, M. B., JAARSMA, D., BRANDT, R. M. C., BARNHOORN, S., VAN VLIET, N., IMHOLZ, S., VAN OOSTROM, C. T., NAGARAJAH, B., PORTILLA FERNÁNDEZ, E., ROKS, A. J. M., ELGERSMA, Y., VAN STEEG, H., FERREIRA, J. A., PENNING, J. L. A., HOEIJMAKERS, J. H. J., VERMEIJ, W. P. & DOLLÉ, M. E. T.

Chapter 6 - References

2021. Unlike dietary restriction, rapamycin fails to extend lifespan and reduce transcription stress in progeroid DNA repair-deficient mice. *Aging Cell*, 20, e13302.

BJEDOV, I., TOIVONEN, J. M., KERR, F., SLACK, C., JACOBSON, J., FOLEY, A. & PARTRIDGE, L. 2010. Mechanisms of Life Span Extension by Rapamycin in the Fruit Fly *Drosophila melanogaster*. *Cell Metabolism*, 11, 35-46.

BLAGOSKLONNY, M. V. 2010. Why men age faster but reproduce longer than women: mTOR and evolutionary perspectives. *Aging*, 2, 265-273.

DOU, X., SUN, Y., LI, J., ZHANG, J., HAO, D., LIU, W., WU, R., KONG, F. & PENG, X. 2017. Short-term rapamycin treatment increases ovarian lifespan in young and middle-aged female mice. *Aging Cell*, 16, 825-836.

EHNINGER, D., NEFF, F. & XIE, K. 2014. Longevity, aging and rapamycin. *Cellular and Molecular Life Sciences*, 71, 4325-4346.

FONTANA, L., PARTRIDGE, L. & LONGO, V. D. 2010. Extending Healthy Life Span—From Yeast to Humans. *Science*, 328, 321-326.

GLADYSHEV, V. N., KRITCHEVSKY, S. B., CLARKE, S. G., CUERVO, A. M., FIEHN, O., DE MAGALHÃES, J. P., MAU, T., MAES, M., MORITZ, R. L., NIEDERNHOFER, L. J., VAN SCHAFTINGEN, E., TRANAH, G. J., WALSH, K., YURA, Y., ZHANG, B. & CUMMINGS, S. R. 2021. Molecular damage in aging. *Nature Aging*, 1, 1096-1106.

HAGHANI, S., KARIA, M., CHENG, R.-K. & MATHURU, A. S. 2019. An Automated Assay System to Study Novel Tank Induced Anxiety. *Frontiers in Behavioral Neuroscience*, 13.

HARRISON, D. E., STRONG, R., ALLISON, D. B., AMES, B. N., ASTLE, C. M., ATAMNA, H., FERNANDEZ, E., FLURKEY, K., JAVORS, M. A., NADON, N. L., NELSON, J. F., PLETCHER, S., SIMPKINS, J. W., SMITH, D., WILKINSON, J. E. &

Chapter 6 - References

- MILLER, R. A. 2014. Acarbose, 17--estradiol, and nordihydroguaiaretic acid extend mouse lifespan preferentially in males. *Aging Cell*, 13, 273-82.
- HARRISON, D. E., STRONG, R., SHARP, Z. D., NELSON, J. F., ASTLE, C. M., FLURKEY, K., NADON, N. L., WILKINSON, J. E., FRENKEL, K., CARTER, C. S., PAHOR, M., JAVORS, M. A., FERNANDEZ, E. & MILLER, R. A. 2009. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature*, 460, 392-395.
- HORVATH, S., LU, A. T., COHEN, H. & RAJ, K. 2019. Rapamycin retards epigenetic ageing of keratinocytes independently of its effects on replicative senescence, proliferation and differentiation. *Aging*, 11, 3238-3249.
- HORVATH, S. & RAJ, K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*, 19, 371-384.
- HORVATH, S., ZOLLER, J. A., HAGHANI, A., LU, A. T., RAJ, K., JASINSKA, A. J., MATTISON, J. A. & SALMON, A. B. 2021. DNA methylation age analysis of rapamycin in common marmosets. *GeroScience*, 43, 2413-2425.
- INGRAM, D. K. & ROTH, G. S. 2015. Calorie restriction mimetics: can you have your cake and eat it, too? *Ageing Res Rev*, 20, 46-62.
- JOHNSON, A., LOH, E., VERBITSKY, R., SLESSOR, J., FRANCAZAK, B. C., SCHALOMON, M. & HAMILTON, T. J. 2023. Examining behavioural test sensitivity and locomotor proxies of anxiety-like behaviour in zebrafish. *Scientific Reports*, 13.
- JOHNSON, S. C., RABINOVITCH, P. S. & KAEBERLEIN, M. 2013. mTOR is a key modulator of ageing and age-related disease. *Nature*, 493, 338-345.
- KAEBERLEIN, M. 2014. Rapamycin and ageing: when, for how long, and how much? *J Genet Genomics*, 41, 459-63.

Chapter 6 - References

KAPAH, P., ZID, B. M., HARPER, T., KOSLOVER, D., SAPIN, V. & BENZER, S. 2004. Regulation of lifespan in *Drosophila* by modulation of genes in the TOR signaling pathway. *Curr Biol*, 14, 885-90.

KENNEDY, B. K., BERGER, S. L., BRUNET, A., CAMPISI, J., CUERVO, A. M., EPEL, E. S., FRANCESCHI, C., LITHGOW, G. J., MORIMOTO, R. I., PESSIN, J. E., RANDO, T. A., RICHARDSON, A., SCHADT, E. E., WYSS-CORAY, T. & SIERRA, F. 2014. Geroscience: Linking Aging to Chronic Disease. *Cell*, 159, 709-713.

KIRKWOOD, T. B. L. 1977. Evolution of ageing. *Nature*, 270, 301-304.

KRUEGER, F. & ANDREWS, S. R. 2011. Bismark: a flexible aligner and methylation caller for Bisulphite-Seq applications. *Bioinformatics*, 27, 1571-2.

LAMMING, D. W., MIHAYLOVA, M. M., KATAJISTO, P., BAAR, E. L., YILMAZ, O. H., HUTCHINS, A., GULTEKIN, Y., GAITHER, R. & SABATINI, D. M. 2014. Depletion of Rictor, an essential protein component of mTORC2, decreases male lifespan. *Aging Cell*, 13, 911-7.

LAMMING, D. W., YE, L., KATAJISTO, P., GONCALVES, M. D., SAITOH, M., STEVENS, D. M., DAVIS, J. G., SALMON, A. B., RICHARDSON, A., AHIMA, R. S., GUERTIN, D. A., SABATINI, D. M. & BAUR, J. A. 2012. Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. *Science*, 335, 1638-43.

LENTH, R. V. 2021. Emmeans: Estimated Marginal Means, Aka Least-Squares Means.

LIND, M. I., CHEN, H. Y., CORTAZAR-CHINARRO, M. & MAKLAKOV, A. A. 2017. Rapamycin additively extends lifespan in short- and long-lived lines of the nematode *Caenorhabditis remanei*. *Exp Gerontol*, 90, 79-82.

LIU, S., HUANG, L., GENG, Y., HE, J., CHEN, X., XU, H., LI, R., WANG, Y., DING, Y. & LIU, X. 2017. Rapamycin inhibits spermatogenesis by changing the

Chapter 6 - References

autophagy status through suppressing mechanistic target of rapamycin-p70S6 kinase in male rats. *Mol Med Rep*, 16, 4029-4037.

LÓPEZ-OTÍN, C., BLASCO, M. A., PARTRIDGE, L., SERRANO, M. & KROEMER, G. 2013. The hallmarks of aging. *Cell*, 153, 1194-217.

LYNCH, H. T., LYNCH, P. M., LANSPA, S. J., SNYDER, C. L., LYNCH, J. F. & BOLAND, C. R. 2009. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin Genet*, 76, 1-18.

MAKLAKOV, A. A. & CHAPMAN, T. 2019. Evolution of ageing as a tangle of trade-offs: energy versus function. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20191604.

MANNICK, J. B., DEL GIUDICE, G., LATTANZI, M., VALIANTE, N. M., PRAESTGAARD, J., HUANG, B., LONETTO, M. A., MAECKER, H. T., KOVARIK, J., CARSON, S., GLASS, D. J. & KLINKSTEIN, L. B. 2014. mTOR inhibition improves immune function in the elderly. *Sci Transl Med*, 6, 268ra179.

MILLER, R. A., HARRISON, D. E., ASTLE, C. M., BAUR, J. A., BOYD, A. R., DE CABO, R., FERNANDEZ, E., FLURKEY, K., JAVORS, M. A., NELSON, J. F., ORIHUELA, C. J., PLETCHER, S., SHARP, Z. D., SINCLAIR, D., STARNES, J. W., WILKINSON, J. E., NADON, N. L. & STRONG, R. 2011. Rapamycin, But Not Resveratrol or Simvastatin, Extends Life Span of Genetically Heterogeneous Mice. *The Journals of Gerontology: Series A*, 66A, 191-201.

MILLER, R. A., HARRISON, D. E., ASTLE, C. M., FERNANDEZ, E., FLURKEY, K., HAN, M., JAVORS, M. A., LI, X., NADON, N. L., NELSON, J. F., PLETCHER, S., SALMON, A. B., SHARP, Z. D., VAN ROEKEL, S., WINKLEMAN, L. & STRONG, R. 2014. Rapamycin-mediated lifespan increase in mice is dose and sex dependent and metabolically distinct from dietary restriction. *Aging Cell*, 13, 468-77.

Chapter 6 - References

- OLIVEIRA, P. F., CHENG, C. Y. & ALVES, M. G. 2017. Emerging Role for Mammalian Target of Rapamycin in Male Fertility. *Trends Endocrinol Metab*, 28, 165-167.
- PAPADOPOLI, D., BOULAY, K., KAZAK, L., POLLAK, M., MALLETTE, F., TOPISIROVIC, I. & HULEA, L. 2019. mTOR as a central regulator of lifespan and aging. *F1000Research*, 8, 998.
- PARTRIDGE, L., DEELEN, J. & SLAGBOOM, P. E. 2018. Facing up to the global challenges of ageing. *Nature*, 561, 45-56.
- PETKOVICH, D. A., PODOLSKIY, D. I., LOBANOV, A. V., LEE, S. G., MILLER, R. A. & GLADYSHEV, V. N. 2017. Using DNA Methylation Profiling to Evaluate Biological Age and Longevity Interventions. *Cell Metab*, 25, 954-960.e6.
- PIFERRER, F. & ANASTASIADI, D. 2023. Age estimation in fishes using epigenetic clocks: Applications to fisheries management and conservation biology. *Frontiers in Marine Science*, 10.
- R CORE TEAM 2022. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.
- ROARK, K. M. & IFFLAND, P. H. 2025. Rapamycin for longevity: the pros, the cons, and future perspectives. *Frontiers in Aging*, 6.
- ROMERO-FERRERO, F., BERGOMI, M. G., HINZ, R. C., HERAS, F. J. H. & DE POLAVIEJA, G. G. 2019. idtracker.ai: tracking all individuals in small or large collectives of unmarked animals. *Nat Methods*, 16, 179-182.
- SAXTON, R. A. & SABATINI, D. M. 2017. mTOR Signaling in Growth, Metabolism, and Disease. *Cell*, 168, 960-976.
- SHINDYAPINA, A. V., CHO, Y., KAYA, A., TYSHKOVSKIY, A., CASTRO, J. P., DEIK, A., GORDEVICIUS, J., POGANIK, J. R., CLISH, C. B., HORVATH, S., PESHKIN, L. &

Chapter 6 - References

GLADYSHEV, V. N. 2022. Rapamycin treatment during development extends life span and health span of male mice and *Daphnia magna*. *Science Advances*, 8.

STEWART, S., GOMEZ, W., ALAN, ARMSTRONG, E., BENJAMIN, HENNER, A. & STANKUNAS, K. 2014. Sequential and Opposing Activities of Wnt and BMP Coordinate Zebrafish Bone Regeneration. *Cell Reports*, 6, 482-498.

URFER, S. R., KAEBERLEIN, T. L., MAILHEAU, S., BERGMAN, P. J., CREEVY, K. E., PROMISLOW, D. E. L. & KAEBERLEIN, M. 2017. A randomized controlled trial to establish effects of short-term rapamycin treatment in 24 middle-aged companion dogs. *GeroScience*, 39, 117-127.

WANG, T., TSUI, B., KREISBERG, J. F., ROBERTSON, N. A., GROSS, A. M., YU, M. K., CARTER, H., BROWN-BORG, H. M., ADAMS, P. D. & IDEKER, T. 2017. Epigenetic aging signatures in mice livers are slowed by dwarfism, calorie restriction and rapamycin treatment. *Genome Biology*, 18.

WEICHHART, T. 2018. mTOR as Regulator of Lifespan, Aging, and Cellular Senescence: A Mini-Review. *Gerontology*, 64, 127-134.

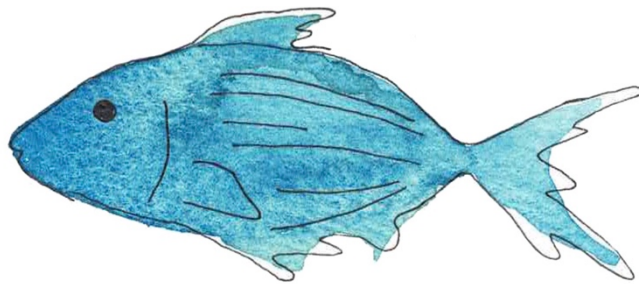
WOOD, S. N. 2017. Generalized Additive Models: An Introduction with R. *Second Edition (2nd ed.)*. Chapman and Hall/CRC.

ZHU, Z., YUE, Q., XIE, J., ZHANG, S., HE, W., BAI, S., TIAN, S., ZHANG, Y., XIONG, M., SUN, Z., HUANG, C., LI, Y., ZHENG, K. & YE, L. 2019. Rapamycin-mediated mTOR inhibition impairs silencing of sex chromosomes and the pachytene piRNA pathway in the mouse testis. *Aging (Albany NY)*, 11, 185-208.

ZIMNIAK, P. 2012. What is the Proximal Cause of Aging? *Frontiers in Genetics*, 3.

CHAPTER SEVEN

Final Discussion



7. Chapter 7

7.1. Overview

This thesis examined how somatic ageing is shaped by tissue context, germline status, and physiological intervention in a vertebrate system. Using zebrafish as an experimental model, a combination of transcriptomic analyses, DNA methylation profiling, longitudinal epigenetic modelling, and behavioural and pathological assays were used to investigate how ageing-related regulatory states are organised and modified.

Across Chapters 2–6, three complementary strategies were used. First, germline ablation was employed as a genetic perturbation that removes reproductive signalling without altering somatic genotype. Second, epigenetic clocks were treated not as fixed biomarkers but as models whose behaviour depends on tissue biology and feature architecture. Third, pharmacological inhibition of mTOR via rapamycin was used to perturb nutrient-sensing and growth signalling under different exposure regimes and genetic backgrounds.

Rather than seeking to define or optimise a single quantitative measure of biological age, this thesis contributes (i) a cross-tissue benchmark of clock architecture, and (ii) longitudinal evidence that germline status shapes the temporal structure and stability of somatic epigenetic ageing. By integrating these approaches, this thesis examined how reproductive status, tissue context, and signalling pathways influence the organisation, stability, and temporal structure of molecular ageing. In this final chapter, these findings are synthesised to draw conceptual conclusions about the regulatory role of the germline in ageing, identify the questions they raise, and position the work within the broader field of ageing biology.

7.2. Conceptual conclusions

7.2.1. Ageing trajectories are regulated, context-dependent processes

A central conclusion of this thesis is that somatic ageing does not proceed as a uniform, cell-autonomous accumulation of molecular damage. Instead, ageing trajectories are regulated processes that depend on tissue identity, reproductive signalling state, and physiological context. Chapter 4 demonstrated that epigenetic ageing across somatic tissues follows a mosaic architecture. Tissue-specific CpGs encode local ageing trajectories shaped by proliferative history and regenerative capacity, while a smaller set of overlapping CpGs captures conserved regulatory vulnerability across tissues. Similar tissue stratification of epigenetic ageing has been observed in mammals, where post-mitotic tissues exhibit more stable methylation trajectories than proliferative epithelia (Horvath, 2013, Hannum et al., 2013, Slieker et al., 2018). Notably, overlapping CpGs did not show homogeneous methylation behaviour between tissues but instead exhibited tissue-specific dynamics at the level of entropy across tissues and correlation structure. This indicates that conserved regulatory targets are differentially modulated depending on tissue context, consistent with models in which ageing reflects destabilisation of regulatory networks rather than uniform molecular decay (Teschendorff et al., 2013, López-Otín et al., 2013).

Chapter 5 further showed that germline ablation alters the shape of somatic ageing trajectories rather than producing a uniform shift in biological age. Germline-free fish exhibited reduced epigenetic age early in life but weakened late-life deceleration and increased regulatory instability. Comparable non-linear ageing trajectories have been reported in longitudinal human epigenetic clock analyses, where age-associated slopes decelerate or plateau at advanced ages (Horvath, 2013, Snir et al., 2019, Marioni et al., 2019).

Chapter 7 - Conceptual conclusions

Together, these findings support a model in which ageing trajectories reflect long-term regulation of somatic maintenance programmes, shaped by tissue dynamics and inter-tissue signalling, rather than inevitable damage accumulation alone. By combining tissue-resolved epigenetic profiling with experimental manipulation of germline status, this work provides experimental evidence that such context dependence reflects active regulatory control rather than passive tissue turnover alone

7.2.2. The germline acts as a regulator of somatic ageing, not merely a cost

A second major conclusion is that the germline influences somatic ageing through regulatory mechanisms rather than through energetic depletion associated with reproduction. In Chapter 2, germline ablation reorganised baseline somatic transcription, altered gene architecture-associated regulation, weakened DREAM-mediated repression, and reshaped the timing and resolution of DNA damage responses. These effects were evident prior to acute stress and extended to transposable element regulation, indicating coordinated chromatin-level control rather than reactive compensation. Similar germline-dependent regulation of somatic maintenance has been described in invertebrate systems, where germline removal activates stress resistance and genome protection pathways independently of fertility (Hsin and Kenyon, 1999, Ermolaeva et al., 2013).

Chapter 5 demonstrated that germline ablation modifies somatic epigenetic ageing trajectories across the life course. Early-life molecular youth in germline-free fish was accompanied by reduced late-life epigenetic stabilisation, increased stochastic divergence at robust CpG sites, and altered behavioural output. This pattern is inconsistent with a simple energetic interpretation of the disposable soma theory, which predicts sustained improvement in somatic maintenance following reduced reproductive

investment (Kirkwood, 1977). Instead, the data are more consistent with frameworks in which the germline provides stabilising regulatory input, that constrains late-life epigenetic drift and maintains long-term regulatory coherence. Removal of this input releases early-life protective programmes but compromises long-term epigenetic stability, consistent with developmental and regulatory trade-off models of ageing (Maklakov and Immler, 2016, Maklakov and Chapman, 2019).

7.2.3. Epigenetic clocks measure different aspects of ageing, not a single truth

A third conceptual conclusion concerns the interpretation of epigenetic clocks themselves. Chapters 3–5 collectively show that epigenetic clocks are not interchangeable measures of a single underlying ageing process. Clock performance and behaviour depended strongly on tissue context and CpG aggregation strategy. Tissue-specific clocks benefited from regional aggregation, whereas multi-tissue clocks required locus-level resolution. Comparable context dependence has been reported in human clock comparisons, where clocks trained on different tissues or objectives show divergent associations with health outcomes and mortality (Horvath and Raj, 2018, Bell et al., 2019). Crucially, different clock architectures responded differently to the same perturbations, including germline ablation and rapamycin exposure. These discrepancies were systematic and reproducible, indicating that clocks emphasise different modes of age-associated methylation change rather than introducing random error. Similar discordance has been reported in comparative intervention studies, where epigenetic clocks respond inconsistently across tissues and species despite clear physiological effects (Horvath et al., 2021).

Chapter 7 - Conceptual conclusions

Epigenetic age should therefore be interpreted as a model-dependent readout of regulatory state, sensitive to tissue biology, feature representation, and experimental context, rather than as a universal scalar measure of biological ageing.

7.2.4. Interventions reshape molecular trajectories rather than uniformly slowing them

Chapter 6 demonstrated that pharmacological mTOR inhibition via rapamycin does not exert uniformly beneficial effects across ageing-related endpoints. Instead, rapamycin reshaped molecular, behavioural, reproductive, and pathological trajectories in a dose-, sex-, strain-, and exposure-dependent manner. Short-term late-life rapamycin exposure increased epigenetic age estimates, altered locomotor behaviour, and reduced male reproductive success. Lifelong exposure modified behavioural ageing trajectories and increased ovarian lesion prevalence in a mismatch repair-deficient background, even though rapamycin produces well-established lifespan-extending effects in survival-optimised mouse studies (Harrison et al., 2009, Miller et al., 2011). Such context dependence is increasingly recognised across taxa. Rapamycin extends lifespan in invertebrates and mammals under specific dosing and timing regimes but can impair metabolic homeostasis, reproduction, or cognitive function under others, particularly when mTORC2 signalling is disrupted (Lamming et al., 2012, Ehninger et al., 2014, Saxton and Sabatini, 2017).

These results indicate that rapamycin acts as a context-dependent geromodulator rather than a universal suppressor of ageing. In this setting, epigenetic clocks functioned as descriptive biomarkers of altered regulatory state rather than proxies for lifespan extension.

7.3. Integration with ageing theory

Taken together, the findings of this thesis align most closely with evolutionary and developmental theories of ageing that emphasise regulatory trade-offs rather than energetic depletion alone.

The stage-dependent effects of germline ablation are difficult to reconcile with a strictly energetic/resource-allocation interpretation of the disposable soma theory (Kirkwood, 1977). Instead, they support models in which reproductive status actively shapes somatic regulatory programmes, influencing both early-life robustness and late-life stability (Gladyshev et al., 2021)

The preferential involvement of developmental and regulatory gene networks in age-associated methylation change, together with Polycomb enrichment at conserved CpGs, supports the developmental theory of ageing, in which ageing reflects the progressive destabilisation of regulatory programmes established during development (Maklakov and Chapman, 2019).

Rapamycin's conditional effects further reinforce this view. Modulation of mTOR signalling did not uniformly preserve youthful states but exposed trade-offs between growth regulation, genome maintenance, behaviour, and pathology, consistent with antagonistic pleiotropy and hyperfunction-based models of ageing (Williams, 1957, Blagosklonny, 2010).

7.4. Conceptual questions raised by this work

This thesis raises several testable conceptual questions. First, which germline-derived signals stabilise late-life somatic epigenetic regulation?

Chapter 7 - Conceptual questions raised by this work

Distinguishing between endocrine, paracrine, and chromatin-mediated mechanisms will require targeted perturbation experiments rather than additional descriptive profiling. Second, when epigenetic clocks yield divergent responses to the same perturbation, do these differences reflect selective alteration of distinct modes of methylation ageing, or changes in cell-type composition within tissues? Single-cell or lineage-resolved DNA methylation profiling would allow these alternatives to be directly disentangled. Third, are the altered ageing trajectories induced by germline ablation established during development, or are they actively maintained by adult reproductive cues? Temporally controlled germline manipulation would provide a direct test of this distinction. Finally, under what conditions can interventions modify ageing trajectories without increasing late-life regulatory instability or pathological burden? Addressing this question will require experimental designs that explicitly quantify trade-offs across molecular, behavioural, and pathological domains, rather than assuming concordance among ageing-related endpoints.

A limitation of the germline-ablation framework used here is that loss of *dnd1* biases zebrafish development towards phenotypic males, constraining inference primarily to male somatic ageing trajectories (Siegfried and Nüsslein-Volhard, 2008, Tzung et al., 2015). Given well-established sex differences in ageing regulation across vertebrates, including divergent lifespan trajectories and molecular ageing dynamics, germline–soma interactions may operate differently in female contexts (Moses et al., 2023, Austad and Fischer, 2016). Future work could address this limitation by combining germline manipulation with hormonal feminisation approaches, such as oestrogen (EE2) exposure, which can induce or maintain female somatic phenotypes in zebrafish, thereby allowing germline status to be disentangled from somatic sex and directly tested for sex-specific regulatory

effects on epigenetic ageing (Andersen et al., 2003, von Hofsten and Olsson, 2005, Hu et al., 2021)

7.5. Final remarks

This thesis demonstrates that somatic ageing in zebrafish is shaped by tissue context, germline-dependent regulation, and physiological intervention. By treating epigenetic clocks as context-sensitive models rather than fixed biomarkers, it shows that ageing trajectories reflect regulated biological processes with inherent trade-offs.

More broadly, the work argues against reducing ageing to a single quantitative variable. Molecular, behavioural, and pathological dimensions of ageing capture distinct aspects of biological change and do not necessarily move in parallel. Understanding how ageing trajectories are regulated, stabilised, and destabilised is therefore essential for both mechanistic insight and rational evaluation of interventions.

7.6. References

- ANDERSEN, L., HOLBECH, H., GESSBO, A., NORRGREN, L. & PETERSEN, G. I. 2003. Effects of exposure to 17 α -ethinylestradiol during early development on sexual differentiation and induction of vitellogenin in zebrafish (*Danio rerio*). *Comp Biochem Physiol C Toxicol Pharmacol*, 134, 365-74.
- AUSTAD, S. N. & FISCHER, K. E. 2016. Sex Differences in Lifespan. *Cell Metab*, 23, 1022-1033.
- BELL, C. G., LOWE, R., ADAMS, P. D., BACCARELLI, A. A., BECK, S., BELL, J. T., CHRISTENSEN, B. C., GLADYSHEV, V. N., HEIJMANS, B. T., HORVATH, S.,

Chapter 7 - References

- IDEKER, T., ISSA, J.-P. J., KELSEY, K. T., MARIONI, R. E., REIK, W., RELTON, C. L., SCHALKWYK, L. C., TESCHENDORFF, A. E., WAGNER, W., ZHANG, K. & RAKYAN, V. K. 2019. DNA methylation aging clocks: challenges and recommendations. *Genome Biology*, 20.
- BLAGOSKLONNY, M. V. 2010. Why men age faster but reproduce longer than women: mTOR and evolutionary perspectives. *Aging*, 2, 265-273.
- EHNINGER, D., NEFF, F. & XIE, K. 2014. Longevity, aging and rapamycin. *Cellular and Molecular Life Sciences*, 71, 4325-4346.
- ERMOLAEVA, M. A., SEGREF, A., DAKHOVNIK, A., OU, H.-L., SCHNEIDER, J. I., UTERMÖHLEN, O., HOPPE, T. & SCHUMACHER, B. 2013. DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance. *Nature*, 501, 416-420.
- GLADYSHEV, V. N., KRITCHEVSKY, S. B., CLARKE, S. G., CUERVO, A. M., FIEHN, O., DE MAGALHÃES, J. P., MAU, T., MAES, M., MORITZ, R. L., NIEDERNHOFER, L. J., VAN SCHAFTINGEN, E., TRANAH, G. J., WALSH, K., YURA, Y., ZHANG, B. & CUMMINGS, S. R. 2021. Molecular damage in aging. *Nature Aging*, 1, 1096-1106.
- HANNUM, G., GUINNEY, J., ZHAO, L., ZHANG, L., HUGHES, G., SADDA, S., KLOTZLE, B., BIBIKOVA, M., FAN, J.-B., GAO, Y., DECONDE, R., CHEN, M., RAJAPAKSE, I., FRIEND, S., IDEKER, T. & ZHANG, K. 2013. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell*, 49, 359-367.
- HARRISON, D. E., STRONG, R., SHARP, Z. D., NELSON, J. F., ASTLE, C. M., FLURKEY, K., NADON, N. L., WILKINSON, J. E., FRENKEL, K., CARTER, C. S., PAHOR, M., JAVORS, M. A., FERNANDEZ, E. & MILLER, R. A. 2009. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature*, 460, 392-395.
- HORVATH, S. 2013. DNA methylation age of human tissues and cell types. *Genome Biology*, 14, R115.

Chapter 7 - References

- HORVATH, S. & RAJ, K. 2018. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*, 19, 371-384.
- HORVATH, S., ZOLLER, J. A., HAGHANI, A., LU, A. T., RAJ, K., JASINSKA, A. J., MATTISON, J. A. & SALMON, A. B. 2021. DNA methylation age analysis of rapamycin in common marmosets. *GeroScience*, 43, 2413-2425.
- HSIN, H. & KENYON, C. 1999. Signals from the reproductive system regulate the lifespan of *C. elegans*. *Nature*, 399, 362-366.
- HU, Y., LI, D., MA, X., LIU, R., QI, Y., YUAN, C. & HUANG, D. 2021. Effects of 2,4-dichlorophenol exposure on zebrafish: Implications for the sex hormone synthesis. *Aquat Toxicol*, 236, 105868.
- KIRKWOOD, T. B. L. 1977. Evolution of ageing. *Nature*, 270, 301-304.
- LAMMING, D. W., YE, L., KATAJISTO, P., GONCALVES, M. D., SAITOH, M., STEVENS, D. M., DAVIS, J. G., SALMON, A. B., RICHARDSON, A., AHIMA, R. S., GUERTIN, D. A., SABATINI, D. M. & BAUR, J. A. 2012. Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. *Science*, 335, 1638-43.
- LÓPEZ-OTÍN, C., BLASCO, M. A., PARTRIDGE, L., SERRANO, M. & KROEMER, G. 2013. The hallmarks of aging. *Cell*, 153, 1194-217.
- MAKLAKOV, A., ALEXEI & IMMLER, S. 2016. The Expensive Germline and the Evolution of Ageing. *Current Biology*, 26, R577-R586.
- MAKLAKOV, A. A. & CHAPMAN, T. 2019. Evolution of ageing as a tangle of trade-offs: energy versus function. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20191604.
- MARIONI, R. E., SUDERMAN, M., CHEN, B. H., HORVATH, S., BANDINELLI, S., MORRIS, T., BECK, S., FERRUCCI, L., PEDERSEN, N. L., RELTON, C. L., DEARY, I. J. & HÄGG, S. 2019. Tracking the Epigenetic Clock Across the Human Life Course: A Meta-analysis of Longitudinal Cohort Data. *J Gerontol A Biol Sci Med Sci*, 74, 57-61.

Chapter 7 - References

- MILLER, R. A., HARRISON, D. E., ASTLE, C. M., BAUR, J. A., BOYD, A. R., DE CABO, R., FERNANDEZ, E., FLURKEY, K., JAVORS, M. A., NELSON, J. F., ORIHUELA, C. J., PLETCHER, S., SHARP, Z. D., SINCLAIR, D., STARNES, J. W., WILKINSON, J. E., NADON, N. L. & STRONG, R. 2011. Rapamycin, But Not Resveratrol or Simvastatin, Extends Life Span of Genetically Heterogeneous Mice. *The Journals of Gerontology: Series A*, 66A, 191-201.
- MOSES, E., ATLAN, T., SUN, X., FRANEK, R., SIDDIQUI, A., MARINOV, G. K., SHIFMAN, S., ZUCKER, D. M., ORON-GOTTESMAN, A., GREENLEAF, W. J., COHEN, E., RAM, O. & HAREL, I. 2023. The germline regulates longevity and somatic repair in a sex-specific manner. Cold Spring Harbor Laboratory.
- SAXTON, R. A. & SABATINI, D. M. 2017. mTOR Signaling in Growth, Metabolism, and Disease. *Cell*, 168, 960-976.
- SIEGFRIED, K. R. & NÜSSLEIN-VOLHARD, C. 2008. Germ line control of female sex determination in zebrafish. *Dev Biol*, 324, 277-87.
- SLIEKER, R. C., RELTON, C. L., GAUNT, T. R., SLAGBOOM, P. E. & HEIJMANS, B. T. 2018. Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception. *Epigenetics & Chromatin*, 11.
- SNIR, S., FARRELL, C. & PELLEGRINI, M. 2019. Human epigenetic ageing is logarithmic with time across the entire lifespan. *Epigenetics*, 14, 912-926.
- TESCHENDORFF, A. E., WEST, J. & BECK, S. 2013. Age-associated epigenetic drift: implications, and a case of epigenetic thrift? *Human Molecular Genetics*, 22, R7-R15.
- TZUNG, K. W., GOTO, R., SAJU, J. M., SREENIVASAN, R., SAITO, T., ARAI, K., YAMAHA, E., HOSSAIN, M. S., CALVERT, M. E. K. & ORBÁN, L. 2015. Early depletion of primordial germ cells in zebrafish promotes testis formation. *Stem Cell Reports*, 4, 61-73.

Chapter 7 - References

- VON HOFSTEN, J. & OLSSON, P. E. 2005. Zebrafish sex determination and differentiation: involvement of FTZ-F1 genes. *Reprod Biol Endocrinol*, 3, 63.
- WILLIAMS, G. C. 1957. PLEIOTROPY, NATURAL SELECTION, AND THE EVOLUTION OF SENESCENCE. *Evolution*, 11, 398-411.

Appendix

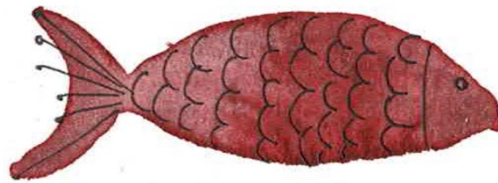


TABLE OF SUPPLEMENTARY FIGURES

FIGURE S 1: EXPERIMENTAL DESIGN.....	336
FIGURE S 2: STRAIN-SPECIFIC TIMING OF SOMATIC TRANSCRIPTIONAL RESPONSE TO GERMLINE ABLATION AND IRRADIATION EXTENDS TO EACH TISSUE.....	344
FIGURE S 3: CELLULAR ORIGIN DEFINES TRANSCRIPTIONAL PROFILES OF BOTH GENES AND TES; AND RESNIK SIMILARITY REVEALS CONSERVED GENOMIC STABILITY FUNCTIONS AND DIVERGENT CHROMATIN-ASSOCIATED PROCESSES ACROSS STRAINS AND TIMEPOINTS POST IRRADIATION.	338
FIGURE S 4 : COMPARATIVE FUNCTIONAL ANALYSIS OF SOMATIC GENE EXPRESSION REVEALS SPECIALISED METABOLIC AND IMMUNE PROGRAMMES IN ALL GLF SOMATIC TISSUES AND TIMEPOINT DEPENDING IMMUNE EXPRESSION.	340
FIGURE S 5: GLF SOMA SHOWS A TRANSIENT SHIFT IN HEDGEHOG PATHWAY REGULATORS (HH) AND DREAM COMPLEX TARGET GENES AT 3HPiR, DRIVEN BY SPLEEN FOR HH REGULATORS AND ALL SOMA FOR DREAM GENES.....	342
FIGURE S 6: MAPPING OF DE GENES AND TES REVEALS CHROMOSOME 4 AS A HOTSPOT OF ENRICHMENT IN GLF FISH IN RESPONSE TO IRRADIATION.	348
FIGURE S 7: STRAIN-SPECIFIC DIVERGENCE IN TRANSPOSABLE ELEMENT COPY NUMBER AND EXPRESSION FOLLOWING IRRADIATION.....	346

FIGURE S 8: PERFORMANCE SCORES (RESCALED R^2 + INVERSE MAE) BY CpG GROUPING METHOD, FACETED BY TISSUE, SHOWING TILE DOMINANCE IN BRAIN/INTESTINE BUT REVERSAL IN MULTI-TISSUE CLOCKS.	362
FIGURE S 9: EPIGENETIC CLOCK PERFORMANCE ACROSS TISSUE × CpG AGGREGATION STRATEGIES.....	381
FIGURE S 10 : NO TISSUE-SPECIFIC MODULATION OF CpG METHYLATION BY SEX AND AGE.	382
FIGURE S 11: COMPREHENSIVE ANALYSES OF SHARED SOMATIC CpGs.....	383
FIGURE S 12: EPIGENETIC CLOCK PERFORMANCE BY CpG GROUPING.	394
FIGURE S 13 : RAPAMYCIN TREATMENT DOES NOT SIGNIFICANTLY AFFECT WITHIN-SAMPLE VARIABILITY, REPRODUCTIVE OUTPUT, OR PREDICTED GROWTH IN THE LIFELONG EXPERIMENT.....	414

TABLE OF SUPPLEMENTARY TABLES

TABLE S 1: GENERALISED LINEAR MODELLING OF GENE EXPRESSION FOR A TARGETED COMPARISON.....	292
TABLE S 2: LINEAR CONTRASTS DERIVED FROM THE FULL MODEL FOR SYSTEMATIC EVALUATION OF GENE EXPRESSION EFFECTS.	296
TABLE S 3: NUMBER OF DIFFERENTIALLY EXPRESSED GENES IDENTIFIED BY EACH GENERALISED LINEAR MODEL.....	294
TABLE S 4: MIXED-EFFECTS MODELLING OF SOMATIC GENE FEATURES.....	319
TABLE S 5: PAIRWISE COMPARISONS OF ESTIMATED MARGINAL MEANS FROM MIXED-EFFECTS MODELS.....	322
TABLE S 6: MIXED-EFFECTS MODELLING OF GERMLINE GENE FEATURES.....	328
TABLE S 7 : TOP 5 DIFFERENTIALLY EXPRESSED GENES IN THE SOMA ACROSS IRRADIATION TIMEPOINTS IN ZEBRAFISH STRAINS.....	299
TABLE S 8: TOP 5 DIFFERENTIALLY EXPRESSED GENES IN EACH SOMATIC TISSUE ACROSS IRRADIATION TIMEPOINTS IN ZEBRAFISH STRAINS.	303
TABLE S 9: GENE INDEX FOR THE ALLUVIAL PLOT SHOWN IN FIGURE 4 OF CHAPTER 2 .	330
TABLE S 10: NUMBER OF SAMPLES PER TISSUE AND SEX USED IN THE EXPERIMENTAL STUDY	349
TABLE S 11: DISTRIBUTION OF SAMPLES ACROSS TECHNICAL AND BIOLOGICAL FACTORS	350
TABLE S 12 : DISTRIBUTION OF CROSS-TISSUE STABILITY CLASSES.....	360
TABLE S 13: LINEAR MIXED-EFFECTS MODELLING OF EPIGENETIC CLOCK PERFORMANCE ACROSS CPG AGGREGATION STRATEGIES.....	357

TABLE S 14: TOP-SELECTED CpG SITES IDENTIFIED IN SINGLE-TISSUE AND MULTI-TISSUE CLOCKS, AND THEIR OVERLAP ACROSS SOMATIC TISSUES.....	352
TABLE S 15 : EFFECT OF CpG SET SIZE ON EPIGENETIC CLOCK PERFORMANCE ASSESSED BY MIXED-EFFECTS MODELLING.....	359
TABLE S 16: NUMBER OF SAMPLES PER TISSUE AND SEX USED IN EXPERIMENTAL STUDY	363
TABLE S 17: DISTRIBUTION OF SAMPLES ACROSS TECHNICAL AND BIOLOGICAL FACTORS	365
TABLE S 18: LINEAR MIXED-EFFECTS MODELLING OF TISSUE- AND AGGREGATION-DEPENDENT AGEING METRICS.....	367
TABLE S 19: LINEAR MIXED-EFFECTS MODELLING OF TISSUE-SPECIFIC CpG METHYLATION FEATURES.....	375
TABLE S 20: LINEAR MIXED-EFFECTS MODELLING OF CpG SITES SHARED ACROSS SOMATIC TISSUES.	377
TABLE S 21: MULTIVARIATE TEST ASSESSING TISSUE SEPARATION ACROSS CHRONOLOGICAL AGE BASED ON CpG SITES SHARED ACROSS SOMATIC TISSUES.	380
TABLE S 22: DISTRIBUTION OF SAMPLES ACROSS TECHNICAL AND BIOLOGICAL FACTORS	384
TABLE S 23: MIXED-EFFECTS AND MULTIVARIATE ANALYSES OF GERMLINE-DEPENDENT EPIGENETIC AGEING METRICS.....	386
TABLE S 24: LINEAR MODELLING OF BEHAVIOURAL METRICS.....	390
TABLE S 25: AGED EXPERIMENT - FISHERS TEST ON SEX-SPECIFIC BREEDING OUTCOMES BY TREATMENT 2 MONTHS AFTER START OF RAPAMYCIN TREATMENT.....	395
TABLE S 26: MIXED-EFFECTS MODELLING OF BEHAVIOURAL METRICS ACROSS RAPAMYCIN TREATMENT CONDITIONS IN AGED FISH.....	398

TABLE S 27 : MIXED-EFFECTS AND NON-LINEAR MODELLING OF EPIGENETIC AGEING METRICS.	401
TABLE S 28 : GENERALISED LINEAR MIXED-EFFECTS MODELLING OF OVARIAN LESION INCIDENCE.....	405
TABLE S 29 : LINEAR EFFECT MODEL (LMER IN R) FOR BEHAVIOUR METRICS IN THE LIFELONG EXPERIMENT.....	408
TABLE S 30 : LINEAR MIXED-EFFECTS MODELLING OF VERTICAL EXPLORATION BEHAVIOUR IN THE NOVEL TANK DIVING TEST.....	411

8. Appendix Chapter two

Supporting Information for Germline ablation affects somatic gene expression resulting in improved somatic maintenance

8.1. Material and Methods

8.1.1. Fish husbandry details

Wild-type AB zebrafish (*Danio rerio*) were maintained under standard conditions at 28 °C in 3 L tanks with a 12:12 h light–dark cycle, following established zebrafish facility protocols. Outbreeding was strictly maintained to preserve genetic variation between generations. Juvenile fish were fed five times per day, whereas adult fish were fed three times daily with live *Artemia* and dried flake food (Zeigler adult zebrafish diet; Aquatic Habitats). Water quality, temperature, and pH were continuously monitored and maintained within standard ranges (pH 7.0–7.5, conductivity 500–550 µS). Fish were reared and handled in accordance with ethical guidelines at the Evolutionary Biology Centre, Uppsala University.

8.1.2. CRISPR/Cas9-mediated germline ablation in zebrafish

To generate germline-free zebrafish, dead-end (*dnd1*) knockout was performed using CRISPR/Cas9-mediated mutagenesis. *dnd1* is essential for primordial germ cell survival, and its loss leads to complete germ cell ablation, resulting in sterile males with somatic testes.

8.1.3. sgRNA design and template preparation

Target sequences within *dnd1* were designed using CRISPR-Scan (<https://crisprscan.org>) and verified for minimal off-target potential. For in vitro transcription using the T7 promoter, each 20 nt target sequence was prefixed with the T7 initiation site ("GG") and appended to a 20 nt overlap sequence complementary to the generic sgRNA scaffold. The targeting oligo (oligo A) was annealed to a universal sgRNA backbone oligo (oligo B; 5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTA ACTTGCTATTTCTAGCTCTAAAAC-3') as described by Varshney et al. (2015).

Annealing and fill-in reactions were carried out using Phusion High-Fidelity DNA Polymerase (NEB M0530S) under the following conditions: 98 °C for 2 min, 50 °C for 10 min, and 72 °C for 10 min. The expected 120 bp sgRNA template product was confirmed by electrophoresis on a 2.5% agarose gel.

8.1.4. In vitro transcription of sgRNA

Approximately 2–3 µL of the assembled sgRNA template was used as input for RNA synthesis using the T7 Quick High Yield RNA Synthesis Kit (NEB E2050S), following the manufacturer's instructions. Following transcription, reactions were treated with 1 µL DNase I (from kit) at 37 °C for 15 min and RNA was purified by isopropanol/sodium acetate precipitation or alternatively using the mirVana miRNA Isolation Kit (Life Technologies AM1560). RNA pellets were dissolved in RNase-free water, quantified using a NanoDrop spectrophotometer (expected yield 1.5–3 µg/µL), and quality-checked on 2–2.5% agarose gels. Aliquots were stored at –80 °C until use.

Appendix Chapter two

8.1.5. Preparation of Cas9 mRNA

Cas9 mRNA was synthesised from the pT3TS-nCas9n plasmid (Addgene #46757). The plasmid was linearised with XbaI (NEB), purified using a QIAprep Spin Column (Qiagen 27104), and quantified by NanoDrop. In vitro transcription was carried out using the mMESSAGE mMACHINE T3 Kit (Life Technologies AM1348). The resulting RNA was purified either by LiCl precipitation or using the RNeasy Mini Kit (Qiagen, 74104). RNA integrity was verified on a 1% agarose gel (expected single 4.5 kb band), and concentration determined ($>1 \mu\text{g}/\mu\text{L}$). Cas9 mRNA was aliquoted and stored at -80°C .

8.1.6. Microinjection

Injection mixtures were prepared immediately prior to use on ice. A standard 10 μL mix contained:

- 1 μL Cas9 mRNA (750 ng/ μL),
- 1 μL sgRNA per target (250 ng/ μL),
- 1 μL phenol red (Sigma P0290),
- RNase-free water to 10 μL total volume.

Injection needles were calibrated to deliver 2 nL per embryo, corresponding to approximately 150 pg Cas9 mRNA and 25 pg sgRNA per target. Injections were performed at the 1-cell stage, typically into the yolk for higher embryo survival, or into the cell for greater mutagenesis efficiency. Embryos were maintained at 28°C in E3 medium and inspected daily.

8.1.7. Verification of mutagenesis and germline loss

At 2–3 days post fertilisation, subsets of embryos were screened for editing efficiency using the T7 endonuclease I (T7E1) assay (NEB M0302L) or

fragment length analysis. Efficient sgRNAs were identified by cleavage of heteroduplexed PCR amplicons. Successfully edited fish were raised to adulthood and confirmed to exhibit complete germ cell ablation via morphological and genotypic analysis.

8.1.8. Gamma irradiation procedure

Twenty-four adult males (GLC and GLF, > 6 months old) were irradiated with a 20 Gy dose from a Caesium-137 gamma source (Gammacell 3000; Stockholm University). Fish were transferred to sterile six-well plates (16.8 mL per well, 1–2 fish/well) containing system water. Wells were sealed with laboratory film to prevent desiccation. Plates were placed in the irradiator for approximately 25 minutes to reach the target dose. Dosimetry was confirmed using a calibrated ionisation chamber (ISO/IEC 17025 standard). Post-irradiation, fish were returned to standard housing conditions and sampled at 3 h and 24 h post-irradiation.

8.1.9. RNA extraction and library preparation

Tissues were manually homogenised in Trizol Reagent (Invitrogen), and RNA was extracted following the manufacturer's protocol. Phase separation was performed with chloroform, followed by centrifugation at $12,000 \times g$ for 20 min at 4 °C. The aqueous phase was recovered, and residual DNA was removed using the TURBO DNA-free Kit (Ambion). RNA concentration and purity were measured on a NanoDrop spectrophotometer, and integrity was verified using an Agilent Bioanalyzer.

RNA libraries were prepared using NEBNext Ultra II Directional RNA Library Prep Kit (poly-A selection, 200–300 bp inserts) and sequenced by Novogene (UK). Samples were sequenced in two balanced batches (intestine, kidney,

and testes in batch 1; brain, spleen, and caudal fin in batch 2) on an Illumina NovaSeq 6000 S1 v1.5 platform, 150 bp paired-end, 60-plex per lane.

8.1.10. RNA-seq processing pipeline

Raw reads were processed using NF-core RNA-seq v3.9 under Nextflow v22.04.5. Quality trimming and adapter removal were performed using Trim Galore, retaining reads $\geq 10,000$ per sample. Reads were aligned to the *Danio rerio* reference genome (GRCz11) using STAR (two-pass mode) and quantified with Salmon in selective-alignment mode. Gene-level count matrices were generated using the NF-core standard parameters. Quality metrics (FastQC, MultiQC) were verified for all samples.

8.1.11. Differential gene expression analyses

Gene-level counts were analysed in edgeR v3.38 (Bioconductor) within RStudio. Lowly expressed genes were filtered using CPM > 1 in ≥ 3 samples. Normalisation was performed using TMM (trimmed mean of M values). Differential expression was tested using generalised linear models with contrasts representing: (i) GLC vs GLF (non-irradiated); (ii) Irradiated vs control (GLC and GLF separately, 3 h and 24 h); (iii) Irradiated GLC vs GLF (time-matched).

False discovery rate (FDR) was controlled using Benjamini–Hochberg correction (FDR < 0.05). Volcano plots were generated using EnhancedVolcano v1.18.

8.1.12. Transposable element (TE) annotation and analysis

TE expression was quantified with Tetrascripts v2.2.3, using the *Danio rerio* genome (GRCz11) and Ensembl TE annotations. TE differential expression was analysed analogously to gene-level data. Because libraries were poly(A)-selected, detected TEs represent only polyadenylated fractions of transposon-derived transcripts.

TE age and divergence were inferred using RepeatMasker (RM) and calcDivergenceFromAlign.pl, which calculates Kimura substitution levels (KS) between TE copies and their consensus sequences. Lower KS values indicate more recent transposon activity. TE density was calculated as the proportion of TE-occupied base pairs per gene, normalised by total gene length.

8.1.13. Chromatin and gene structure analyses

To assess chromatin openness, gene coordinates were mapped to ATAC-seq chromatin accessibility tracks derived from the zebrafish COPES dataset described by Baranasic et al., Nat. Genet. 54:1682–1693 (2022). These tracks identify Constitutively Open Predicted Elements (COPES), representing genomic regions consistently accessible throughout zebrafish development.

Gene exon–intron structures were obtained from Ensembl v107 (*Danio rerio* GRCz11) annotations. Exonic regions were merged by chromosomal start sites to avoid artificial elongation from alternative splicing. Total exonic length per gene was computed as the sum of non-overlapping regions; intronic length was derived by subtracting exonic from total gene length. The proportion of exonic content was expressed as (exonic length / gene length) × 100.

8.2. Supplementary Tables

Table S 1: Generalised linear modelling of gene expression for a targeted comparison.

Two generalised linear models (GLMs; edgeR) were fitted to a biologically defined subset of samples to isolate the specified comparison and effect. Model dispersion (global overdispersion across genes), mean tagwise dispersion (average gene-level variability), and residual degrees of freedom (reflecting effective model power) are reported.

Model	EdgeR model	Comparison of interest	Number of samples / Total	Model Dispersion	Mean Tagwise Dispersion	Residual df
Full Soma	~ Strain * Irradiation * Tissue * Timepoint + Individual	Overall difference in expression between somas considering the effect of genotoxic stress	111 / 123	2.4	3.7	55

Appendix Chapter two

Germlin e	~ Irradiatio n * Timepoi nt + Individua l	Effect of irradiation on expression in the germline	12/123	0.6	1.2	8
--------------	---	--	--------	-----	-----	---

Table S 2: Number of differentially expressed genes identified by each generalised linear model.

Differential expression was assessed using generalised linear models implemented in edgeR. Genes were classified as differentially expressed at a false discovery rate (FDR) < 0.05. Upregulated genes correspond to positive $\log_2$ fold change ($\log\text{FC} > 0$), while downregulated genes correspond to negative $\log_2$ fold change ($\log\text{FC} < 0$)

Comparison	Number of upregulated genes	Number of downregulated genes	Number of non-differentially expressed genes	Total number of expressed genes
C1	2357	4498	16659	23514
C2.24	3372	4619	15523	23514
C2.3	3779	4473	15262	23514
C3.br	1592	3891	18031	23514
C3.int	3211	6476	13827	23514
C3.kid	3405	6823	13286	23514
C3.ski	5990	3766	13758	23514
C3.spl	3850	5692	13972	23514
C4.br.3	4498	4110	14906	23514
C4.int.3	5188	4688	13638	23514

Appendix Chapter two

C4.kid.3	4997	4759	13758	23514
C4.ski.3	4806	4654	14054	23514
C4.spl.3	4819	6876	11819	23514
C5.br.24	1884	3050	18580	23514
C5.int.24	4465	7002	12047	23514
C5.kid.24	6117	5270	12127	23514
C5.ski.24	5629	5516	12369	23514
C5.spl.24	4122	4024	15368	23514
C6.24	6989	7736	8167	22892
C6.3	6710	9118	7064	22892

Table S 3: Linear contrasts derived from the full model for systematic evaluation of gene expression effects.

A comprehensive set of linear contrasts was specified from the full model to test predefined hypotheses, including strain-specific effects within individual tissues, irradiation responses within each strain, and strain-treatment interactions.

Model	Contrast ID	Comparison	Interpretation
Soma model	C1	Non-irrad. GLC vs Non-irrad. GLF	Effect of germline ablation on non-irradiated somatic tissues
Soma model	C2.3	3hpir GLC vs 3hpir GL	Germline-irradiation interaction at 3 hours post-irradiation (3hpir) in soma
Soma model	C2.24	24hpir GLC vs 24hpir GLF	Germline-irradiation interaction at 24 hours post-irradiation (24hpir) in soma
Soma model	C3.brain	Non-irrad. GLC brain vs Non-irrad. GLF brain	Effect of germline ablation in brain tissue
Soma model	C3.intestine	Non-irrad. GLC intestine vs Non-irrad. GLF intestine	Effect of germline ablation in intestinal tissue

Appendix Chapter two

Soma model	C3.kidney	Non-irrad. GLC kidney vs Non-irrad. GLF kidney	Effect of germline ablation in kidney tissue
Soma model	C3.caudal fin	Non-irrad. GLC caudal fin vs Non-irrad. GLF caudal fin	Effect of germline ablation in caudal fin tissue
Soma model	C3.spleen	Non-irrad. GLC spleen vs Non-irrad. GLF spleen	Effect of germline ablation in spleen tissue
Soma model	C4.brain.3	3hpir GLC brain vs 3hpir GLF brain	Germline–irradiation interaction in brain tissue at 3hpir
Soma model	C4.intestine.3	3hpir GLC intestine vs 3hpir GLF intestine	Germline–irradiation interaction in intestinal tissue at 3hpir
Soma model	C4.kidney.3	3hpir GLC kidney vs 3hpir GLF kidney	Germline–irradiation interaction in kidney tissue at 3hpir
Soma model	C4.caudal fin.3	3hpir GLC caudal fin vs 3hpir GLF caudal fin	Germline–irradiation interaction in caudal fin tissue at 3hpir
Soma model	C4.spleen.3	3hpir GLC spleen vs 3hpir GLF spleen	Germline ablation–irradiation interaction in spleen tissue at 3hpir

Appendix Chapter two

Soma model	C5.brain.24	24hpir GLC brain vs 24hpir GLF brain	Germline–irradiation interaction in brain tissue at 24hpir
Soma model	C5.intestine.24	24hpir GLC intestine vs 24hpir GLF intestine	Germline–irradiation interaction in intestinal tissue at 24hpir
Soma model	C5.kidney.24	24hpir GLC kidney vs 124hpir GLF kidney	Germline–irradiation interaction in kidney tissue at 24hpir
Soma model	C5.caudal fin. 24	24hpir GLC caudal fin vs 24hpir GLF caudal fin	Germline–irradiation interaction in caudal fin tissue at 24hpir
Soma model	C5.spleen.24	24hpir GLC spleen vs 24hpir GLF spleen	Germline ablation–irradiation interaction in spleen tissue at 24hpir
Germline model	C6.3	Non-irrad. GLC testes vs 3hpir GLC testes	Irradiation at 24 hours post-irradiation in the germline (24hpir)
Germline model	C6.24	Non-irrad. GLC testes vs 24hpir GLC testes	Irradiation at 3 hours post-irradiation in the germline (3hpir)

Table S 4 : Top 5 differentially expressed genes in the soma across irradiation timepoints in zebrafish strains.

This table presents the top ten differentially expressed genes (ranked by false discovery rate, FDR) for each zebrafish strain (GLC vs GLF) at three timepoints post-irradiation: **A)** non-irradiated controls (nonir), **B)** 3 hours post-irradiation (3hpir), and **C)** 24 hours post-irradiation (24hpir). Genes were selected from a linear model-based differential expression analysis using an FDR threshold of <0.05, stratified by strain and direction of regulation (upregulated or downregulated). For each gene, Ensembl gene identifiers, gene symbols, associated model groupings, and relevant Gene Ontology (GO) terms, including GO ID, term name, and definition, are provided.

A)

Ensemble Gene ID	Symbol	Model	Label	Strain	GO ID
ENSDARG000000079105	mhc2db	full_model	nonir_soma	GLC	GO:0042613
ENSDARG000000097137	NA	full_model	nonir_soma	GLC	
ENSDARG000000071024	zgc:171679	full_model	nonir_soma	GLC	GO:0005515
ENSDARG000000030448	sppl2	full_model	nonir_soma	GLC	GO:0042500
ENSDARG000000044074	lox12b	full_model	nonir_soma	GLC	GO:0016641

Appendix Chapter two

ENSDARG00000086418	si:ch211-236p5.3	full_model	nonir_soma	GLF	GO:0005515
ENSDARG00000087299	armc9	full_model	nonir_soma	GLF	GO:0036064
ENSDARG00000112702	NA	full_model	nonir_soma	GLF	
ENSDARG00000068621	nirc11	full_model	nonir_soma	GLF	GO:0005515
ENSDARG00000104899	si:dkey-242k1.4	full_model	nonir_soma	GLF	

B)

Ensemble Gene ID	Symbol	Model	Label	Strain	GO ID
ENSDARG00000071024	zgc:171679	full_model	3hpir_soma	GLC	GO:0005515
ENSDARG00000003132	apip	full_model	3hpir_soma	GLC	
ENSDARG000000097137	NA	full_model	3hpir_soma	GLC	
ENSDARG000000098030	NA	full_model	3hpir_soma	GLC	

Appendix Chapter two

ENSDARG00000103 841	NA	full_mo del	3hpir_so ma	GLC	GO:00038 24
ENSDARG00000006 8621	nlrc1l	full_mo del	3hpir_so ma	GLF	GO:00055 15
ENSDARG00000010 4635	mhc2dgb	full_mo del	3hpir_so ma	GLF	
ENSDARG00000009 1847	LOC110437 747	full_mo del	3hpir_so ma	GLF	GO:00055 15
ENSDARG000000116 524	mhc2dhb	full_mo del	3hpir_so ma	GLF	GO:00426 13
ENSDARG00000007 9645	NA	full_mo del	3hpir_so ma	GLF	

c)

Ensemble Gene ID	Symbol	Model	Label	Strain	GO ID
ENSDARG00000007 9105	mhc2dab	full_mo del	24hpir_so ma	GLC	GO:00426 13
ENSDARG00000010 4635	mhc2dgb	full_mo del	24hpir_so ma	GLC	
ENSDARG00000007 1024	zgc:171679	full_mo del	24hpir_so ma	GLC	GO:00055 15

Appendix Chapter two

ENSDARG00000004 4074	lox12b	full_mo del	24hpir_so ma	GLC	GO:00166 41
ENSDARG00000007 4653	si:ch211- 233m11.1	full_mo del	24hpir_so ma	GLC	GO:00055 15
ENSDARG00000006 8621	n1rc11	full_mo del	24hpir_so ma	GLF	GO:00055 15
ENSDARG00000009 1847	LOC110437 747	full_mo del	24hpir_so ma	GLF	GO:00055 15
ENSDARG00000009 4210	fth131	full_mo del	24hpir_so ma	GLF	GO:00068 79
ENSDARG00000008 7299	armc9	full_mo del	24hpir_so ma	GLF	GO:00360 64
ENSDARG00000010 4613	NA	full_mo del	24hpir_so ma	GLF	

Table S 5: Top 5 differentially expressed genes in each somatic tissue across irradiation timepoints in zebrafish strains.

This table presents the top ten differentially expressed genes (ranked by false discovery rate, FDR) for each zebrafish strain (GLC vs GLF) at three timepoints post-irradiation: **A)** non-irradiated controls (nonir), **B)** 3 hours post-irradiation (3hpir), and **C)** 24 hours post-irradiation (24hpir). Genes were selected from a linear model-based differential expression analysis using an FDR threshold of <0.05, stratified by strain and direction of regulation (upregulated or downregulated). For each gene, Ensembl gene identifiers, gene symbols, associated model groupings, and relevant Gene Ontology (GO) terms, including GO ID, term name, and definition, are provided.

A)

Ensemble Gene ID	Symbol	Model	Label	Strain	GO ID
ENSDARG00000079105	mhc2dab	full_model	nonir_brain	GLC	GO:0042613
ENSDARG00000044074	lox12b	full_model	nonir_brain	GLC	GO:0016641
ENSDARG00000097137	NA	full_model	nonir_brain	GLC	
ENSDARG000000113971	LOC100149352	full_model	nonir_brain	GLC	GO:0005634
ENSDARG00000004141	dhrs11b.2	full_model	nonir_brain	GLC	GO:0005575

Appendix Chapter two

ENSDARG000000 87299	armc9	full_mo del	nonir_brain	GLF	GO:0036 064
ENSDARG0000011 2702	NA	full_mo del	nonir_brain	GLF	
ENSDARG000000 86418	si:ch211- 236p5.3	full_mo del	nonir_brain	GLF	GO:0005 515
ENSDARG000000 27992	hao2	full_mo del	nonir_brain	GLF	GO:00164 91
ENSDARG000000 95949	si:dkey- 22i16.9	full_mo del	nonir_brain	GLF	
ENSDARG000000 79105	mhc2dab	full_mo del	nonir_intes tine	GLC	GO:0042 613
ENSDARG0000010 4387	slc4a5b	full_mo del	nonir_intes tine	GLC	GO:0008 509
ENSDARG000000 97137	NA	full_mo del	nonir_intes tine	GLC	
ENSDARG000000 87012	NA	full_mo del	nonir_intes tine	GLC	GO:0005 525
ENSDARG000000 36045	penkb	full_mo del	nonir_intes tine	GLC	GO:0007 218
ENSDARG0000011 2702	NA	full_mo del	nonir_intes tine	GLF	

Appendix Chapter two

ENSDARG000000 76043	si:dkeyp- 73d8.9	full_mo del	nonir_intes tine	GLF	
ENSDARG00000010 5731	LOC103908 715	full_mo del	nonir_intes tine	GLF	GO:0005 515
ENSDARG000000 86418	si:ch21l- 236p5.3	full_mo del	nonir_intes tine	GLF	GO:0005 515
ENSDARG000000 78844	LOC100007 883	full_mo del	nonir_intes tine	GLF	
ENSDARG000000 79105	mhc2dab	full_mo del	nonir_kidn ey	GLC	GO:0042 613
ENSDARG000000 38587	LOC100004 199	full_mo del	nonir_kidn ey	GLC	
ENSDARG000000 78828	npb	full_mo del	nonir_kidn ey	GLC	GO:00016 64
ENSDARG000000 55240	xdh	full_mo del	nonir_kidn ey	GLC	GO:00435 46
ENSDARG000000 97137	NA	full_mo del	nonir_kidn ey	GLC	
ENSDARG00000011 2702	NA	full_mo del	nonir_kidn ey	GLF	
ENSDARG00000010 0851	gphnb	full_mo del	nonir_kidn ey	GLF	GO:00323 24

Appendix Chapter two

ENSDARG000000 86418	si:ch21l- 236p5.3	full_mo del	nonir_kidn ey	GLF	GO:0005 515
ENSDARG000000 86569	zgc:172051	full_mo del	nonir_kidn ey	GLF	
ENSDARG000000 35694	stm	full_mo del	nonir_kidn ey	GLF	
ENSDARG000000 79105	mhc2dab	full_mo del	nonir_caud al fin	GLC	GO:0042 613
ENSDARG000000 97137	NA	full_mo del	nonir_caud al fin	GLC	
ENSDARG000000 88625	si:ch21l- 142d6.2	full_mo del	nonir_caud al fin	GLC	GO:0005 515
ENSDARG0000011 3971	LOC100149 352	full_mo del	nonir_caud al fin	GLC	GO:0005 634
ENSDARG000000 30448	sppl2	full_mo del	nonir_caud al fin	GLC	GO:0042 500
ENSDARG000000 86418	si:ch21l- 236p5.3	full_mo del	nonir_caud al fin	GLF	GO:0005 515
ENSDARG000000 87299	armc9	full_mo del	nonir_caud al fin	GLF	GO:0036 064
ENSDARG0000011 6832	NA	full_mo del	nonir_caud al fin	GLF	GO:00150 74

Appendix Chapter two

ENSDARG000000 43002	vmola	full_mo del	nonir_caud al fin	GLF	
ENSDARG000000 37789	pvalb1	full_mo del	nonir_caud al fin	GLF	GO:0005 509
ENSDARG000000 79105	mhc2dab	full_mo del	nonir_splee n	GLC	GO:0042 613
ENSDARG000000 38587	LOC100004 199	full_mo del	nonir_splee n	GLC	
ENSDARG000000 30448	sppl2	full_mo del	nonir_splee n	GLC	GO:0042 500
ENSDARG000000 31952	mb	full_mo del	nonir_splee n	GLC	GO:00156 71
ENSDARG000000 97137	NA	full_mo del	nonir_splee n	GLC	
ENSDARG000000 20084	tg	full_mo del	nonir_splee n	GLF	
ENSDARG000000 86418	si:ch211- 236p5.3	full_mo del	nonir_splee n	GLF	GO:0005 515
ENSDARG000000 28148	pax2a	full_mo del	nonir_splee n	GLF	GO:0006 355
ENSDARG000000 05392	slc5a5	full_mo del	nonir_splee n	GLF	GO:0055 085

Appendix Chapter two

ENSDARG00000069296	moxd1l	full_model	nonir_spleen	GLF	GO:0004500
--------------------	--------	------------	--------------	-----	------------

B)

Ensemble Gene ID	Symbol	Model	Label	Strain	GO ID
ENSDARG00000014577	LOC108183819	full_model	3hpir_brain	GLC	GO:0005515
ENSDARG00000044074	lox12b	full_model	3hpir_brain	GLC	GO:0016641
ENSDARG00000003132	apip	full_model	3hpir_brain	GLC	
ENSDARG00000098030	NA	full_model	3hpir_brain	GLC	
ENSDARG00000097137	NA	full_model	3hpir_brain	GLC	
ENSDARG00000068621	nlrcl1	full_model	3hpir_brain	GLF	GO:0005515
ENSDARG000000104635	mhc2dgb	full_model	3hpir_brain	GLF	
ENSDARG00000097610	NA	full_model	3hpir_brain	GLF	

Appendix Chapter two

ENSDARG0000011 6524	mhc2dhb	full_mo del	3hpir_brain	GLF	GO:0042 613
ENSDARG000000 91847	LOC110437 747	full_mo del	3hpir_brain	GLF	GO:00055 15
ENSDARG000000 88524	acot17	full_mo del	3hpir_intes tine	GLC	GO:00167 90
ENSDARG0000010 4387	slc4a5b	full_mo del	3hpir_intes tine	GLC	GO:0008 509
ENSDARG000000 77068	si:ch211- 11p18.6	full_mo del	3hpir_intes tine	GLC	GO:00055 15
ENSDARG0000001 6623	si:ch211- 195b13.1	full_mo del	3hpir_intes tine	GLC	
ENSDARG000000 87012	NA	full_mo del	3hpir_intes tine	GLC	GO:00055 25
ENSDARG0000010 4635	mhc2dgb	full_mo del	3hpir_intes tine	GLF	
ENSDARG000000 71216	si:ch211- 133n4.9	full_mo del	3hpir_intes tine	GLF	GO:0003 676
ENSDARG000000 92833	si:dkeyp- 1h4.8	full_mo del	3hpir_intes tine	GLF	
ENSDARG000000 79105	mhc2dab	full_mo del	3hpir_intes tine	GLF	GO:0042 613

Appendix Chapter two

ENSDARG000000 94210	fthl3l	full_mo del	3hpir_intes tine	GLF	GO:0006 879
ENSDARG000000 71223	zgc:158445	full_mo del	3hpir_kidn ey	GLC	GO:00055 75
ENSDARG000000 97137	NA	full_mo del	3hpir_kidn ey	GLC	
ENSDARG000000 03132	apip	full_mo del	3hpir_kidn ey	GLC	
ENSDARG0000003 6846	anks4b	full_mo del	3hpir_kidn ey	GLC	GO:00055 15
ENSDARG000000 78828	npb	full_mo del	3hpir_kidn ey	GLC	GO:00016 64
ENSDARG000000 91847	LOC110437 747	full_mo del	3hpir_kidn ey	GLF	GO:00055 15
ENSDARG00000010 4635	mhc2dgb	full_mo del	3hpir_kidn ey	GLF	
ENSDARG000000 68621	nlrc1l	full_mo del	3hpir_kidn ey	GLF	GO:00055 15
ENSDARG000000 78529	adgrb1b	full_mo del	3hpir_kidn ey	GLF	
ENSDARG000000 79645	NA	full_mo del	3hpir_kidn ey	GLF	

Appendix Chapter two

ENSDARG000000 42988	slc24a2	full_mo del	3hpir_caud al fin	GLC	
ENSDARG000000 98030	NA	full_mo del	3hpir_caud al fin	GLC	
ENSDARG000000 97480	NA	full_mo del	3hpir_caud al fin	GLC	
ENSDARG000000 25783	si:ch211- 125e6.11	full_mo del	3hpir_caud al fin	GLC	
ENSDARG000000 74754	ifi27.6	full_mo del	3hpir_caud al fin	GLC	GO:00160 20
ENSDARG000000 68621	nlrc11	full_mo del	3hpir_caud al fin	GLF	GO:00055 15
ENSDARG0000010 4635	mhc2dgb	full_mo del	3hpir_caud al fin	GLF	
ENSDARG0000010 3716	mhc2dga	full_mo del	3hpir_caud al fin	GLF	GO:0042 613
ENSDARG000000 24829	tnn	full_mo del	3hpir_caud al fin	GLF	GO:00055 15
ENSDARG0000011 6524	mhc2dhb	full_mo del	3hpir_caud al fin	GLF	GO:0042 613
ENSDARG0000001 6494	ddc	full_mo del	3hpir_splee n	GLC	

Appendix Chapter two

ENSDARG000000 71601	pvalb9	full_mo del	3hpir_splee n	GLC	GO:00055 09
ENSDARG000000 68372	zgc:158427	full_mo del	3hpir_splee n	GLC	GO:00160 20
ENSDARG000000 57789	lyz	full_mo del	3hpir_splee n	GLC	GO:00037 96
ENSDARG000000 88524	acot17	full_mo del	3hpir_splee n	GLC	GO:00167 90
ENSDARG000000 20084	tg	full_mo del	3hpir_splee n	GLF	
ENSDARG00000010 4635	mhc2dgb	full_mo del	3hpir_splee n	GLF	
ENSDARG000000 68621	nlrc11	full_mo del	3hpir_splee n	GLF	GO:00055 15
ENSDARG00000010 3716	mhc2dga	full_mo del	3hpir_splee n	GLF	GO:0042 613
ENSDARG00000011 6524	mhc2dhb	full_mo del	3hpir_splee n	GLF	GO:0042 613
ENSDARG000000 78683	LOC110437 942	germlin e	3hpir_teste s	GLC	GO:00165 67
ENSDARG00000011 5978	NA	germlin e	3hpir_teste s	GLC	

Appendix Chapter two

ENSDARG00000075352	brinp3b	germline	3hpir_testes	GLC	GO:0005737
ENSDARG000000032631	ltb4r	germline	3hpir_testes	GLC	
ENSDARG00000069282	bbc3	germline	3hpir_testes	GLC	GO:0090200

C)

Ensemble Gene ID	Symbol	Model	Label	Strain	GO ID
ENSDARG00000079105	mhc2dab	full_model	24hpir_brain	GLC	GO:0042613
ENSDARG00000044074	lox12b	full_model	24hpir_brain	GLC	GO:0016641
ENSDARG000000104635	mhc2dgb	full_model	24hpir_brain	GLC	
ENSDARG00000037954	tnnt1	full_model	24hpir_brain	GLC	GO:0005861
ENSDARG00000091280	si:ch21l-66k16.27	full_model	24hpir_brain	GLC	GO:0042981
ENSDARG00000068621	nirc1l	full_model	24hpir_brain	GLF	GO:0005515

Appendix Chapter two

ENSDARG000000 91847	LOC110437 747	full_mo del	24hpir_brai n	GLF	GO:0005 515
ENSDARG000000 76043	si:dkeyp- 73d8.9	full_mo del	24hpir_brai n	GLF	
ENSDARG000000 87299	armc9	full_mo del	24hpir_brai n	GLF	GO:0036 064
ENSDARG000000 86418	si:ch21l- 236p5.3	full_mo del	24hpir_brai n	GLF	GO:0005 515
ENSDARG0000001 04387	slc4a5b	full_mo del	24hpir_intes tine	GLC	GO:0008 509
ENSDARG000000 79105	mhc2dab	full_mo del	24hpir_intes tine	GLC	GO:0042 613
ENSDARG0000001 04635	mhc2dgb	full_mo del	24hpir_intes tine	GLC	
ENSDARG000000 36045	penkb	full_mo del	24hpir_intes tine	GLC	GO:0007 218
ENSDARG0000001 03716	mhc2dga	full_mo del	24hpir_intes tine	GLC	GO:0042 613
ENSDARG000000 16319	c9	full_mo del	24hpir_intes tine	GLF	
ENSDARG000000 68621	nlrcl1	full_mo del	24hpir_intes tine	GLF	GO:0005 515

Appendix Chapter two

ENSDARG000000 12694	c3a.1	full_mo del	24hpir_intes tine	GLF	GO:0004 866
ENSDARG000000 55278	cfb	full_mo del	24hpir_intes tine	GLF	GO:0006 956
ENSDARG000001 02456	cfhl4	full_mo del	24hpir_intes tine	GLF	
ENSDARG000000 79105	mhc2dab	full_mo del	24hpir_kidn ey	GLC	GO:0042 613
ENSDARG000000 91280	si:ch211- 66k16.27	full_mo del	24hpir_kidn ey	GLC	GO:0042 981
ENSDARG000000 67848	nmrk2	full_mo del	24hpir_kidn ey	GLC	GO:0005 737
ENSDARG000000 38587	LOC10000 4199	full_mo del	24hpir_kidn ey	GLC	
ENSDARG000000 37550	cyp21a2	full_mo del	24hpir_kidn ey	GLC	GO:00167 05
ENSDARG000000 68621	nlrcl1	full_mo del	24hpir_kidn ey	GLF	GO:0005 515
ENSDARG000000 91847	LOC110437 747	full_mo del	24hpir_kidn ey	GLF	GO:0005 515
ENSDARG000000 94210	fthl31	full_mo del	24hpir_kidn ey	GLF	GO:0006 879

Appendix Chapter two

ENSDARG000000 16771	tfa	full_mo del	24hpir_kidn ey	GLF	
ENSDARG000000 86569	zgc:172051	full_mo del	24hpir_kidn ey	GLF	
ENSDARG000000 79105	mhc2dab	full_mo del	24hpir_caud al fin	GLC	GO:0042 613
ENSDARG000000 42988	slc24a2	full_mo del	24hpir_caud al fin	GLC	
ENSDARG000001 04635	mhc2dgb	full_mo del	24hpir_caud al fin	GLC	
ENSDARG000000 44074	lox12b	full_mo del	24hpir_caud al fin	GLC	GO:00166 41
ENSDARG000000 95082	si:ch211- 194p6.12	full_mo del	24hpir_caud al fin	GLC	
ENSDARG000000 68621	nlrcl1	full_mo del	24hpir_caud al fin	GLF	GO:0005 515
ENSDARG000000 91847	LOC110437 747	full_mo del	24hpir_caud al fin	GLF	GO:0005 515
ENSDARG000000 62788	irg1l	full_mo del	24hpir_caud al fin	GLF	GO:00168 29
ENSDARG000001 6832	NA	full_mo del	24hpir_caud al fin	GLF	GO:00150 74

Appendix Chapter two

ENSDARG000000 07769	sult5a1	full_mo del	24hpir_caud al fin	GLF	GO:00167 40
ENSDARG000000 79105	mhc2dab	full_mo del	24hpir_sple en	GLC	GO:0042 613
ENSDARG000001 04635	mhc2dgb	full_mo del	24hpir_sple en	GLC	
ENSDARG000000 38587	LOC10000 4199	full_mo del	24hpir_sple en	GLC	
ENSDARG000001 6524	mhc2dhb	full_mo del	24hpir_sple en	GLC	GO:0042 613
ENSDARG000000 71024	zgc:171679	full_mo del	24hpir_sple en	GLC	GO:0005 515
ENSDARG000000 20084	tg	full_mo del	24hpir_sple en	GLF	
ENSDARG000000 68621	nlrcl1	full_mo del	24hpir_sple en	GLF	GO:0005 515
ENSDARG000000 69296	moxd1l	full_mo del	24hpir_sple en	GLF	GO:0004 500
ENSDARG000000 91847	LOC110437 747	full_mo del	24hpir_sple en	GLF	GO:0005 515
ENSDARG000000 76043	si:dkeyp- 73d8.9	full_mo del	24hpir_sple en	GLF	

Appendix Chapter two

ENSDARG000000 15866	apoa2	germlin e	24hpir_tests	GLC	
ENSDARG000000 75261	timp2b	germlin e	24hpir_tests	GLC	GO:00081 91
ENSDARG000000 03533	col8a1b	germlin e	24hpir_tests	GLC	GO:0005 515
ENSDARG000000 69282	bbc3	germlin e	24hpir_tests	GLC	GO:0090 200
ENSDARG000001 01324	apoa1b	germlin e	24hpir_tests	GLC	GO:00421 57

Table S 6: Mixed-effects modelling of somatic gene features.

Linear mixed-effects models (LMMs; lmer, R) fitted to the full somatic dataset were used to analyse **A)** gene expression and **B)** gene length, while generalised linear mixed-effects models (GLMMs; glmer, R) were applied to (C) the proportion of exonic sequence. All models include random effects to account for hierarchical structure within the data.

A)

Effect	Gene Expression (logCPM)		
	X ²	df	p
Strain	63	1	< 0.0001
Contrast	2412	17	< 0.0001
Gene Length	770	1	< 0.0001
Chromatin	11	1	< 0.001
Strain : Contrast	2252	17	< 0.0001
Random Effects Variance	Chromosome : 0.08, SD : 0.28 Strand : 0.001, SD : 0.03		

Appendix Chapter two

B)

Effect	Gene Length (bp)		
	X ²	df	p
Strain	34	1	< 0.0001
Contrast	717	17	< 0.0001
Chromatin	7052	1	< 0.0001
Strain : Contrast	1538	17	< 0.0001
Random Effects Variance	Chromosome : 0.007, SD : 0.08 Strand : 0.0002, SD = 0.01		

Appendix Chapter two

C)

Effect	Gene Exonic Region		
	X ²	df	p
Strain	18731	1	< 0.0001
Contrast	146372	17	< 0.0001
Gene Length	124737573	1	< 0.0001
Chromatin	321416	1	< 0.0001
Strain : Contrast	301334	17	< 0.0001
Random Effects Variance	Chromosome: 0.005, SD : 0.07 Strand : 0.0001, SD = 0.1		

Table S 7: Pairwise comparisons of estimated marginal means from mixed-effects models.

Pairwise comparisons of estimated marginal means (EMMs) were derived from linear mixed-effects models (LMMs; lmer, R) for **A)** gene expression and **B)** gene length, and from generalised linear mixed-effects models (GLMMs; glmer, R) for **C)** the proportion of exonic sequence. P-values were adjusted for multiple comparisons using Tukey's honestly significant difference (HSD) procedure.

A)

Contrast	Estimate	SE	df	z.ratio	p-value
GLC C1 - GLF C1	-0.492	0.062	Inf	-7.932	<.0001
GLC C2.3 - GLF C2.3	0.725	0.05392563	Inf	13.442	<.0001
GLC C2.24 - GLF C2.24	-0.711	0.0553	Inf	-12.874	<.0001
GLC C1 - GLC C2.3	-0.572	0.052	Inf	-11.097	<.0001
GLC C1 - GLC C2.24	0.409	0.051	Inf	8.009	<.0001
GLC C2.24 - GLC C2.3	-0.981	0.051	Inf	-19.167	<.0001
GLF C1 - GLF C2.3	0.645	0.064	Inf	10.077	<.0001
GLF C1 - GLF C2.24	0.190	0.065	Inf	2.898	<.0001
GLF C2.24 - GLF C2.3	0.455	0.058	Inf	7.882	<.0001
GLC C3.br - GLF C3.br	-0.249	0.073	Inf	-3.428	<.0001

Appendix Chapter two

GLC C3.int - GLF C3.int	-0.451	0.053	Inf	-8.564	<.0001
GLC C3.kid - GLF C3.kid	-0.416	0.051	Inf	-8.133	<.0001
GLC C3.ski - GLF C3.ski	0.108	0.051	Inf	2.123	<.0001
GLC C3.spl - GLF C3.spl	-0.533	0.051	Inf	-10.465	<.0001
GLC C4.br.3 - GLF C4.br.3	1.387	0.053	Inf	26.314	<.0001
GLC C4.int.3 - GLF C4.int.3	0.2821	0.0491	Inf	5.7407	<.0001
GLC C4.kid.3 - GLF C4.kid.3	-0.048	0.049	Inf	-0.979	<.0001
GLC C4.ski.3 - GLF C4.ski.3	0.765	0.050	Inf	15.238	<.0001
GLC C4.spl.3 - GLF C4.spl.3	0.9058	0.0459	Inf	19.7469	<.0001
GLC C5.br.24 - GLF C5.br.24	-0.258	0.071	Inf	-3.611	<.0001
GLC C5.int.24 - GLF C5.int.24	-0.525	0.047	Inf	-11.236	<.0001
GLC C5.kid.24 - GLF C5.kid.24	-0.367	0.046	Inf	-8.000	<.0001
GLC C5.ski.24 - GLF C5.ski.24	-0.129	0.046	Inf	-2.786	<.0001
GLC C5.spl.24 - GLF C5.spl.24	-0.307	0.054	Inf	-5.681	<.0001

Appendix Chapter two

B)

Contrast	Estimate	SE	df	z.ratio	p-value
GLC C1 - GLF C1	0.1793	0.0309	Inf	5.8037	<.0001
GLC C2.3 - GLF C2.3	-0.3542	0.0268	Inf	-13.1951	<.0001
GLC C2.24 - GLF C2.24	0.1444	0.0275	Inf	5.2462	<.0001
GLC C1 - GLC C2.3	0.2508	0.0256	Inf	9.7795	<.0001
GLC C1 - GLC C2.24	0.0307	0.0254	Inf	1.2079	<.0001
GLC C2.24 - GLC C2.3	0.2201	0.0255	Inf	8.6381	<.0001
GLF C1 - GLF C2.3	-0.2826	0.0319	Inf	-8.8646	<.0001
GLF C1 - GLF C2.24	-0.0042	0.0326	Inf	-0.1274	<.0001
GLF C2.24 - GLF C2.3	-0.2784	0.0288	Inf	-9.6776	<.0001
GLC C3.br - GLF C3.br	0.3898	0.0361	Inf	10.7859	<.0001
GLC C3.int - GLF C3.int	0.1611	0.0262	Inf	6.1448	<.0001
GLC C3.kid - GLF C3.kid	0.0510	0.0255	Inf	1.9999	<.0001
GLC C3.ski - GLF C3.ski	0.0344	0.0253	Inf	1.3622	<.0001
GLC C3.spl - GLF C3.spl	0.0758	0.0253	Inf	2.9904	<.0001
GLC C4.br.3 - GLF C4.br.3	-0.5949	0.0262	Inf	-22.6935	<.0001
GLC C4.int.3 - GLF C4.int.3	-0.0864	0.0245	Inf	-3.5301	<.0001
GLC C4.kid.3 - GLF C4.kid.3	-0.0889	0.0246	Inf	-3.6125	<.0001

Appendix Chapter two

GLC C4.ski.3 - GLF C4.ski.3	-0.3537	0.0250	Inf	-14.1572	<.0001
GLC C4.spl.3 - GLF C4.spl.3	-0.3856	0.0228	Inf	-16.8984	<.0001
GLC C5.br.24 - GLF C5.br.24	0.2516	0.0356	Inf	7.0681	<.0001
GLC C5.int.24 - GLF C5.int.24	0.1738	0.0233	Inf	7.4723	<.0001
GLC C5.kid.24 - GLF C5.kid.24	0.0165	0.0228	Inf	0.7247	<.0001
GLC C5.ski.24 - GLF C5.ski.24	0.1368	0.0230	Inf	5.9434	<.0001
GLC C5.spl.24 - GLF C5.spl.24	0.0195	0.0269	Inf	0.7238	<.0001

Appendix Chapter two

C)

Contrast	Estimate	SE	df	z.ratio	p-value
GLC C1 - GLF C1	-0.0732	0.0005	Inf	-136.8623	<.0001
GLC C2.3 - GLF C2.3	0.0358	0.0005	Inf	77.2684	<.0001
GLC C2.24 - GLF C2.24	-0.0537	0.0005	Inf	-111.2968	<.0001
GLC C1 - GLC C2.3	-0.0499	0.0004	Inf	-111.3901	<.0001
GLC C1 - GLC C2.24	0.0077	0.0004	Inf	17.6583	<.0001
GLF C1 - GLF C2.3	0.0591	0.0005	Inf	107.8751	<.0001
GLF C1 - GLF C2.24	0.0272	0.0006	Inf	47.3808	<.0001
GLF C2.24 - GLF C2.3	0.0319	0.0005	Inf	64.2225	<.0001
GLC C3.br - GLF C3.br	-0.1395	0.0007	Inf	-210.0547	<.0001
GLC C3.int - GLF C3.int	-0.0271	0.0005	Inf	-57.3460	<.0001
GLC C3.kid - GLF C3.kid	-0.0704	0.0004	Inf	-158.4571	<.0001
GLC C3.ski - GLF C3.ski	-0.0305	0.0004	Inf	-67.8595	<.0001
GLC C3.spl - GLF C3.spl	-0.0809	0.0005	Inf	-178.1585	<.0001
GLC C4.br.3 - GLF C4.br.3	0.0932	0.0005	Inf	193.8331	<.0001
GLC C4.int.3 - GLF C4.int.3	0.0443	0.0004	Inf	101.3887	<.0001
GLC C4.kid.3 - GLF C4.kid.3	0.0037	0.0004	Inf	8.4026	<.0001
GLC C4.ski.3 - GLF C4.ski.3	0.0419	0.0004	Inf	94.7956	<.0001

Appendix Chapter two

GLC C4.spl.3 - GLF C4.spl.3	0.0941	0.0004	Inf	232.0648	<.0001
GLC C5.br.24 - GLF C5.br.24	-0.0599	0.0006	Inf	-92.7293	<.0001
GLC C5.int.24 - GLF C5.int.24	-0.0443	0.0004	Inf	-106.2674	<.0001
GLC C5.kid.24 - GLF C5.kid.24	-0.0533	0.0004	Inf	-130.1877	<.0001
GLC C5.ski.24 - GLF C5.ski.24	0.0219	0.0004	Inf	53.4500	<.0001
GLC C5.ski.24 - GLF C5.spl.24	-0.0247	0.0004	Inf	-56.3571	<.0001
GLC C5.spl.24 - GLF C5.spl.24	-0.0706	0.0005	Inf	-147.3544	<.0001

Table S 8: Mixed-effects modelling of germline gene features.

Linear mixed-effects models (LMMs; lmer, R) fitted to the germline dataset were used to analyse **A)** gene expression, **B)** gene length, and **C)** the proportion of exonic sequence.

A)

Effect	Gene Expression (logCPM)		
	X ²	df	p
Irradiation Effect	132	1	< 0.0001
Gene Length	15	1	< 0.0001
Chromatin	20	1	< 0.0001
Random Effects Variance	Chromosome : 0.08, SD : 0.3		

B)

Effect	Gene Length (bp)		
	X ²	df	p
Irradiation Effect	498	1	< 0.0001
Chromatin	1277	1	< 0.0001
Random Effects Variance	Chromosome: 0.008, SD = 0.09		

Appendix Chapter two

C)

Effect	Exonic Region Length		
	X ²	df	p
Irradiation Effect	25951	1	< 0.0001
Gene Length	23358721	1	< 0.0001
Chromatin	36931	1	< 0.0001
Random Effects Variance	Chromosome: 0.005, SD : 0.07 Strand: 0.00008, SD : 0.009		

Table S 9: Gene Index for the alluvial plot shown in Figure 5 of Chapter 2

Table lists all genes contributing to the alluvial flows, provided as a reference to enable identification of individual genes underlying each pathway-strain-timepoint interaction.

	Timepoint		
Gene Index	nonir	3hpir	24hpir
1	abat	abat	abl1
2	abl1	agt	acta2
3	ace	amh	agt
4	aco1	arg2	akap9
5	aco2	atr	appl2
6	adsl	aurkb	atr
7	agap2	cab39	b2m
8	agt	calca	bcar1
9	ak4	calr	blink
10	atp7a	casp9	brca2
11	atp8a2	cd68	clqa
12	atr	cdk2	clqc
13	bnip3	chek2	c4b
14	clqa	cisd2	c5

Appendix Chapter two

15	calca	cited2	c8a
16	calr	cln8	c8g
17	canx	cnp	cacna1g
18	card14	col4a2	calca
19	casp9	comp	calr
20	cat	cops8	cat
21	chchd10	ctcl	cblb
22	cisd2	ctnna1	ccdc88c
23	comp	ctsc	cd276
24	cox15	daxx	cd68
25	cox4i2	dbf4b	cdk2
26	cox6a1	ddc	cdk6
27	cox6a2	dnmbp	cfb
28	cycl	edn1	cfi
29	dag1	ercc1	clu
30	edn1	fbxo4	cmklr1
31	emp2	fgf2	cmtm3
32	endog	foxm1	colec1l
33	ercc1	foxo3b	csk

Appendix Chapter two

34	fh	gclc	ctcl
35	foxm1	gprc5c	cytl
36	foxo3b	hamp	dagl
37	fxn	il23r	ddc
38	gas6	irak1	dnmbp
39	gprc5c	jund	eef1e1
40	ikbkb	kif14	eif2b3
41	il6r	kmo	endog
42	inpp5d	lmna	fbxo4
43	ireb2	lrrk2	fcho1
44	itgb1bp1	map2k1	fgf2
45	kat6a	map3k3	flna
46	malt1	map3k5	foxo3b
47	map3k3	mapkapk5	gata3
48	mb12	mbd2	gbp1
49	mob3c	mif	gbp2
50	msh6	mmp2	gclc
51	nckap1l	mob3c	gpd1l
52	ndnl2	ndnl2	hck

Appendix Chapter two

53	ndufa2	nsmce2	irak1
54	ndufa6	nudt1	itk
55	ndufa7	p2ry1	klhl6
56	ndufaf1	park7	lck
57	ndufb2	pawr	malt1
58	ndufb3	pck1	mapk1
59	ndufb6	pdx1	mb12
60	ndufc1	phb	mybpc3
61	ndufs5	pibf1	myh10
62	ndufv3	pink1	myh6
63	nrg1	pitx3	myl6
64	park7	ppp3ca	myo5b
65	pawr	prelp	myo5c
66	pck1	rad54b	nedd4l
67	pibf1	rad54l	nek6
68	pik3ca	rgcc	nono
69	ppp2r3c	ripk3	pck1
70	pqbp1	romo1	pde4d
71	prkdc	rplp1	pik3cd

Appendix Chapter two

72	pycard	rpn2	pkp2
73	rgcc	rwdd1	pla2r1
74	ripk3	scap	polb
75	rptor	sec63	pqbp1
76	sdhaf2	serpine1	prkd2
77	sdhc	serpinf1	ptprc
78	shmt2	slc1a1	pycard
79	sirt3	slc30a10	rb1l
80	slc1a1	smc5	rc3h2
81	slc25a33	sod1	rela
82	src	srebf1	serping1
83	stoml2	srr	skap1
84	suc1g2	stk11	slc12a2
85	tbk1	stradb	src
86	tlr1	tal1	srebf1
87	tlr3	tert	srr
88	tomm70a	tfcp2l1	susd4
89	tpx2	tgfb3	thyl
90	traf6	tlr1	tnfrsf21

Appendix Chapter two

91	trpv4	tlr3	tomm70a
92	uqcrc1	tp53	tp53
93	uqcrfs1	tpra1	tyrobp
94	wnt11	ucp2	ucp2
95	znfx1	ulk3	ulk3
96		vash1	vash1

8.3. Supplementary Figures

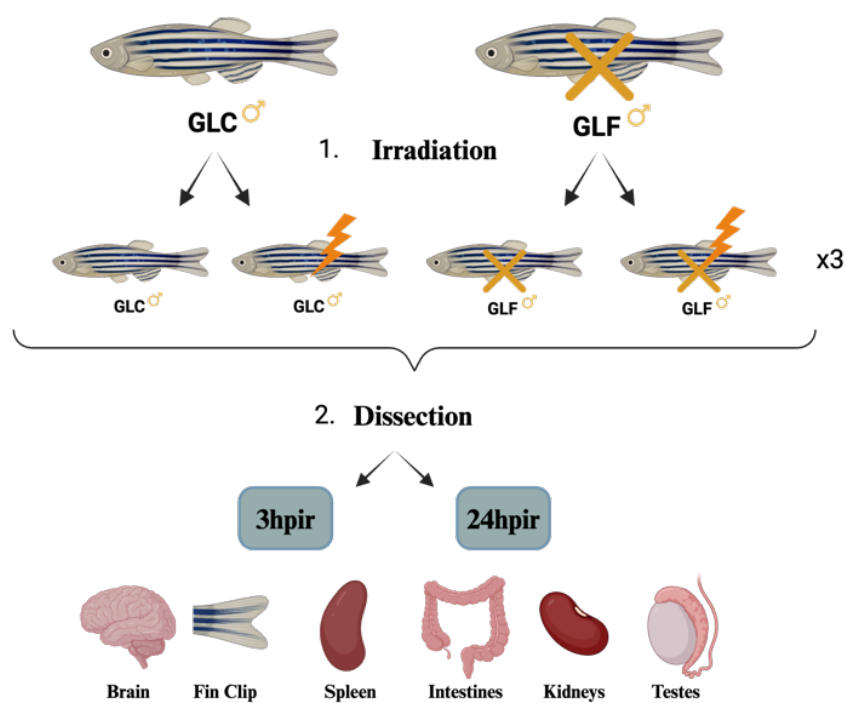


Figure S 1: Experimental design.

Samples from three males per treatment (GLC or GLF), irradiation (non-irradiated or irradiated) and timepoint post-irradiation (3hours post irradiation (hpir) or 24hpir) were collected under a full-factorial design adding up to a total of 24 males.

Appendix Chapter two

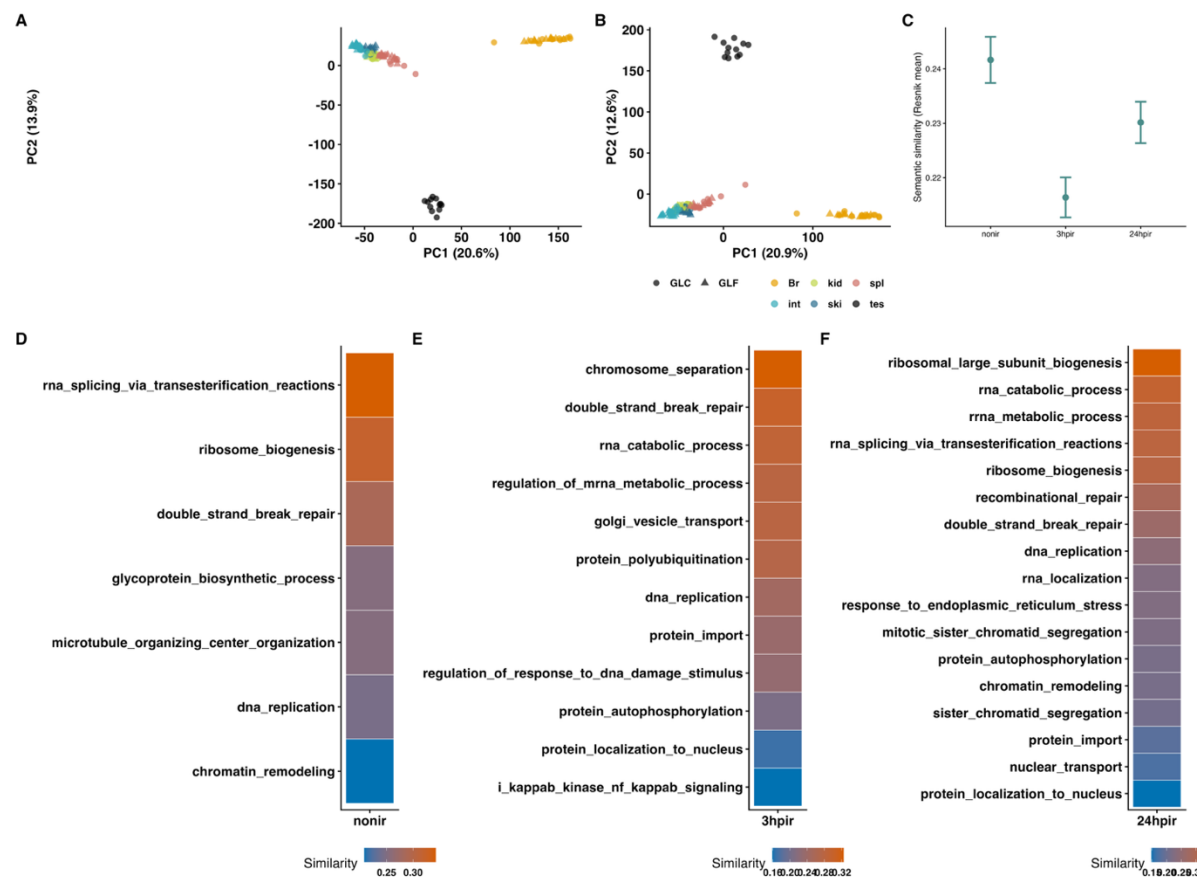


Figure S 2: Cellular Origin defines transcriptional profiles of both genes and TEs; and Resnik similarity reveals conserved genomic stability functions and divergent chromatin-associated processes across strains and timepoints post irradiation.

Variation across all samples used in the contrasts is shown, with strain (GLC and GLF fish) as shapes and colour-coded by tissue. **A)** shows a PCA of differentially expressed genes per sample, and **B)** shows a PCA of differentially expressed TEs per sample. **C)** shows mean Resnik similarity between GLC and GLF fish between timepoints, with lower values indicating lower similarity. Resnik heatmaps visualise the semantic similarity between GLC and GLF gene sets for the 100 most statistically enriched GO terms (lowest *p*-values) per condition. GO terms were selected from differential expression results and represent biologically salient responses to **D)** non-irradiated (nonir), **E)** 3 h post-irradiation, and **F)** 24 h post-irradiation states. Colour intensity reflects the degree of functional similarity between gene sets associated with each GO term, highlighting patterns of convergence or divergence in biological processes between strains. GO terms are ordered by increasing similarity within each condition.

Appendix Chapter two

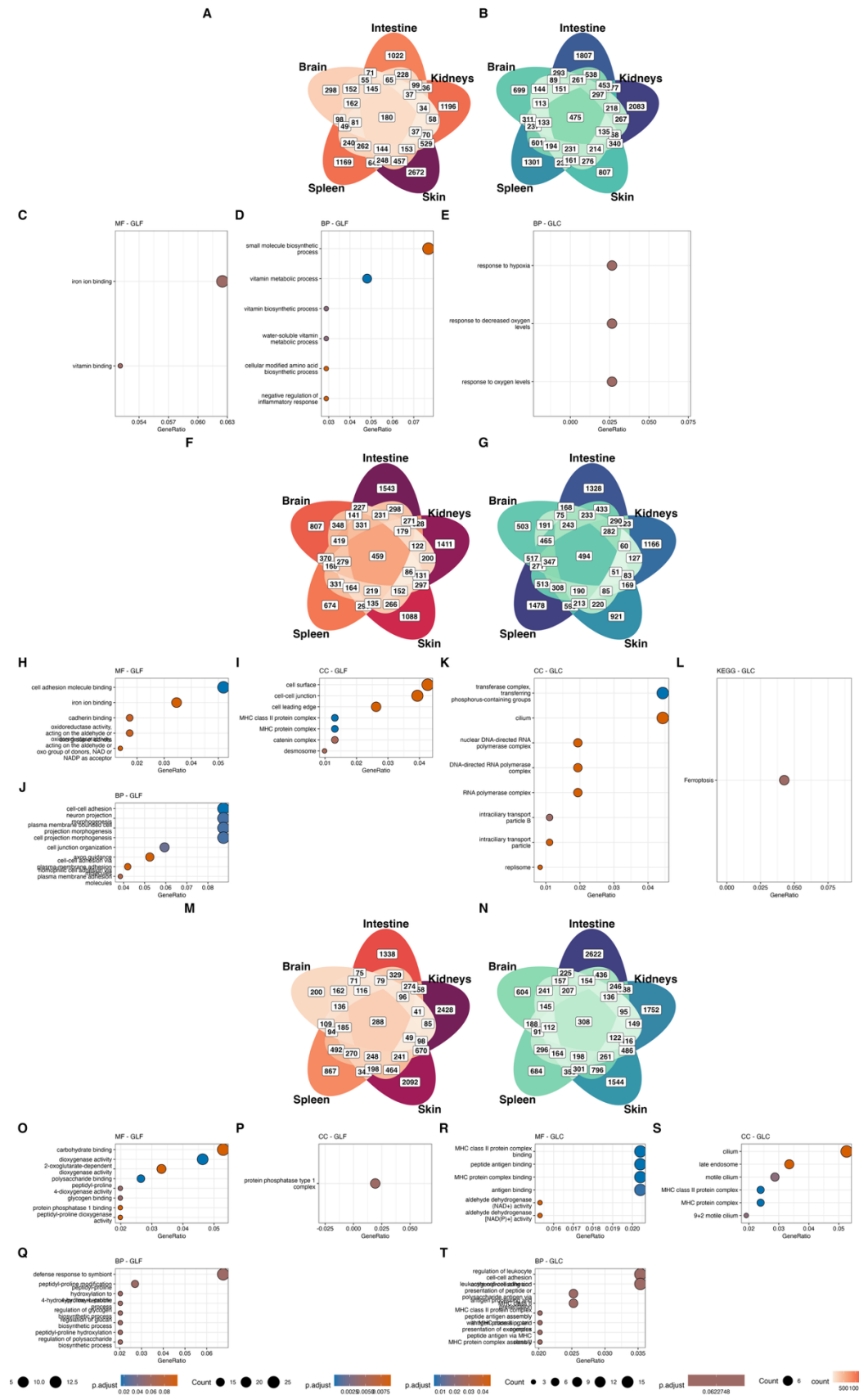


Figure S 3 : Comparative functional analysis of somatic gene expression reveals specialised metabolic and immune programmes in all GLF somatic tissues and timepoint depending immune expression.

A) Differentially expressed genes upregulated in the soma of non-IR GLF fish. **B)** Differentially expressed genes upregulated in the soma of non-IR GLC fish. **(C-D)** Functional enrichment of 180 overlapping genes in soma of non-IR GLF fish, **C)** Molecular Function **(MF)**, **D)** Biological Processes (BP). **E)** function enrichment of 475 overlapping genes in the soma of non-IR GLC fish based on BP. **F)** Differentially expressed genes upregulated in the soma of GLF fish 3hpir. **G)** Differentially expressed genes upregulated in the soma of GLC fish. **(H-J)** Functional enrichment of 459 overlapping genes in the soma of GLF fish 3hpir, based on **H)** MF, **I)** CC and **J)** BP. **(K-L)** Functional enrichment of 494 overlapping genes in the soma of GLC fish 3hpir, based on **K)** CC and **L)** KEGG. **M)** Differentially expressed genes upregulated in the soma of GLF fish 24hpir. **N)** Differentially expressed genes upregulated in the soma of GLC fish 24hpir. **(O-Q)** Functional enrichment of 288 overlapping genes in the soma of GLF fish 24hpir., based on **O)** MF, **P)** CC) and **Q)** BP. **(R-T)** Functional enrichment of 308 overlapping genes in the soma of GLC fish 24hpir., based on **R)** MF, **S)** CC and **T)** BP. Only terms with adjusted $p < 0.1$ are displayed.

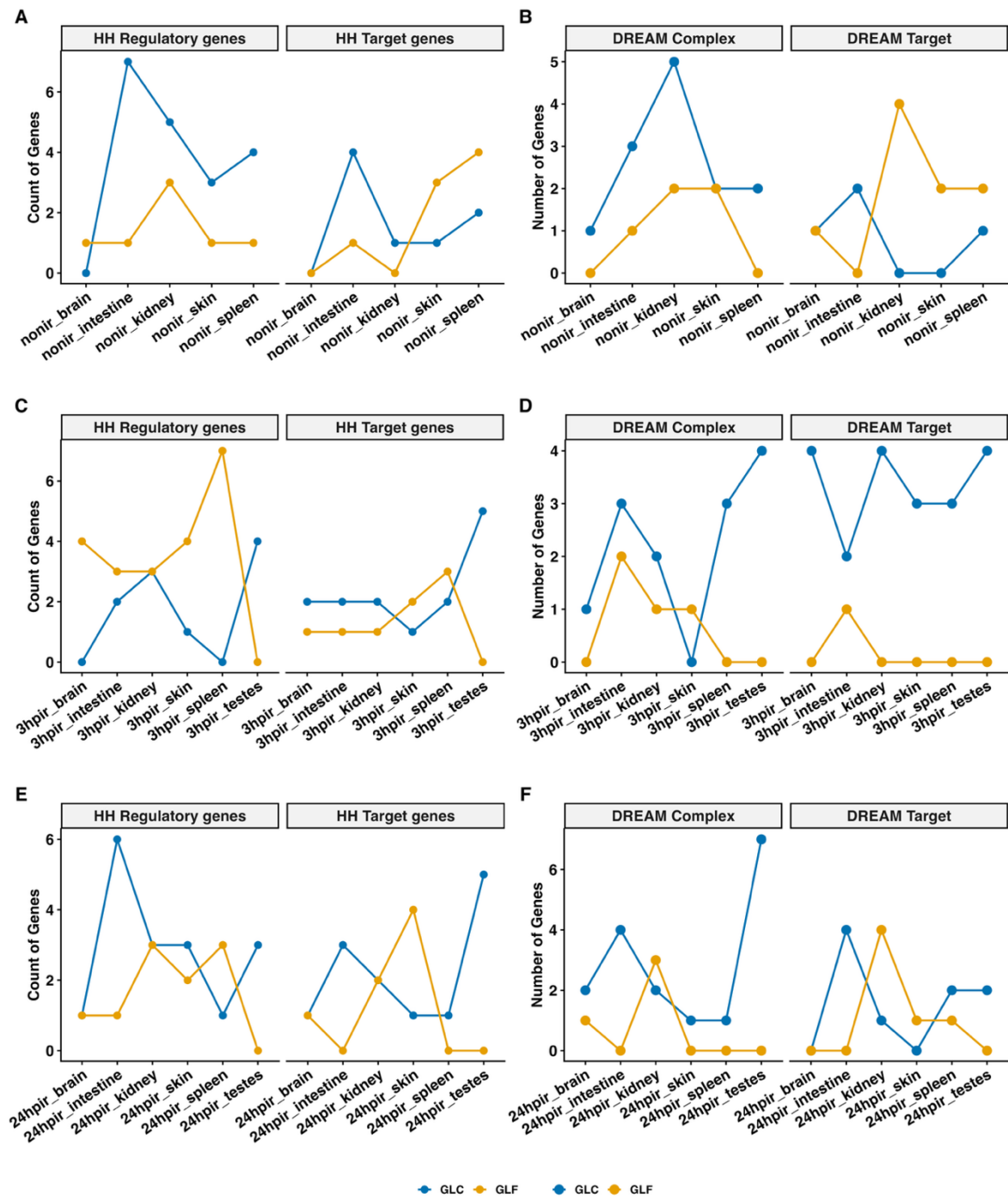


Figure S 4: GLF soma shows a transient shift in Hedgehog pathway regulators (HH) and Dream Complex target genes at 3hpir, driven by spleen for HH regulators and all soma for Dream genes.

Line plot showing the number of differentially expressed genes (FDR < 0.05) across timepoints for each fish tissue between GLC and GLF fish, across Hedgehog regulatory genes and target genes **A), C), E)** and Dream complex target and regulatory genes **B), D), F)**. **(A-B)** non-irradiated somatic tissues, **(C-D)** tissues 3hpir, and **(E-F)** tissues 24hpir. Gene counts are grouped by gene type (e.g., Hedgehog pathway complex vs targets panels or Dream complex vs target genes; see facet). Lines and points represent the dynamics of DE gene counts across developmental timepoints for each fish, coloured according to population-specific palettes.

Appendix Chapter two

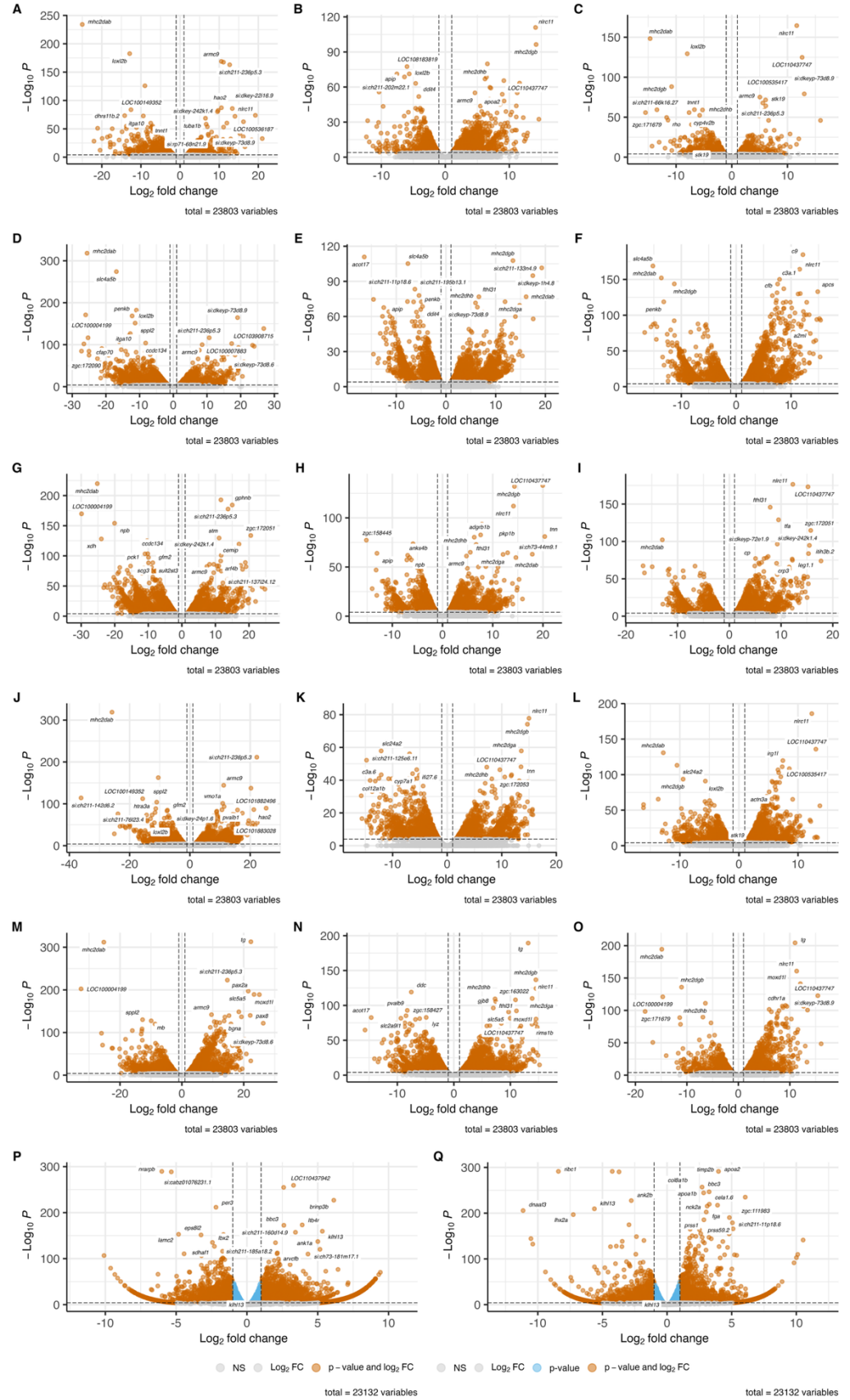


Figure S 5: Strain-specific timing of somatic transcriptional response to germline ablation and irradiation extends to each tissue.

Each point represents a single gene, with the x-axis denoting the $\log_2$ fold change in expression between GLC and GLF strains, and the y-axis showing the $-\log_{10}(\text{p-value})$, indicating statistical significance. Genes with a positive logFC are more highly expressed in GLF, whereas those with negative logFC values are more highly expressed in GLC. Genes not reaching significance ($\text{FDR} \geq 0.05$) are depicted in grey. Vertical and horizontal dashed lines mark fold-change thresholds ($\log_2\text{FC} = \pm 0.6$) and the significance threshold ($-\log_{10}(\text{p-value}) = 1.3$), respectively. **A)** presents differential expression in non-irradiated Brain, **B)** in Brain at 3 hours post-irradiation, and **C)** in Brain at 24 hours post-irradiation (D–F) show corresponding plots for Intestine, **(G–I)** for Kidneys, **(J–L)** for Caudal fin, and **(M–O)** for Spleen. **P)** shows the differential expression of contrast C6.3 (Non-IR GLC testes vs 3hpir GLC testes) & **Q)** the differential expression of contrast C5.24 (Non-IR GLC testes vs 24hpir GLC testes).

Appendix Chapter two

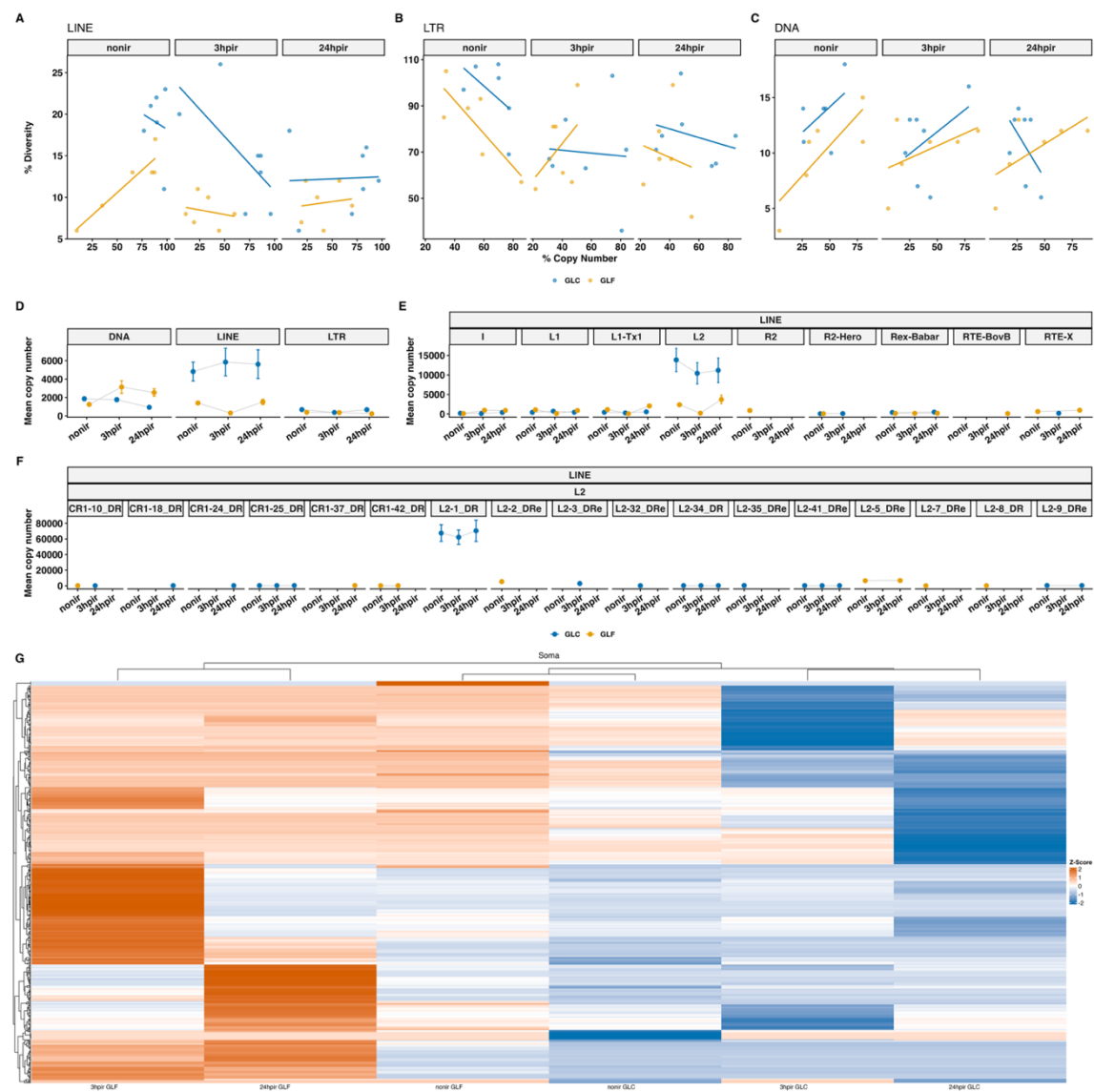


Figure S 6: Strain-specific divergence in transposable element copy number and expression following irradiation

(A–C) show scatterplots of the percentage of total copy number expressed (x-axis) versus the number of significantly expressed repeats (y-axis) for LINE **A**), LTR **B**), and DNA **C**) transposons, respectively. Each point represents a strain-specific contrast at a given timepoint (non-irradiated, 3 hours post-irradiation or 24 hours post-irradiation) with colour indicating strain identity. Linear models are fitted per panel to visualise the relationship between copy number representation and TE diversity. **D**) shows the mean copy number ($\pm$ standard error) of significantly expressed LINE, LTR, and DNA transposable element classes across timepoints (non-irradiated, 3 hours post-irradiation 24 hours post-irradiation coloured by fish strain. **E**) focuses on LINE subfamilies, showing mean copy number trajectories for each significantly expressed subfamily across time, highlighting distinct patterns of activation among strains. **F**) presents repeat-level copy number dynamics within the L2 subfamily, identifying specific L2 elements driving differential expression across timepoints. Data represent mean copy number values of significantly expressed repeats ($\text{FDR} < 0.05$), with error bars indicating standard error. **G**) Heatmap shows scaled log fold change (Z-score) of TE Repeats (rows) across somatic tissues (columns) from the nonirVShpir contrast group. Each value represents the mean logFC per TE Repeat per tissue, computed across significant comparisons ($\text{FDR} < 0.05$). Row-wise Z-score normalisation was applied to centre TE-specific responses. Both rows and columns are hierarchically clustered. Only TE subfamilies with complete expression profiles across tissues are shown. Colour indicates relative upregulation (orange) or downregulation (blue) based on the Okabe–Ito colour palette.

Appendix Chapter two

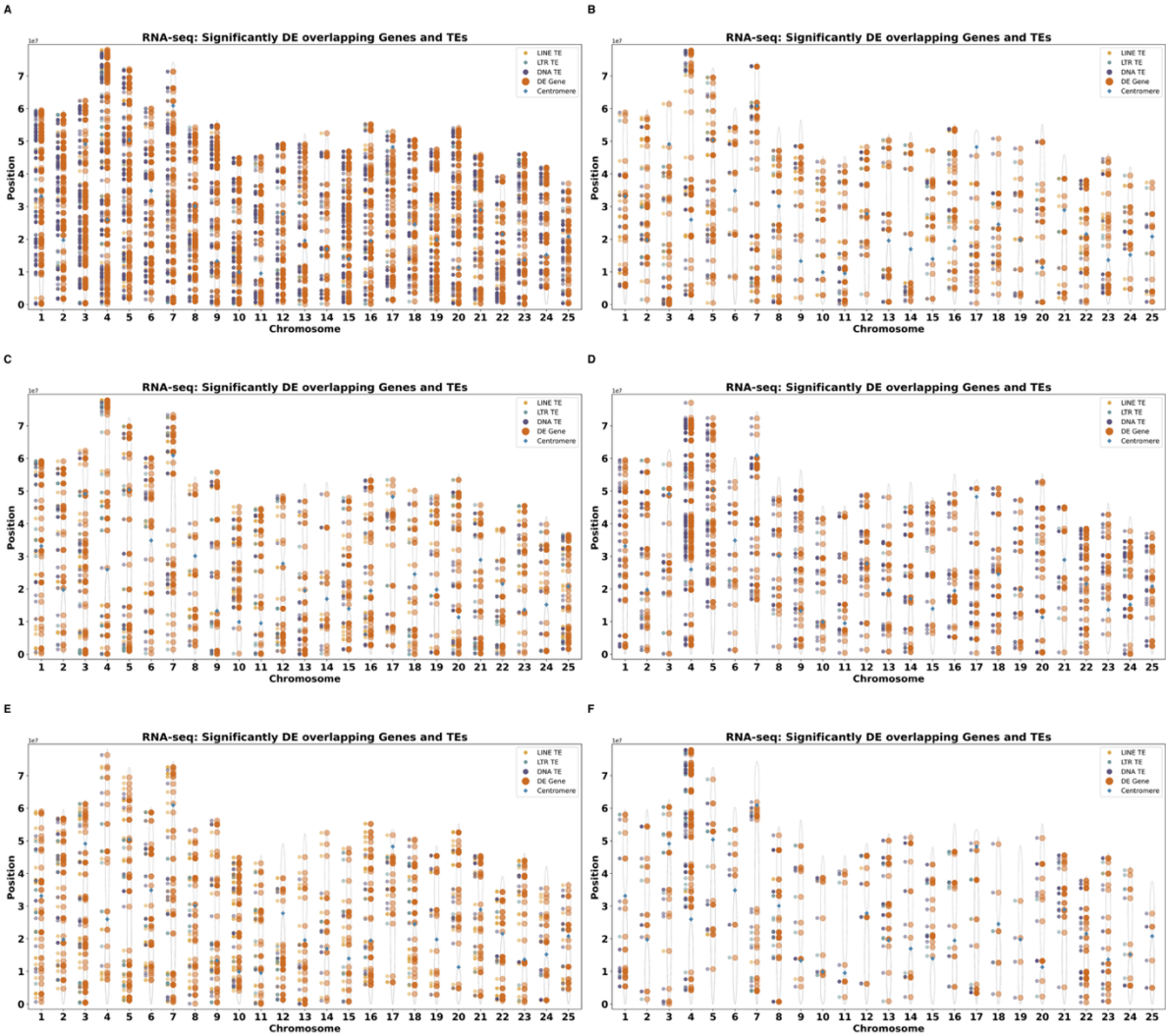


Figure S 7: Mapping of DE genes and TEs reveals chromosome 4 as a hotspot of enrichment in GLF fish in response to irradiation.

(A-F) represent strain and timepoint-specific comparisons. **A)** GLC nonir, **B)** GLC nonir, **C)** GLC 3hpir, **D)** GLF 3hpir, **E)** GLC 24hpir, and **F)** GLF 24hpir. Differentially expressed genes are shown in orange, whilst overlapping differentially expressed TEs are classified through colour by Class : LINE = yellow, LTR = Blye, and DNA = Purple. Each point corresponds to the genomic start position of either a TE or a gene, plotted on the Y axis against chromosome identity on the x-axis. TE positions are offset laterally to improve visual interpretation. Centromere locations are marked with blue diamonds, and grey ellipses approximate chromosome arms based on known centromeric coordinates

9. Appendix Chapter three

Supporting information for Defining the determinants of a robust Epigenetic Clock

9.1. Supplementary tables

Table S 10: Number of samples per tissue and sex used in the experimental study

Age	Female			Male		
	Brain	Intestine	Caudal Fin	Brain	Intestine	Caudal Fin
3	6	6	6	7	7	5
4	2	4	4	2	2	1
5	5	3	5	6	4	7
8	2	2	2	3	5	3
10	4	4	4	2	2	2
12	2	2	2	2	2	2
14	3	3	3	3	3	1
16	3	3	2	3	3	3
18	3	3	3	2	2	2
22	6	6	6	7	7	5

Table S 11: Distribution of samples across technical and biological factors

CPG Aggregation		Regression model		R ² Threshold		Feature Filter		Eps		Tissue	
Cluster	1312	Elastic Net	1738	No filter	315	No Filter	806	1	111	Brain	859
Point	111	Lasso	1568	0.05	731	Min 5	810	50	412	Caudal Fin	816
Tile	1883			0.1	755	Min 10	845	100	419	Intestine	824
				0.2	762	Min 20	845	500	417	Multi-tissue	807
				0.3	743			750	149		
								1000	417		
								1500	149		
								2000	421		
								3000	139		

Appendix Chapter three

								5000	27 2		
								2250 0	12 8		
								2500 0	27 2		

Table S 12: Top-selected CpG sites identified in single-tissue and multi-tissue clocks, and their overlap across somatic tissues.

Brain	Caudal Fin	Intestine	Multi-tissue	Overlapping Soma
2:58868001:5 8869240	1:712001:7130 00	1:51950001:51 955000	1:24076877:2 4076977	1:14372733:14 373618
4:1658618:165 8618	1:24077077:2 4077150	2:3575001:36 00000	1:24077077:2 4077126	1:40644559:4 0644788
5:22750001:2 2775000	1:41264001:41 266000	2:11200001:112 25000	2:34600001:3 4625000	3:13557565:13 558122
6:14882001:14 883000	2:32775001:32 800000	2:57232001:57 233798	2:58868741:5 8869240	4:3011851:301 2149
8:5001:12000	2:57232001:57 233044	3:15875001:15 900000	3:1105225:1105 225	4:51645740:5 1646027
9:15788482:15 788482	3:3650001:36 51000	3:23525001:23 576000	3:15875001:15 900000	4:77559642:7 7559642
10:38075001: 38100000	3:11600001:116 25000	4:12928001:12 929000	3:23570001:23 576000	4:77560396:7 7560396
11:42714001:4 2716028	3:23550001:23 576000	4:77550001:7 7575000	4:17675001:17 700000	4:77561200:7 7561200
12:16564778:1 6564778	4:17675001:17 700000	5:24200001:2 4205000	4:31950001:31 975000	5:38774677:3 8775285

Appendix Chapter three

12:34320001:3 4325000	4:34312001:34 314000	5:43683051:4 3683100	4:34312001:34 314000	6:4935441:49 35768
14:7245001:7 247000	5:39427001:3 9428000	5:43683104:4 3683105	4:42625001:4 2650229	7:34162239:3 4162861
14:13550001:1 3575000	5:65610001:65 616000	6:58690001:5 8696000	6:64024:6402 4	8:31120071:311 20427
17:174609:179 608	7:20967884:2 0970000	7:1:25000	7:5001:10000	8:50628644: 50628910
19:32030001: 32035000	7:68415001:68 420000	7:18598490:18 598711	8:53670663:5 3671090	8:51371072:51 371541
20:37140001: 37145000	9:11340947:113 41050	7:21885001:21 890000	13:13670001:13 672000	13:3704291:37 04340
22:11184001:11 190000	10:720001:724 000	7:32720001:32 725000	13:14567598:1 4567598	13:4766104:4 766350
22:36300001: 36325000	10:3746001:37 46500	7:63325001:63 330000	13:51245001:51 250000	15:42582708: 42583050
23:16300001:1 6325000	11:26750001:2 6775000	8:42620001:4 2626000	14:214333:214 482	16:5093555:5 093769
	11:41769001:41 770000	10:10875001:1 0900000	14:4182269:41 82401	16:815631:816 213
	12:33192127:33 192400	10:28150001:2 8175000	14:28052001:2 8055000	17:9936654:9 936703

Appendix Chapter three

	12:35875001:3 5900000	11:10828040:1 0828428	15:21200001:2 1226000	2:44181891:44 182390
	13:1532001:153 4000	14:3570001:35 75000	15:28199001:2 8199500	2:6262120:62 62331
	13:1755001:176 0000	15:25802001:2 5804000	15:31047627:3 1047627	21:23174621:2 3174850
	13:17470292:17 470491	16:10345001:1 0350000	15:43540924: 43540924	22:11185769:11 186071
	14:28050001: 28055050	16:22780001:2 2782000	16:22780001:2 2782000	22:12224812:1 2225254
	15:21715001:21 720000	16:46765001:4 6770000	16:44542001: 44544000	22:34108053: 34108286
	16:2807224:2 809000	16:53015001:5 3020000	16:46765344: 46766406	22:37419600: 37419816
	16:22780001:2 2782000	17:28880001: 28885000	17:7760001:77 62000	24:3782666:3 783423
	16:44542001: 44544000	18:48950001: 48955000	17:9936401:99 36898	25:35134471:3 5134811
	17:7760001:77 62000	20:1764029:17 64100	17:14458405:1 4458505	25:35502450: 35502998
	17:52099561:5 2099561	20:32100001:3 2125000	17:41075001:4 1100000	

Appendix Chapter three

	18:18042001:1 8044000	21:6775001:68 00000	17:52708001:5 2710000	
	19:21574001:21 600000	22:709001:70 9500	18:44827390: 44827390	
	20:6999077:7 002000	22:709512:70 9611	19:6310001:63 16000	
	21:201001:2013 22	22:4250001:4 255000	20:15770001:1 5775060	
	21:9780001:97 85000	24:18178001:1 8180000	20:54175001:5 4180000	
	21:25175001:2 5200000	24:37080001: 37085000	20:54390001: 54395000	
	22:3300001:3 325000	25:4420001:4 425000	22:695001:70 0000	
	22:3631001:36 32000		22:709301:70 9611	
	22:11148001:111 55000		22:20910001:2 0915000	
	22:33820001: 33825000		24:986415:98 6416	
	23:7000001:7 025000		24:3017976:3 018219	

Appendix Chapter three

	24:25350001: 25375000		24:3059536:3 059585	
	24:31942001:3 1944000			
	25:10685001:1 0690000			
	25:30800001: 30825000			
	25:35500001: 35505000			

Table S 13: Linear mixed-effects modelling of epigenetic clock performance across CpG aggregation strategies.

Linear mixed-effects models (LMMs; lmer, R) were used to analyse **A)** overall performance scores across CpG aggregation methods and **B)** performance scores as a function of CpG aggregation method and tissue.

A)

Effect	Performance Score per CpG aggregation		
	χ^2	df	p
CpG Aggregation	97	2	< 0.0001
Tile method	13.6	1	< 0.0001
Model type	8.6	1	0.003
R-squared threshold	23.8	4	< 0.001
Feature filter	48.9	3	< 0.0001
Epsize size	0.4	1	0.5
Random Effects Variance	Tissue 0.01, 0.1		

Appendix Chapter three

B)

Effect	Performance Score per cpG aggregation & tissue		
	X ²	df	p
CpG Aggregation	199	2	< 0.0001
Tissue	21.6	3	< 0.0001
Tile method	13.6	1	< 0.0002
Model type	87	1	0.007
R-squared threshold	24.8	4	< 0.0001
Feature filter	54.6	3	< 0.0001
Eps size	0.3	1	0.5
CpG Aggregation : Tissue	182.4	6	< 0.0001
Random Effects Variance	Tissue 0.01, 0.1		

Table S 14 : Effect of CpG set size on epigenetic clock performance assessed by mixed-effects modelling.

Linear mixed-effects models (LMMs; lmer, R) were used to assess the effect of the number of candidate and selected CpG sites included in clock models on the final performance score.

Effect	Performance Score per Number of Candidate and Selected CpG sites in models		
	X ²	df	p
Number Selected CpG	450.6	1	< 0.0001
Number Candidate CpG	396.3	1	< 0.0001
Random Effects Variance	Selected CpG ID : 0.002, 0.2 Chromosome : 0.00003, 0.005 Tissue : 0.002, 0.01		

Table S 15 : Distribution of Cross-Tissue Stability Classes

A) By CpG Aggregation Method: Counts and percentages of epigenetic clock models classified by cross-tissue stability. Models were categorised into four stability classes: High-performing stable, High-performing unstable, Low-performing stable, and Low-performing unstable, based on their mean rank across tissues and the variability of rank (rank SD). This panel summarises these metrics for each CpG aggregation strategy (Point, Tile, Cluster), highlighting differences in overall performance and robustness. **B)** By Tissue: Counts and percentages of models within each stability class stratified by tissue. This panel illustrates tissue-specific patterns in model performance and stability, showing how individual tissues differ in their sensitivity to clock design choices.

A)

CpG aggregation methods	High-performing (%)		Low performing (%)		Total
	Stable	Unstable	Stable	Unstable	
Cluster	14.4	20.3	32.1	33.2	355
Point	8.6	40	25.7	25.7	35
Tile	36.5	29.7	14.3	19.5	266

Appendix Chapter three

B)

Tissue	High-performing (%)		Low performing (%)		Total
	Stable	Unstable	Stable	Unstable	
Brain	24.3	14	5.3	56.4	456
Caudal Fin	6.3	5.3	16.5	71.8	419
Intestine	13	6.5	23	57.4	430
Multi-tissue	5.4	8.5	16.5	69.6	425

9.2. Supplementary Figure

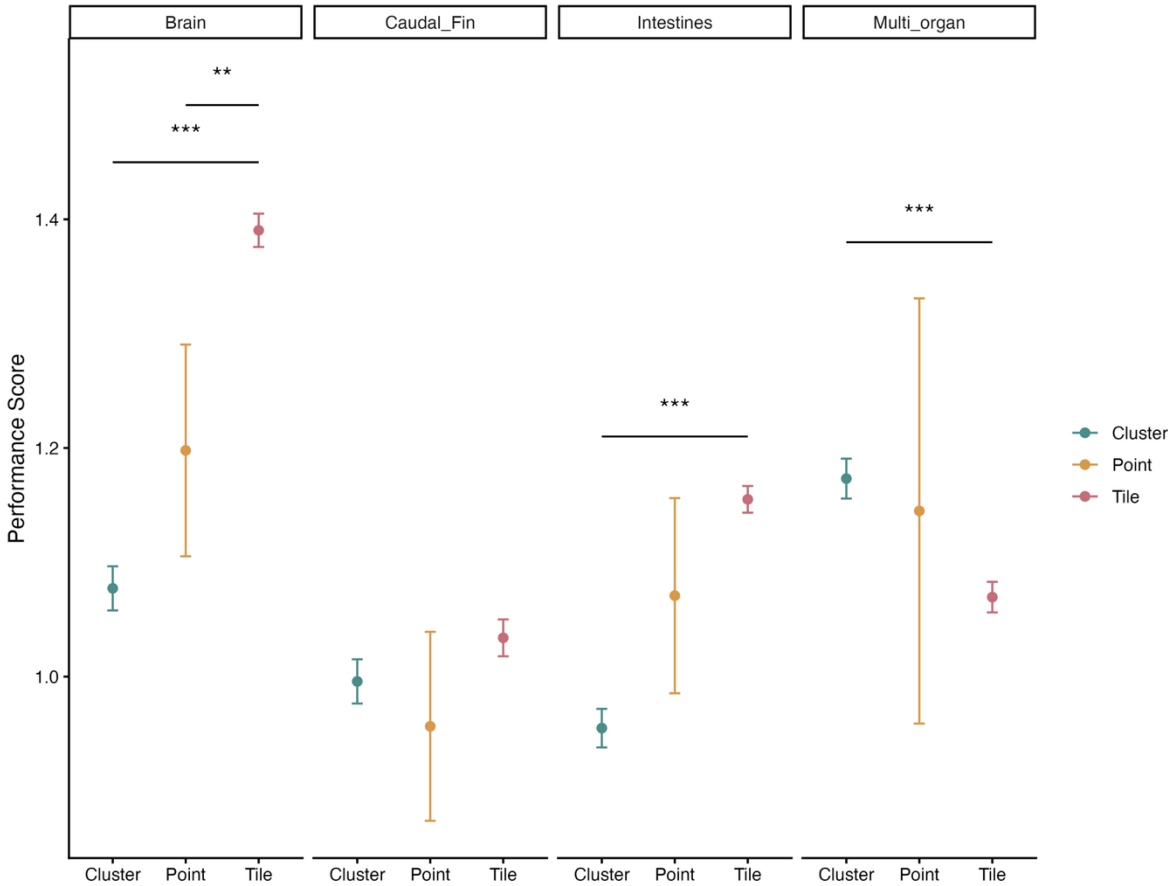


Figure S 8: Performance scores (rescaled R^2 + inverse MAE) by CpG grouping method, faceted by tissue, showing Tile dominance in brain/intestine but reversal in multi-tissue clocks.

Significance is indicated with asterisks, with * = $p < 0.05$, ** = $p < 0.001$ & *** = $p < 0.0001$. Point indicate the mean and error bars represent standard errors.

10. Appendix Chapter Four

Supporting information for Cross-tissue comparison of ageing in zebrafish using epigenetic clocks

10.1. Supplementary Tables

Table S 16: Number of samples per tissue and Sex used in Experimental Study

Age	Female			Male		
	Brain	Intestine	Caudal Fin	Brain	Intestine	Caudal Fin
3	6	6	6	7	7	5
4	2	4	4	2	2	1
5	5	3	5	6	4	7
8	2	2	2	3	5	3
10	4	4	4	2	2	2
12	2	2	2	2	2	2
14	3	3	3	3	3	1
16	3	3	2	3	3	3
18	3	3	3	2	2	2
22	6	6	6	7	7	5

Table S 17: Distribution of samples across technical and biological factors

CPG Aggregation		Regression model		R ² Threshold		Feature Filter		Eipse		Tissue	
Cluster	1312	Elastic Net	1738	No filter	315	No Filter	806	1	111	Brain	859
Point	111	Lasso	1568	0.05	731	Min 5	810	50	412	Caudal Fin	816
Tile	1883			0.1	755	Min 10	845	100	419	Intestine	824
				0.2	762	Min 20	845	500	417	Multi - Tissue	807
				0.3	743			750	149		
								1000	417		
								1500	149		
								2000	421		

Appendix Chapter Four

								3000	13 9		
								5000	27 2		
								2250 0	12 8		
								2500 0	27 2		

Table S 18: Linear mixed-effects modelling of tissue- and aggregation-dependent ageing metrics.

Linear mixed-effects models (LMMs; lmer, R) were used to analyse **A)** overall performance scores across somatic tissues and CpG aggregation methods, **B)** overall delta age across somatic tissues and CpG aggregation methods, and **C)** the relationship between overall delta age, somatic tissue, and chronological age.

A)

Effect	Performance Score per tissue and CpG aggregation		
	X ²	df	p
Tissue	85.6	3	< 0.0001
CpG Aggregation	188.5	2	< 0.0001
Tissue : CpG Aggregation	188.2	6	< 0.0001
Random Effects Variance	Individual_fish_ID: 0.002, 0.05		

B)

Effect	Delta Age per tissue and CpG aggregation		
	X ²	df	p
Tissue	35.6	3	< 0.0001
CpG Aggregation	7.9	2	0.02
Tissue : CpG Aggregation	30.2	6	< 0.0001
Random Effects Variance	Individual_fish_ID : 32.7, 5.7		

Appendix Chapter Four

C)

Effect	Delta Age per tissue and chronological age		
	X ²	df	p
Tissue	699.6	3	< 0.0001
Chronological Age	114,804.5	1	< 0.0001
Tissue : Chronological Age	1073.8	3	< 0.0001
Random Effects Variance	Individual_fish_ID : 1.9, 1.4		

Table S 19: Robustly selected tissue-specific and overlapping CpG sites identified from epigenetic clock model selection

CpG sites were classified as tissue-specific if they were selected in at least 10% of final epigenetic clock models trained within a given tissue. Overlapping CpGs represent sites selected across all somatic tissues, reflecting age-associated methylation features shared across tissues. Selection frequency was used as a stability criterion to identify CpGs with reproducible age-associated signal across modelling configurations.

Brain	Intestine	Caudal fin	Multi-tissue	Somatic overlapp
2:58868001:5 8869240	1:51950001:51 955000	1:712001:7130 00	1:24076877:2 4076977 M	6:4935441:49 35768
4:1658618:165 8618	2:3575001:36 00000	1:24077077:2 4077150	1:24077077:2 4077126	8:50628644: 50628910
5:22750001:2 2775000	2:11200001:112 25000	1:41264001:41 266000	2:34600001:3 4625000	4:51645740:5 1646027
6:14882001:14 883000	2:57232001:57 233798	2:32775001:32 800000	2:58868741:5 8869240	5:38774677:3 8775285
8:5001:12000	3:15875001:15 900000	2:57232001:57 233044	3:1105225:1105 225	8:51371072:51 371541
9:15788482:15 788482	3:23525001:23 576000	3:3650001:36 51000	3:15875001:15 900000	24:3782666:3 783423
10:38075001: 38100000	4:12928001:12 929000	3:11600001:116 25000	3:23570001:23 576000	2:44181891:44 182390

Appendix Chapter Four

11:42714001:4 2716028	4:77550001:7 7575000	3:23550001:23 576000	4:17675001:17 700000	4:3011851:301 2149
12:16564778:1 6564778	5:24200001:2 4205000	4:17675001:17 700000	4:31950001:31 975000	22:11185769:11 186071
12:34320001:3 4325000	5:43683051:4 3683100	4:34312001:34 314000	4:34312001:34 314000	13:3704291:37 04340
14:7245001:7 247000	5:43683104:4 3683105	5:39427001:3 9428000	4:42625001:4 2650229	16:815631:816 213
14:13550001:1 3575000	6:58690001:5 8696000	5:65610001:65 616000	6:64024:6402 4	15:42582708: 42583050
17:174609:179 608	7:1:25000	7:20967884:2 0970000	7:5001:10000	21:23174621:2 3174850
19:32030001: 32035000	7:18598490:18 598711	7:68415001:68 420000	8:53670663:5 3671090	1:14372733:14 373618
20:37140001: 37145000	7:21885001:21 890000	9:11340947:113 41050	13:13670001:13 672000	25:35134471:3 5134811
22:11184001:11 190000	7:32720001:32 725000	10:720001:724 000	13:14567598:1 4567598	22:34108053: 34108286
22:36300001: 36325000	7:63325001:63 330000	10:3746001:37 46500	13:51245001:51 250000	17:9936654:9 936703
23:16300001:1 6325000	8:42620001:4 2626000	11:26750001:2 6775000	14:214333:214 482	25:35502450: 35502998

Appendix Chapter Four

	10:10875001:1 0900000	11:41769001:41 770000	14:4182269:41 82401	8:31120071:311 20427
	10:28150001:2 8175000	12:33192127:33 192400	14:28052001:2 8055000	22:37419600: 37419816
	11:10828040:1 0828428	12:35875001:3 5900000	15:21200001:2 1226000	3:13557565:13 558122
	14:3570001:35 75000	13:1532001:153 4000	15:28199001:2 8199500	7:34162239:3 4162861
	15:25802001:2 5804000	13:1755001:176 0000	15:31047627:3 1047627	16:5093555:5 093769
	16:10345001:1 0350000	13:17470292:17 470491	15:43540924: 43540924	1:40644559:4 0644788
	16:22780001:2 2782000	14:28050001: 28055050	16:22780001:2 2782000	13:4766104:4 766350
	16:46765001:4 6770000	15:21715001:21 720000	16:44542001: 44544000	22:12224812:1 2225254
	16:53015001:5 3020000	16:2807224:2 809000	16:46765344: 46766406	2:6262120:62 62331
	17:28880001: 28885000	16:22780001:2 2782000	17:7760001:77 62000	4:77561200:7 7561200
	18:48950001: 48955000	16:44542001: 44544000	17:9936401:99 36898	4:77559642:7 7559642

Appendix Chapter Four

	20:1764029:17 64100	17:7760001:77 62000	17:14458405:1 4458505	4:77560396:7 7560396
	20:32100001:3 2125000	17:52099561:5 2099561	17:41075001:4 1100000	
	21:6775001:68 00000	18:18042001:1 8044000	17:52708001:5 2710000	
	22:709001:70 9500	19:21574001:21 600000	18:44827390: 44827390	
	22:709512:70 9611	20:6999077:7 002000	19:6310001:63 16000	
	22:4250001:4 255000	21:201001:2013 22	20:15770001:1 5775060	
	24:18178001:1 8180000	21:9780001:97 85000	20:54175001:5 4180000	
	24:37080001: 37085000	21:25175001:2 5200000	20:54390001: 54395000	
	25:4420001:4 425000	22:3300001:3 325000	22:695001:70 0000	
		22:3631001:36 32000	22:709301:70 9611	
		22:11148001:111 55000	22:20910001:2 0915000	

Appendix Chapter Four

		22:33820001: 33825000	24:986415:98 6416	
		23:7000001:7 025000	24:3017976:3 018219	
		24:25350001: 25375000	24:3059536:3 059585	
		24:31942001:3 1944000	24:11825001:11 850000	
		25:10685001:1 0690000		
		25:30800001: 30825000		
		25:35500001: 35505000		

Table S 20: Linear mixed-effects modelling of tissue-specific CpG methylation features.

Linear mixed-effects models (LMMs; lmer, R) were applied to tissue-specific selected CpG sites to analyse **A)** mean methylation as a function of tissue, sex, and chronological age, **B)** age-associated methylation slope by tissue, and **C)** stochastic methylation drift by tissue.

A)

Effect	Methylation per tissue and chronological age		
	X ²	df	p
Tissue	2407.4	3	< 0.0001
Chronological Age	9	9	0.4
Sex	0.2	1	0.6
Tissue : Chronological Age	14.7	27	0.9
Random Effects Variance	Individual_fish_ID : 0, 0		

Appendix Chapter Four

B)

Effect	Slopes per tissue		
	X ²	df	p
Tissue	8.9	2	<0.01
Random Effects Variance	CpG ID : 45.8, 6.8		

C)

Effect	Stochastic drift per tissue		
	X ²	df	p
Deviation	0.1	1	0.7
Tissue	9.6	2	0.008
Deviation : Tissue	5.3	2	0.06
Random Effects Variance	CpG ID : 43.6, 6.6		

Table S 21: Linear mixed-effects modelling of CpG sites shared across somatic tissues.

Linear mixed-effects models (LMMs; lmer, R) were applied to selected CpG sites overlapping all three somatic tissues to analyse **A)** mean methylation as a function of tissue, sex, and chronological age, **B)** age-associated methylation slope by tissue, **C)** stochastic methylation drift by tissue, and **D)** methylation entropy.

A)

Effect	Delta Age per tissue and chronological age		
	X ²	df	p
Tissue	0.9	3	0.6
Chronological Age	9 6.2	9	0.7
Sex	0.3	1	0.6
Tissue : Chronological Age	4.8	18	1
Random Effects Variance	Individual_fish_ID : 0, 0		

Appendix Chapter Four

B)

Effect	Slopes per tissue		
	χ^2	df	p
Tissue	2	2	0.4
Random Effects Variance	CpG ID : 4.1, 2		

C)

Effect	Stochastic drift per tissue		
	χ^2	df	p
Deviation	0.0002	1	1
Tissue	2.5	2	0.3
Deviation : Tissue	2.4	2	0.3
Random Effects Variance	CpG ID : 0,0		

Appendix Chapter Four

D)

Effect	Sample-wise methylation entropy per tissue		
	X ²	df	p
Tissue	92.7	2	<0.0001
Random Effects Variance	Sample ID : 0.0006, 0.02\		

Table S 22: Multivariate test assessing tissue separation across chronological age based on CpG sites shared across somatic tissues.

Effect	Multivariate test on tissue separation across overlapping selected CpG sites				
	df	SumofSqs	R2	F	p
Model	2	407.7	0.5	11.9	0.001

10.2. Supplementary Figures

Performance by organ × CpG grouping (Supplementary)

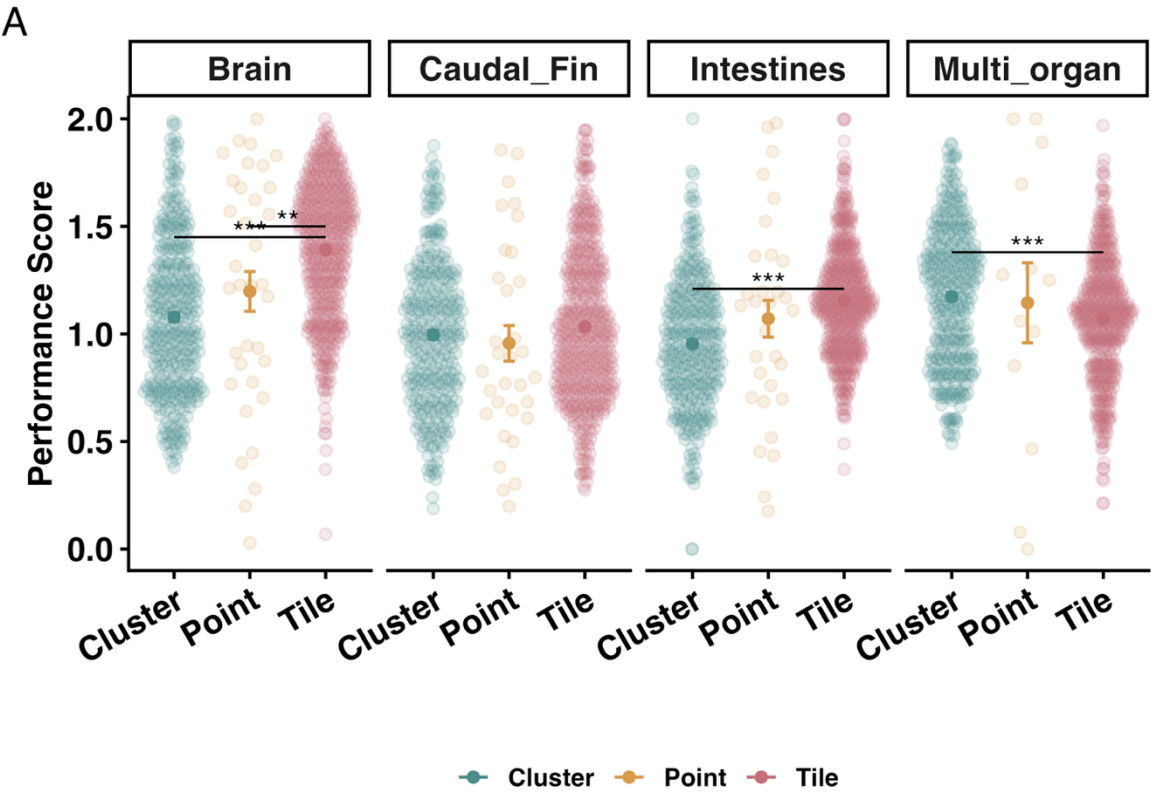


Figure S 9: Epigenetic clock performance across tissue × CpG aggregation strategies.

(A) Performance scores are shown for epigenetic clock models stratified by tissue and CpG aggregation method (Point, Cluster, and Tile). Individual model outcomes are shown as jittered points, with overlaid group means and $\pm$ standard error of the mean (SE). Panels are faceted by tissue with independent y-axis scaling. Pairwise statistical significance between CpG aggregation strategies within each tissue is indicated by asterisks.

Organ-specific CpG methylation by sex, age, and density (Supplementary)

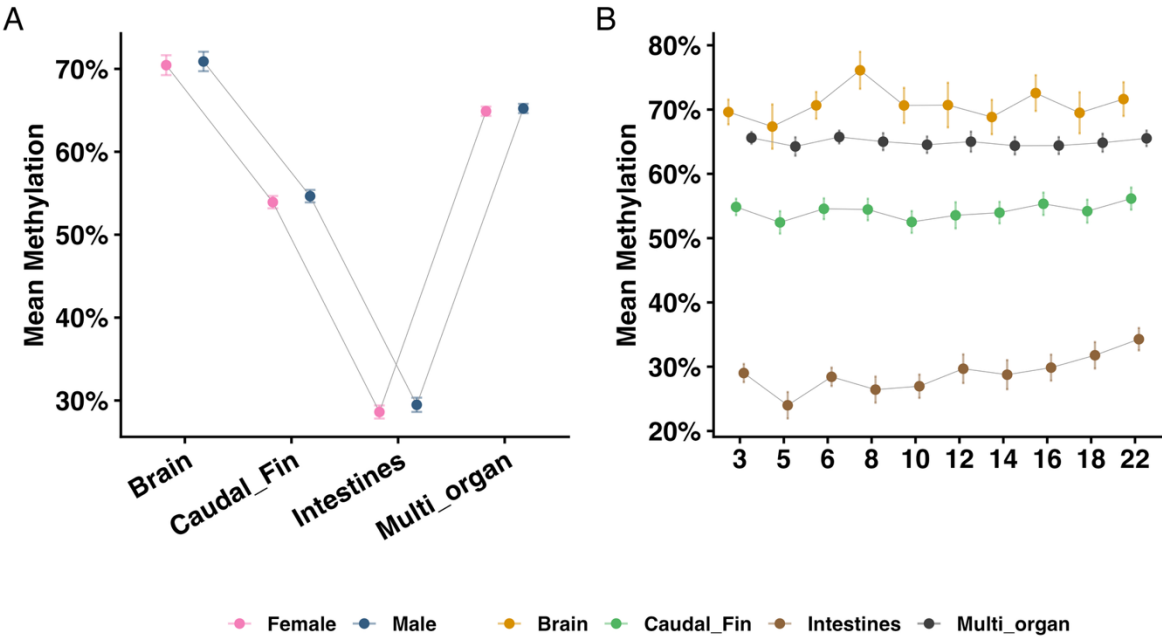


Figure S 10 : No tissue-specific modulation of CpG methylation by sex and age.

A) Mean CpG methylation stratified by tissue and sex, showing sex-specific differences in methylation levels across tissues. Points indicate group means, lines connect means within sex, and error bars represent $\pm$ standard error of the mean (SE). **B)** Mean CpG methylation across chronological age for each tissue, illustrating tissue-specific temporal trajectories. Points and lines denote tissue means with $\pm$ SE. Together, these analyses demonstrate that tissue-specific methylation patterns are not modulated by either sex and age.

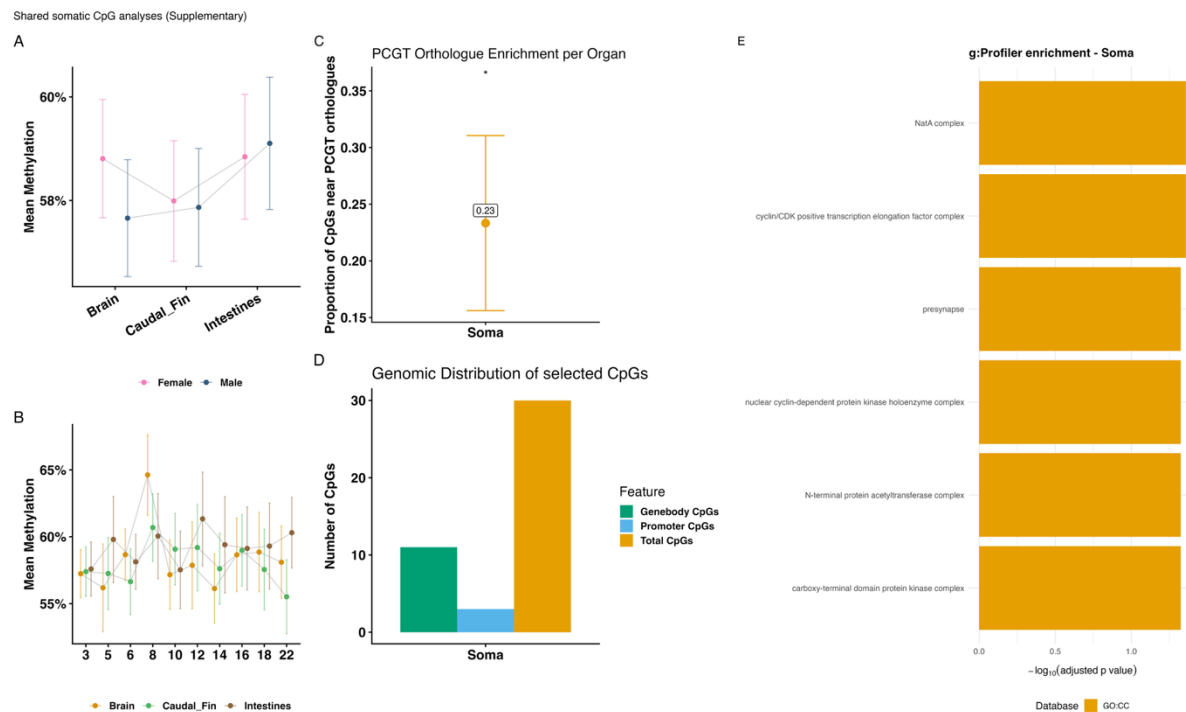


Figure S 11: Comprehensive analyses of shared somatic CpGs.

Analyses are restricted to CpGs shared across all somatic tissues. **A)** Mean methylation stratified by tissue and sex, shown as group means $\pm$ standard error of the mean (SE). **B)** Mean methylation trajectories across chronological age for each tissue, with points and lines indicating tissue means $\pm$ SE. **C)** Enrichment of shared CpGs proximal to Polycomb Group Target (PCGT) orthologues, shown as proportion $\pm$ SE with significance indicated. **D)** Genomic context of shared CpGs, partitioned into promoter, genebody, and total CpG categories. **E)** Functional enrichment analysis of genes associated with shared somatic CpGs, highlighting over-represented biological processes. Together, these analyses characterise the regulatory context and functional relevance of CpGs conserved across somatic tissues.

11. Appendix Chapter Five

Supporting information for Defining the determinants of a robust Epigenetic Clock

11.1. Supplementary Tables

Table S 23: Distribution of samples across technical and biological factors

CPG Aggregation		Regression model		R ² Threshold		Feature Filter		Epse (Window size)	
Cluster	560	Elastic Net	240	No filter	190	No Filter	239	1	38
Point	360	Lasso	278	0.05	192	Min 5	239	50	120
Tile	38			0.1	192	Min 10	240	100	120
				0.2	192	Min 20	240	500	120
				0.3	192			750	40
								1000	120
								1500	40
								2000	120
								3000	40
								5000	80

Appendix Chapter Five

								22500	40
								25000	80

Table S 24: Mixed-effects and multivariate analyses of germline-dependent epigenetic ageing metrics.

Linear mixed-effects models (LMMs; lmer, R) were used to analyse **A)** overall performance scores across CpG aggregation methods, **B)** delta age differences between germline-carrying (GLC) and germline-free (GLF) fish across chronological age, **C)** methylation differences between GLC and GLF across chronological age, **E)** hypo- and hypermethylation density between GLC and GLF across chronological age, and **F)** stochastic methylation drift between GLC and GLF. A multivariate permutation test was applied to D) selected CpG sites.

A)

Effect	Performance Score per CpG aggregation		
	X ²	df	p
CpG Aggregation	59.9	2	<0.0001
Random Effects Variance	R-squared threshold : 0.006, 0.08		
	Feature filter: 0.0008, 0.03		

B)

Appendix Chapter Five

Effect	Delta Age between GLC and GLF across chronological age		
	X ²	df	p
Germline Ablation Treatment	122.5	1	<0.0001
Chronological Age	93.3	1	<0.0001
Germline Ablation : Chronological Age	314.9	1	<0.0001
Random Effects Variance	Unique model ID : 0.3, 0.5 Chronological Age : 2.5, 1.6		

C)

Effect	Methylation between GLC and GLF across chronological age		
	X ²	df	p
Germline Ablation Treatment	0.02	1	0.9
Chronological Age	1.6	3	0.7
Germline Ablation : Chronological Age	7	3	0.06
Random Effects Variance	Chronological Age : 0.1, 0.3		

D)

Effect	Multivariate permutation test at selected CpG sites				
	df	Sum of Sqs	R ²	F	p
Germline Ablation Treatment	4	337.7	0.7	2.3	0.01

E)

Effect	Hypo/hypermethylation density between GLC and GLF across chronological age		
	X ²	df	p
Germline Ablation Treatment	0.03	1	0.9
Random Effects Variance	CpG ID : 43.8, 6.6		

F)

Effect	Stochastic drift between GLC and GLF across chronological age		
	X ²	df	p
Deviation	0.2	1	0.6
Germline Ablation Treatment	1.5	1	0.2
Deviation :Germline Ablation	1.5	1	0.2
Random Effects Variance	CpG ID: 44.5, 6.7		

Table S 25: Linear modelling of behavioural metrics.

Linear models (LMs; lm, R) were used to analyse behavioural measures including **A)** total distance travelled (mm), **B)** mean speed (mm s⁻¹), **C)** maximum speed (mm s⁻¹), **D)** mean acceleration (mm s⁻²), **E)** maximum acceleration (mm s⁻²), (F) proportion of time spent moving (%), **G)** mean angular velocity, and (H) thigmotaxis (%).

A)

Effect	Behaviour metrics : Total distance		
	Estimate	t	p
Germline Ablation Treatment - GLC	-155,309	-6.1	0.05

B)

Effect	Behaviour metrics : Mean Speed		
	Estimate	t	p
Germline Ablation Treatment- GLC	-58.8	6	0.1

C)

Effect	Behaviour metrics : Max Speed		
	Estimate	t	p
Germline Ablation Treatment- GLC	-5,149	-2.6	0.01

Appendix Chapter Five

D)

Effect	Behaviour metrics : Mean Acceleration		
	Estimate	t	p
Germline Ablation Treatment- GLC	-931.6	-1.2	0.3

E)

Effect	Behaviour metrics : Max Acceleration		
	Estimate	t	p
Germline Ablation Treatment- GLC	-138,672	-2.4	0.02

F)

Effect	Behaviour metrics : Time Moving		
	Estimate	t	p
Germline Ablation Treatment- GLC	0.3	2.1	0.04

G)

Effect	Behaviour metrics : Mean Angular Velocity		
	Estimate	t	p
Germline Ablation Treatment-GLC	256	0.7	0.5

H)

Effect	Behaviour metrics : Thigmotaxis		
	Estimate	t	p
Germline Ablation Treatment- GLC	14.7	1.6	0.1

11.2. Supplementary Figures

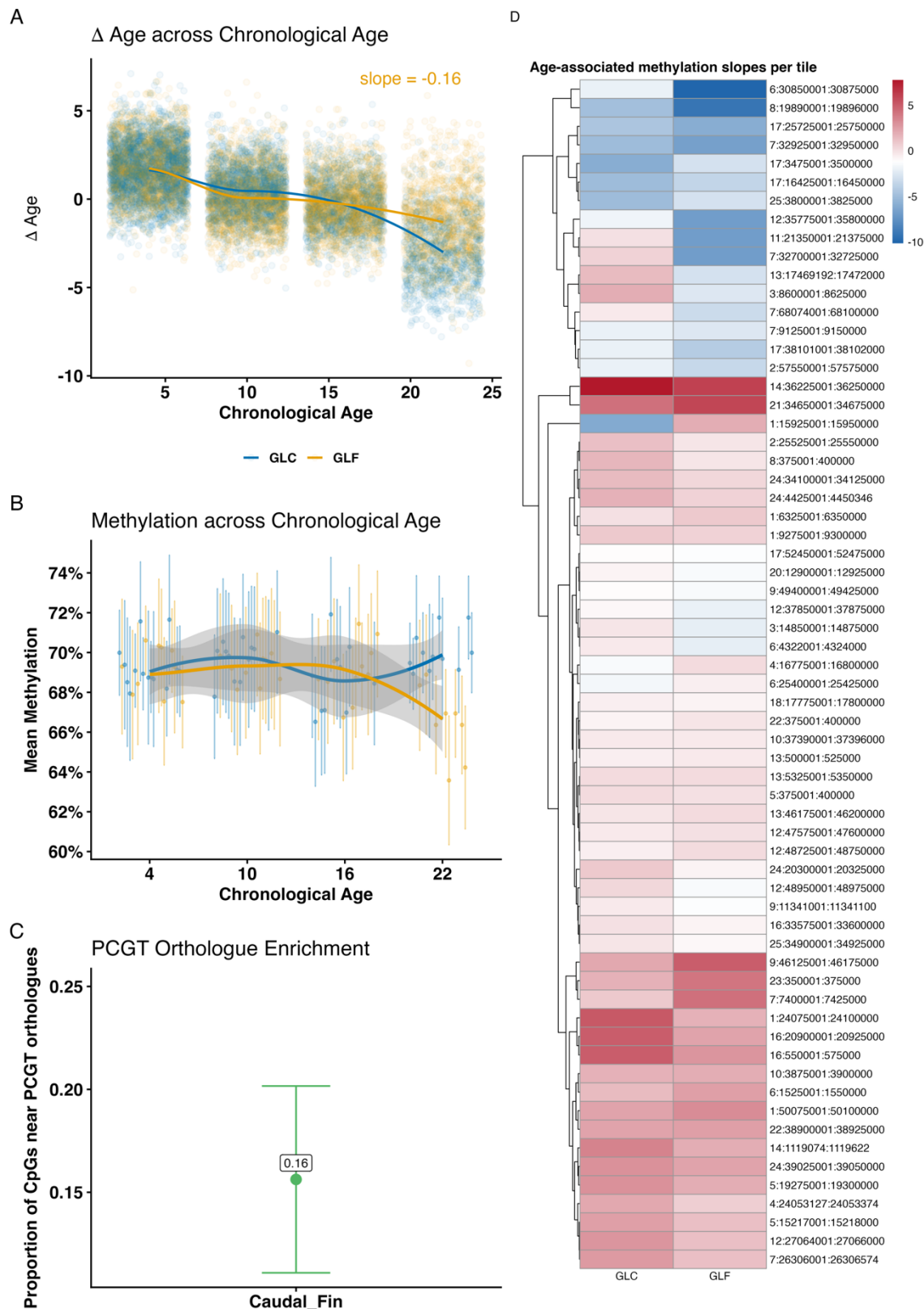


Figure S 12: Epigenetic clock performance by CpG grouping.

A) Δ Mean epigenetic age deviation (Δ -Age), defined as the difference between predicted epigenetic age and chronological age stratified by treatment, with LOESS smoothing and individual data points overlaid; slope values are indicated for each treatment. **B)** Mean methylation across chronological age per individual, highlighting treatment-specific trajectories. **C)** Proportion of selected CpGs overlapping Polycomb group target (PCGT) orthologues, with error bars representing standard error; significance is denoted above each point (** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$). **D)** Heatmap of age-associated methylation slopes per tile, clustered by similarity across CpG tiles. Data are shown for a single tissue (caudal fin) and stratified by experimental treatment.

12. Appendix Chapter Six

Supporting information for Rapamycin accelerates epigenetic age ,alters behaviour and ovarian lesions trajectories In zebrafish (Danio rerio)

12.1. Supplementary Tables

Table S 26: Experiment 1 - fishers test on sex-specific breeding outcomes by treatment 2 months after start of rapamycin treatment.

Sex	Rapamycin Treatment	Number Bred	Number non-Bred	Proportion Bred	Fish p-value (vs Control)
Female	Control	1	19	0.05	-
Female	Low	6	14	0.30	0.1
Female	High	2	18	0.10	0.1
Male	Control	7	11	0.39	-
Male	Low	1	15	0.06	0.001
Male	High	0	20	0	0.001

Table S 27: Distribution of samples used for epigenetic clock construction

This table summarises the number of samples used for epigenetic clock analyses in Experiment 1 of Chapter 6, stratified by treatment, sex, and time point. Only samples passing sequencing, alignment, and coverage quality control were included

Timepoint	Control		Low		High	
	Female	Male	Female	Male	Female	Male
1	4	4	4	4	4	4
2	4	4	4	4	4	4

Table S 28: Mixed-effects modelling of behavioural metrics across rapamycin treatment conditions in aged fish.

Linear mixed-effects models (LMMs; lmer, R) were fitted separately to each behavioural metric to assess differences between rapamycin treatment conditions in the aged experiment.

Behaviour Metric	Effect	Breeding success between Rapamycin conditions and Sex		
		X ²	df	p
Total Distance (mm)	Rapamycin Treatment	171.2	2	< 0.0001
	Random Effects Variance	Video Minute :792,080, 290 Fish ID :10,746, 103.7 Block : 5,372, 73.2 Timepoint : 1,430,747, 1,196.1		
Mean Speed (mm / s)	Rapamycin Treatment	4.4	2	0.1
	Random Effects Variance	Fish ID :10,746, 103.7		
Max Speed (mm / s)	Rapamycin Treatment	102.5	2	< 0.0001

Appendix Chapter Six

	Random Effects Variance	Video Minute :6,978,83.5 Fish ID :9,856,99.3 Block : 1,780, 42.2 Timepoint : 11,300, 106.3		
Mean Acceleration (mm / s ²)	Rapamycin Treatment	4	2	0.1
	Random Effects Variance	Fish ID :126,11.2		
Max Acceleration (mm / s ²)	Rapamycin Treatment	102.5	2	< 0.0001
	Random Effects Variance	Video Minute :6,280,315 , 2,506 Fish ID :8,947,938 , 2,991 Block : 1,605,225 , 1,267 Timepoint : 10,208,043 , 19,296		
Time moving (%)	Rapamycin Treatment	217.6	2	< 0.0001
	Random Effects Variance	Fish ID :0.02,0.2 Block : 0.008,0.09 Timepoint : 0.02,0.1		
Thigmotaxis (%)	Rapamycin Treatment	249.9	2	< 0.0001

	Random Effects Variance	Video Minute :11.1 , 3.3 Fish ID :11.2 , 3.3 Block : 8.3 , 2.9 Timepoint : 33.2 , 5.8		
Mean Angular Velocity	Rapamycin Treatment	352.4	2	< 0.0001
	Random Effects Variance	Fish ID :11,504 , 107.3 Block : 1,449 , 38.1 Timepoint : 17,740 , 133.2		

Table S 29 : Mixed-effects and non-linear modelling of epigenetic ageing metrics.

Linear mixed-effects models (LMMs; lmer, R) were used to analyse **A)** treatment-associated changes in delta age, **B)** treatment-specific age-associated slopes, **C)** tissue heterogeneity in epigenetic ageing, and **E)** sex-specific treatment effects on delta age. Non-linear ageing architecture was assessed using generalised additive models (GAMs) in **D).**

A)

Effect	Delta Age differences between Rapamycin Conditions across chronological age		
	X ²	df	p
Rapamycin Conditions	553.3	2	< 0.0001
Chronological Age	16,468.8	1	< 0.0001
Rapamycin Timepoint	810.7	1	< 0.0001
Rapamycin Conditions : Chronological Age	506.5	2	< 0.0001
Random Effects Variance	Unique model ID : 135.9, 11.7 CPG Aggregation : 7/8, 2.8		

B)

Treatme nt	slope (β)	SE	95% CI	Comparis on	Δ slope	SE	z	p- value
Control	-1.123	0.0088	[-1.140 , -1.106]	Control - Low	-0.088	0.01	-8.52	<0.0001
Low	-1.035	0.0076	[-1.050, -1.020]	Control - High	-0.229	0.01	-21.95	<0.0001
High	-0.894	0.0078	[-0.909, -0.878]	Low - High	-0.141	0.0096	-14.78	<0.0001

Appendix Chapter Six

C)

Effect	Delta Age differences between Rapamycin Conditions across chronological age		
	χ^2	df	p
Rapamycin Conditions	3.8	2	0.1
Rapamycin Timepoint	1.6	1	0.2
Rapamycin Conditions : Rapamycin Timepoint	3	2	0.2
Random Effects Variance	Chronological Age : 0.003, 0.05		

D)

Rapamycin Treatment	EDF	p-value
Control	1.41	<0.001
Low	1.98	<0.001
High	2	<0.001

E)

Effect	Delta Age differences between Rapamycin Conditions and Sex		
	X ²	df	p
Rapamycin Conditions	280	2	< 0.0001
Sex	2	1	0.1
Rapamycin Conditions : Sex	1,093.9	2	< 0.0001
Random Effects Variance	Unique Model ID : 138.9,11.8 Chronological Age : 4.1,2 Rapamycin Timepoint : 0.09, 0.3		

Table S 30 : Generalised linear mixed-effects modelling of ovarian lesion incidence.

Generalised linear mixed-effects models (GLMMs; glmer, R) were used to analyse **A)** the proportion of fish with at least one ovarian lesion across genotypes and **B)** the proportion of fish with multiple ovarian lesions.

A)

Effect	Proportion of fish with at least one ovarian lesion		
	X ²	df	p
Rapamycin Conditions	0.1	2	0.9
Genotype	2.2	1	0.1
Rapamycin Conditions : Genotype	2.9	2	0.2
Random Effects Variance	Weight : 0.4, 0.6 Tank : 0.6, 0.8		
Pairwise contrasts (Tukey adjusted)	Estimate	z value	p
WT : Control – Low	0.035	0.859	0.999
WT : Control – High	-0.246	0.854	0.955
WT : Low - High	-0.281	0.858	0.943
Sa22101 : Control – Low	-1.269	0.97	0.39

Appendix Chapter Six

Sa22101 : Control – High	-2.409	0.966	0.034
Sa22101 : Low - High	-1.14	0.875	0.393

B)

Effect	Proportion of fish with multiple ovarian lesions		
	X ²	df	p
Rapamycin Conditions	3	2	0.2
Genotype	0.2	1	0.7
Rapamycin Conditions : Genotype	0.7	2	0.7
Random Effects Variance	Weight : 3.6 e -9 , 6 e – 5 Tank : 0.4 , 0.6		
Pairwise contrasts (Tukey adjusted)	Estimate	z value	p
WT : Control – Low	-1.365	-1.464	0.3082
WT : Control – High	-1.517	-1.639	0.2293
WT : Low - High	-0.152	-0.191	0.9801
Sa22101 : Control – Low	-1.698	-1.648	0.2256
Sa22101 : Control – High	-2.543	-2.536	0.0301

Appendix Chapter Six

Sa22101 : Low - High	-0.846	-1.096	0.5168
----------------------	--------	--------	--------

Table S 31 : Linear effect model (lmer in R) for Behaviour metrics in the Lifelong experiment

Behaviour Metric	Effect	Breeding success between Rapamycin conditions and Sex		
		X ²	df	p
Total Distance (mm)	Rapamycin Treatment	216.5	2	< 0.0001
	Random Effects Variance	Video Minute : 941, 950 ; 970.5 Fish ID : 18,534 ; 136.1 Rapamycin Timepoint : 1,024,417 ; 1,012.1		
Mean Speed (mm / s)	Rapamycin Treatment	9.3	2	0.009
	Random Effects Variance	Video Minute : 0.6 ; 0.8 Rapamycin Timepoint : 3.1 ; 1.8		
Max Speed (mm / s)	Rapamycin Treatment	128.2	2	< 0.0001
	Random Effects Variance	Video Minute : 53,250 ; 230.8 Fish ID : 2,167 ; 46.5 Rapamycin Timepoint : 762, 605 ; 873.3		

Appendix Chapter Six

Mean Acceleration (mm / s ²)	Rapamycin Treatment	6.2	2	0.04
	Random Effects Variance	Video Minute : 336.5 ; 18.3 Rapamycin Timepoint : 3. 583.1 ; 59.9		
Max Acceleration (mm / s ²)	Rapamycin Treatment	129.5	2	< 0.0001
	Random Effects Variance	Video Minute : 4.8 e + 7 ; 6,904 Fish ID : 2 e + 6 ; 1, 397 Rapamycin Timepoint : 6.7 e + 8 ; 39,517		
Time moving (%)	Rapamycin Treatment	11.8	2	0.002
	Random Effects Variance	Video Minute : 0.007 ; 0.09 Fish ID : 0.006 ; 0.08 Rapamycin Timepoint : 5.2 ; 2.3		
Thigmotaxis (%)	Rapamycin Treatment	5.2	2	0.02
	Random Effects Variance	Video Minute : 0.6; 0.8 Fish ID : 0.3 ; 0.6 Rapamycin Timepoint : 35.8 ; 8.9		

Appendix Chapter Six

Mean Angular Velocity	Rapamycin Treatment	213.2	2	< 0.0001
	Random Effects Variance	Video Minute : 390.8 ; 19.8 Fish ID : 200.2 ; 14.2 Rapamycin Timepoint : 65, 974 ; 256.9		

Table S 32 : Linear mixed-effects modelling of vertical exploration behaviour in the Novel Tank Diving Test.

Linear mixed-effects models (LMMs; lmer, R) were used to analyse **A)** the percentage of time fish spent in the upper zone of the tank and **B)** the percentage of time spent in the lower zone of the tank during the Novel Tank Diving Test (NTDT).

A)

Effect	Percentage of time spent at top of tank		
	X ²	df	p
Rapamycin Conditions	1.8638	2	0.39381
Genotype	1.5466	1	0.21363
Age	50326.0161	1	< 0.0001
Rapamycin Conditions : Genotype	1.6239	2	0.44399
Rapamycin Conditions : Age	12555.8712	2	< 0.0001
Genotype : Age	12119.4132	1	< 0.0001
Rapamycin Conditions : Genotype : Age	32996.0709	2	< 0.0001
Random Effects Variance	Tank : 0.2; 0.5 Fish ID : 0.05 ; 0.2 Video minute : 0.2 ; 0.4		

Appendix Chapter Six

B)

Effect	Percentage of time spent at frozen		
	X ²	df	p
Rapamycin Conditions	0.513	2	0.773771
Genotype	0.8429	1	0.358562
Age	88466.9661	1	< 0.0001
Rapamycin Conditions : Genotype	0.0239	2	0.988116
Rapamycin Conditions : Age	45873.4166	2	< 0.0001
Genotype : Age	29454.8347	1	< 0.0001
Rapamycin Conditions : Genotype : Age	45423.5062	2	< 0.0001
Random Effects Variance	Tank : 0.2 ; 0.5 Fish ID : 0.4 ; 0.6 Video minute : 0.2 ; 0.5		

12.2. Supplementary Figures

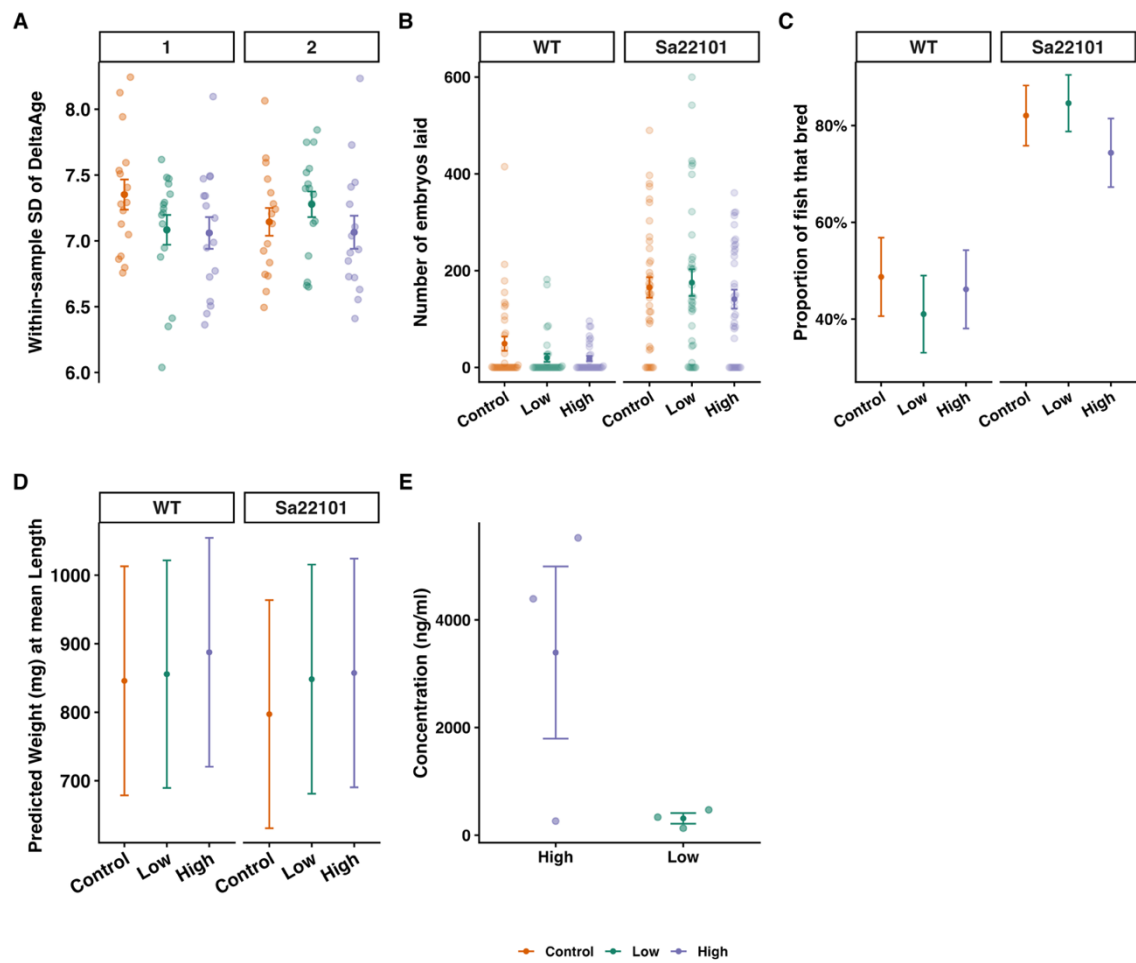


Figure S 13 : Rapamycin treatment does not significantly affect within-sample variability, reproductive output, or predicted growth in the lifelong experiment

A) Within-sample variability in epigenetic age, quantified as the median absolute deviation (MAD) of ΔAge (predicted – chronological age) across tissues within individual fish, shown across rapamycin treatments and timepoints in Aged Experiment. No significant treatment effects were detected, **B)** Number of embryos laid per breeding event across rapamycin treatments in the lifelong experiment. No significant differences were observed between treatments, **C)** Breeding success (binary outcome: production of viable embryos) across rapamycin treatments in the lifelong experiment. No significant treatment effects were detected, **D)** Predicted body weight from LMM model across rapamycin treatments in the lifelong experiment. No significant differences were observed between treatments, and **E)** whole-blood rapamycin concentrations for Low- and High-dose treatments are shown for reference; although statistical testing was not performed due to limited sample size, High-dose fish exhibited markedly higher circulating rapamycin levels than Low-dose fish. Data are shown as individual observations with mean $\pm$ s.e.m. overlaid.

