Age-related nuclear mutation profiling across somatic and germline tissues in MutaMouse males

Elena Esina

Thesis submitted to the University of Ottawa

in partial Fulfillment of the requirements for the

Master of Science in Biology

With a Specialization in Bioinformatics

Department of Biology

Faculty of Science

University of Ottawa

ABSTRACT

Aging is associated with progressive declines in tissue function and increased susceptibility to chronic disease, but the molecular changes that accumulate across tissues during aging remain incompletely understood. Nuclear mutations are permanent genomic changes that can provide a cumulative record of DNA damage, repair, replication errors, and cellular population dynamics experienced by tissues over time. However, age-related nuclear mutagenesis remains difficult to characterize because spontaneous mutations are rare and existing evidence is fragmented across methods, species, tissues, and genomic compartments. Duplex Sequencing (DS) enables high-accuracy detection of rare mutations in endogenous DNA, but accurate mutation detection alone is not sufficient for biological interpretation. The overarching objective of this thesis was therefore to determine whether nuclear mutation frequency (MF) and mutation spectra change with age across MutaMouse male tissues in a tissue-dependent manner.

First, a simulation-based framework was developed to evaluate how DS study-design parameters influence the ability to detect differences in MF. Using empirical liver and bone marrow data from adult MutaMouse males as reference conditions, the analyses examined the effects of baseline MF, inter-animal variability, expected effect size, sample size, and informative duplex sequencing depth. As expected, power increased with larger sample sizes, greater sequencing depth, and larger expected effect sizes, but decreased as inter-animal variability increased. Increasing sequencing depth improved power only up to a point, after which additional informative duplex bases provided limited benefit relative to increasing the number of animals per group. Under the modeled conditions, sample sizes of 4–5 animals per group were sufficient to detect at least a 1.5-fold change in MF with feasible sequencing depth, whereas routine detection of smaller changes remained challenging.

Second, DS was used to characterize endogenous nuclear mutagenesis across liver, lung, bone marrow, and male germ cells from MutaMouse males spanning early to late life. These tissues were selected to represent distinct biological contexts, including differences in proliferative activity, metabolic demand, environmental exposure, and somatic versus germline

identity. Early in life, MF was broadly similar across tissues. Across the full age range, however, tissue-specific trajectories emerged. MF increased with age in all three somatic tissues, with the unique-mutation estimate, MFMin, rising approximately 1.5–2.0-fold between 3 weeks and ≥78 weeks. Liver and lung showed gradual age-related accumulation, whereas bone marrow showed a delayed increase that became most apparent in late life. In contrast, male germ cells did not show a significant age-related increase in MF over the ages examined. Mutation spectra and genomic-context analyses further showed that tissue identity shaped broader mutational patterns, although higher-resolution signature-level analyses were limited by sparse mutation counts after stratification.

Overall, this thesis shows that age-related nuclear mutation accumulation in MutaMouse males is tissue-dependent rather than uniform across the organism. Somatic tissues showed detectable age-related increases in MF, whereas male germ cells did not show a significant increase over the ages examined. The simulation-based framework developed here further shows that high-accuracy mutation detection is most informative when paired with sufficient statistical power, especially when expected effects are modest or rare mutations are subdivided for higher-resolution analyses. These findings show that age-related nuclear mutagenesis depends on the tissue being studied and provide practical guidance for designing DS-based studies that can distinguish biological patterns from limitations of study design.

ACKNOWLEDGEMENTS

They say it takes a village, and this thesis certainly did. This work would not have been possible without the support of many people who helped me along the way.

I am deeply grateful to my supervisors, Dr. Carole Yauk and Dr. Francesco Marchetti, for guiding this work and supporting me throughout my degree. Thank you for pushing me to think more carefully, write more clearly, and trust myself more as a scientist. I am especially grateful for the time you invested in this project and for your help during the more difficult parts of graduate school.

I would also like to thank my committee members, Dr. John Stead and Dr. Stéphane Aris-Brosou, for their thoughtful feedback during our meetings. Your comments helped strengthen this thesis and encouraged me to see my work from new perspectives.

Thank you to everyone in the GReAT lab and at Health Canada who helped make this project possible. I am especially grateful to Annette Dodge for introducing me to Dr. Yauk and Dr. Marchetti and inspiring me to pursue a Master's degree with them, and to Alper James Alcaraz for the many hours spent mentoring and training me in the lab. A big thank you also to Andrew Williams for sharing his statistical expertise and for his patience with my many questions.

To my friends and fellow students, thank you for making graduate school feel less lonely. I am grateful for every conversation and message that helped me keep going.

Finally, thank you to my family for believing in me through every stage of this degree. Your love carried me through the hardest parts, and I am so grateful.

Table of Contents

LIST OF TABLES

Chapter One

Chapter Two

Chapter Three

Supplementary Materials

LIST OF FIGURES

Chapter Two

Chapter Three

Supplementary Materials

LIST OF ABBREVIATIONS

BM bone marrow

bp base pair

COPD chronic obstructive pulmonary disease

DCS duplex consensus sequence

DS duplex sequencing

DSB double stranded break

ecNGS error-corrected next-generation sequencing

ECS error-corrected sequencing

GLMM generalized linear mixed model

HSC haematopoietic stem cell

indel insertion and deletion

MF mutation frequency

MF_{Max} mutation frequency maximum

MF_{Min} mutation frequency minimum

MNV multi-nucleotide variants

mtDNA mitochondrial DNA

NGS next-generation sequencing

PCR polymerase chain reaction

ROS reactive oxygen species

SNV single-nucleotide variants

SSC spermatogonial stem cell

SSCS single-strand consensus sequence

ST seminiferous tubules

SV structural variant

TGR transgenic rodent

UMI unique molecular identifier

VAF variant allele frequency

VC vehicle control

STATEMENT OF CONTRIBUTION

Chapter 2: **Power analyses to inform Duplex Sequencing study designs for MutaMouse liver and bone marrow**

This chapter has been published in Environmental and Molecular Mutagenesis in Oct 2024.

Authors:

Elena Esina [1], Annette E Dodge [2], Andrew Williams [2], David M Schuster [1], Danielle P M LeBlanc [2], Francesco Marchetti [2 3], Carole L Yauk [1]

Affiliations:

[1] Department of Biology, University of Ottawa, Ottawa, Ontario, Canada.

[2] Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada.

[3] Department of Biology, Carleton University, Ottawa, Ontario, Canada.

Study Concept and Design……. Elena Esina
Carole L. Yauk
Francesco Marchetti

Data Resources and Curation…….. Danielle P M LeBlanc
David M Schuster
Annette E Dodge

Data Analysis and Bioinformatics Processing…… Elena Esina
Andrew Williams
Annette E Dodge

Manuscript Preparation …….. Elena Esina
Carole L. Yauk
Francesco Marchetti
Andrew Williams

Chapter 3: **Age-related nuclear mutation profiling across somatic and germline tissues in MutaMouse males**

Authors:

Elena Esina[1], Alper James Alcaraz[1], Andrew Williams[2], David M Schuster[1], Madison Stewart[1,2], Stephanie Smith-Roe[4], Francesco Marchetti[2,3], Carole L. Yauk[1]

Affiliations:

[1] Department of Biology, University of Ottawa, Ottawa, Ontario, Canada.

[2] Environmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada.

[3] Department of Biology, Carleton University, Ottawa, Ontario, Canada.

[4] Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA.

Study Concept and Design……Elena Esina
Carole L. Yauk
Francesco Marchetti

Laboratory Work…….. Elena Esina
Alper James Alcaraz

Data Resources and Curation…….. Danielle P M LeBlanc
David M Schuster
Madison Stewart
Stephanie Smith-Roe

Bioinformatics and Data Analysis……Elena Esina
Andrew Williams
Alper James Alcaraz

Manuscript Preparation …….. Elena Esina
Carole L. Yauk
Francesco Marchetti

CHAPTER ONE

INTRODUCTION
Elena Esina

1.1 Aging, health span, and disease burden

Advances in sanitation, housing, nutrition, and medical care have roughly doubled average life expectancy in most high-income countries over the past two centuries (Partridge et al., 2018). As mortality declines, populations worldwide are experiencing substantial aging, a trend that is projected to continue in the coming decades (*Ageing and Health*, 2025). However, living longer has not necessarily meant living healthier (i.e., with a longer health span), and the burden of chronic disease in later life remains high (Bell et al., 2019; Partridge et al., 2018).

With advancing age, major degenerative pathologies can develop, including cancers, cardiovascular disorders, musculoskeletal and endocrine disorders, and neurodegenerative diseases (Erickson, 2003; Jackson et al., 2018; López-Otín et al., 2013). These conditions not only reduce quality of life but also impose substantial economic, healthcare, and social burdens. The central challenge is, therefore, to ensure that the additional years of life gained are not dominated by disease or poor health, but instead correspond to an extension of health span (Melzer et al., 2020). Understanding how and why tissues age is a fundamental step towards meeting this challenge and developing interventions that preserve tissue and organ function into later stages of life.

1.2 Mutagenesis and biological aging

Age can be described in two distinct ways: chronological age, denoting the number of calendar years a living system has existed, and biological age, reflecting physiological or functional aging. Chronological age is strongly linked with deteriorating health, but poorly reflects internal biological processes (Chen et al., 2023). As such, the scientific community has shifted its attention from chronological signs of aging to biological ones (Chen et al., 2023).

Despite increasing research on biological markers of aging, there is surprisingly little consensus on how to define aging. For example, a recent survey showed that experts attending the 2022 Systems Aging Gordon Research Conference disagree on fundamental questions such as what aging is, what causes it, and when it begins, with no single answer to any of these questions receiving majority support (Gladyshev et al., 2024). Free-text definitions of aging

were grouped into several broad categories. The most common responses described aging as a loss of function over time (~30% of respondents). Other frequently-mentioned themes included the accumulation of damage or other detrimental changes over time, viewing aging as multifaceted, and describing it as a systemic decline in health. Although these categories emphasize different aspects of aging, the authors noted some shared tendencies across responses even if they disagreed strongly on definitions and mechanisms; most researchers viewed aging as a time-dependent process that is inherently deleterious and involves the accumulation of harmful changes, damage, degeneration, and loss of function (Gladyshev et al., 2024). Consistent with this view, for the purposes of this thesis, biological aging will be defined as a gradual, time-dependent deterioration of biological function, marked by accumulating damage and other deleterious changes that progressively reduce resilience and increase susceptibility to age-related disease and mortality (Gladyshev et al., 2024; López-Otín et al., 2013).

Within this framework, aging is associated with multiple molecular and cellular changes. Factors include mutation accumulation, telomere shortening, gene expression alterations, mitochondrial dysfunction, changes in the proteome, and shifts in epigenetic patterns (Jones et al., 2015; Leuthner & Meyer, 2021; Oh et al., 2023). Among these processes, mutation accumulation has emerged as a measurable indicator of biological age. Evidence for age-associated mutation accumulation and the methods used to detect these changes are summarized in Tables 1.1 and 1.2. Mutations are defined as irreversible changes in the DNA sequence. Mutation burden can be used as a biomarker that captures the long-term integration of DNA damage and repair outcomes across an organism's life. Mutational endpoints include single- and multi-nucleotide variants (SNVs and MNVs), small insertions and deletions (indels) of less than 1000bp, and larger structural variants (SVs) in nuclear DNA. At the chromosome level, aging-related genome instability can also include somatic aneuploidy, in which individual cells gain or lose whole chromosomes. Although distinct from the sequence-level mutation endpoints emphasized in this thesis, aneuploidy is relevant because chromosome gains or losses can induce cellular stress and have been linked to aging-associated tissue dysfunction and disease (Ricke & van Deursen, 2013; Zhu et al., 2018). Nuclear mutation accumulation can

therefore be considered one quantifiable manifestation of biological aging, particularly when measured at high resolution in specific tissues or cell types.

Age-associated increases in mutation burden can reflect both *de novo* mutagenesis and the clonal expansion of cells carrying pre-existing mutations. *De novo* mutagenesis refers to the formation of new variants in the DNA sequence of individual cells through replication-associated errors or the misrepair of DNA lesions (Vijg, 2021). In contrast, clonal expansion does not necessarily require a higher rate of new mutation formation; instead, a mutation that arose earlier may become more abundant because the cell carrying it, or its descendants, proliferate or persist within the tissue (Martincorena, 2019). As a result, the mutations detected in a tissue sample reflect both the history of mutational processes and the population dynamics of mutant cell lineages, including stochastic changes in clone size through neutral genetic drift and selective expansion of clones carrying advantageous mutations (Martincorena, 2019; Yizhak et al., 2019). Distinguishing these processes is important because an increase in measured mutation burden with age may indicate the accumulation of independent mutational events, expansion of mutant clones, or both.

Mutations can arise naturally from a variety of sources over the course of an individual's lifespan and have consequences for tissue function and disease risk, as well as offspring health (Shelby, 1994; Vijg, 2021). For example, errors in DNA replication that evade DNA polymerase proofreading and post-replicative mismatch repair can introduce mutations within coding sequences, where single-nucleotide substitutions may be classified by their effects on the encoded protein as synonymous, missense, or nonsense mutations, and indels can disrupt the reading frame (Brown, 2002a). Exposure to endogenous genotoxic stressors, such as reactive oxygen species (ROS), also contributes to mutagenesis; ROS is implicated in 100 different oxidative base lesions and modifications (Durland & Ahmadian-Moghadam, 2023). Even naturally occurring processes like spontaneous base hydrolysis, DNA methylation, and DNA glycosylase cleavage can result in mutagenesis (Durland & Ahmadian-Moghadam, 2023).

Table 1.1: Select studies of age-related mutation accumulation in somatic tissues of mouse models.

Reference	Method of Mutation Detection	Tissues of Interest	Age Groups (months)	Key Findings	
				Mutation Frequency	Mutation Spectra & Landscape
Ono et al., 2000	MutaMouse *lacZ* transgenic reporter assay; rescue of λ shuttle vector and sequencing at 0 and 23 months of *lacZ* and flanking regions in recovered mutants	spleen, liver, heart, brain, back skin, testis	0 2 6 12 23	• MFs are similar at birth; increase ~linearly with age in all tissues with tissue-specific slopes • Tissue-specific mutation accumulation is inconsistent with cell division rates • Mutations consistent with DNA replication errors, DNA methylation, DNA damage	• Point mutations prevalent in all tissues, predominantly C:G>T:A transitions at CpG sites; infrequent indels (mostly 1-20 nts) in repeat tracts • Rare complex mutations occur only in old tissues
Turker et al., 2007	Endogenous Aprt autosomal reporter; in vitro selection for DAP-resistant Aprt^{-} clones, followed by LOH mapping at chr8 microsatellites to classify mutation types	kidney epithelial, ear mesenchymal fibroblasts	2–8 16–20 26–35	• Aprt MFs increase with age in kidneys of both sexes and in female ear fibroblasts; no significant age-related increase is detected in male ear fibroblasts • Differences in mutation accumulation are gender- and tissue-specific	• Mutation spectra are tissue-specific: kidney spectra enriched for chromosome loss, ear spectra enriched for mitotic recombination and interstitial deletions • Mutational spectra show only small age-related trends and no major shifts with age
Fischer & Stringer, 2008	G11 PLAP mononucleotide repeat frameshift reporter; in situ histochemical detection of PLAP^{+} cells	heart, brain cerebrum and cerebellum, kidney cortex and medulla	0–5 6–16 17–24	• Number of PLAP^{+} mutant cells increases ~linearly with age in heart, brain and kidney (cortex and medulla) • Mutant-cell accumulation occurs at a relatively steady rate across age; there's no evidence for a late-life acceleration or deceleration • Mutant cells accumulate in both proliferative and nominally non-proliferative cells	*(spectrum/genomic landscape not reported by the assay)*
Kimoto et al., 2017	EGFP homologous recombination reporter, fluorescent imaging of lung sections	lung	7–9 wks 8–15	• Mutant (EGFP^{+}) lung cells are rare in young mice but much more common in older mice (~an order-of-magnitude increase) • Older lungs also contain many more small clusters of neighbouring EGFP^{+} cells, indicating more clonal expansion of mutant cells with age	*(spectrum/genomic landscape not reported by the assay)*
Dollé et al., 1997	*lacZ* plasmid-based reporter; plasmid rescue and positive selection for *lacZ* mutants, PCR/restriction and hybridization; sequencing of selected rearrangements	brain, liver	1 day 4–6 10–24 25–34	• *lacZ* mutant frequency has tissue-specific changes with age: in liver, mutant frequency keeps rising with age; in brain, mutant frequency only increases early (birth→young) and then plateaus; at any age, liver has more mutants than brain	• Both organs show a mix of small mutations and internal deletions • SVs stay rare and non–age-related in the brain but become much more frequent in very old liver, shifting the liver spectrum toward SVs
Dollé et al., 2000	*lacZ* plasmid reporter; plasmid rescue + positive selection for *lacZ*-loss mutants, then restriction mapping and sequencing to classify spectra	heart, small intestine	3–33	• Spontaneous MFs increased in both tissues with age • Mutation accumulation rate is tissue-specific: the small intestine gains mutants about twice as fast as the heart and has a higher MF at every age • Mutations consistent w/ oxidative stress, deamination of 5-methylcytosine, and error-prone DSB repair (for the large rearrangements)	• At the young age, spectra are similar in the heart and small intestine • By old age, spectra are tissue-specific: old small intestine's extra mutations are almost all point mutations, while old heart has many SVs + some point mutations • Recurrent CpG C:G>T:A hotspots present in both tissues
Busuttil et al., 2007	*lacZ* gene reporter-based assay, classify mutants by PCR/AvaI digest, and sequence subsets for spectra	brain, small intestine	7 19 30	• MF rose with age in hippocampus, hypothalamus and total brain, but not cortex or the rest of the brain (intra-brain variation) • In the small intestine, mutation burden is higher in mucosal epithelium than serosa; in duodenal serosa, both point and SVs increase with age • Point mutation accumulation tracks with cell proliferation, whereas genome-rearrangement mutations can also build up in low-proliferative cells	• Brain subparts (hippocampus, hypothalamus) show spectra dominated by (C:G>T:A) transitions, consistent with oxidative damage • Intestinal layers have more diverse point-mutation spectra; SVs increase with age in serosa but stay rare and non–age-related in the brain

Note. This table summarizes representative studies identified through a targeted narrative literature review and is not intended to represent a systematic review.

Table 1.2: Select research on age-related mutation accumulation in mitochondrial and nuclear DNA in mouse and human models using ecNGS methods.

Reference	Method of Mutation Detection	Tissues of Interest	Age Groups	Key Findings	
				Mutation Frequency	Mutation Spectra & Landscape
Sanchez-Contreras et al., 2023 [a]	Duplex Sequencing	heart, skeletal muscle, kidney, liver, brain hippocampus and cerebellum, retina, RPE/choroid	4.5 months 26 months	• mtDNA MF was low in young tissues and increased significantly with age in all tissues examined; when combined with prior data from shared tissues, these results suggested approximately linear accumulation of mtDNA SNVs across the mouse lifespan • High mutation burden/clonality in kidney and liver suggests a role for proliferation, although increased retinal clonality indicates additional tissue-specific mechanisms • Increases in clonally expanded mutations varied between tissues: proportion of clonal mutations rises only in kidney, liver, retina	• Across tissues and ages, mtDNA mutations are dominated by C:G>T:A and T:A>C:G transitions, consistent with replication errors and/or base deamination; ROS-linked C:G>A:T and C:G>G:C transversions are the second most common class but do not increase with age • The spectrum of clonally expanded mtDNA variants differs from that of unique/collapsed variants: across tissues, clonal mutations are enriched for transitions and show an almost complete suppression of ROS-linked C:G>A:T and C:G>G:C transversions (so oxidative-damage-associated transversions less likely to form clones)
Kennedy et al., 2013 [b]	Duplex Sequencing	brain	<1 year >75 years	• Low-frequency/non-clonal mtDNA point mutations increase with age • Mutations consistent with DNA replication errors and spontaneous base hydrolysis, but not oxidative damage	• Mutation spectrum is strongly transition-biased (C:G>T:A and T:A>C:G); the relative proportions of mutation types are similar in young and old brains • The mtDNA D-loop (non-coding control region) has a higher mutation burden than coding regions in both young and old samples, but D-loop and coding regions gain mutations at similar rates with age
Arbeithuber et al., 2020 [a]	Duplex Sequencing	brain, muscle, oocytes	20 days 10 months	• Frequency of de novo mtDNA mutations increased with age in both somatic and germ cells • Mutations consistent with replication errors, spontaneous deamination of methylated nucleotide bases, with oxidative-damage-type mutations playing a minor role	• With age, all transition types increase, but T:A>C:G transitions show the strongest age-related increase in oocytes, whereas C:G>T:A transitions increase most in brain and muscle • The non-coding D-loop has a higher mutation burden than protein-coding, tRNA, and rRNA regions in both young and old mice, especially in somatic tissues of older animals • C:G>T:A transitions were more frequent at CpG sites than non-CpG sites in both age groups, especially in older animals
Neville et al., 2025 [b]	NanoSeq, deep exome/targeted NanoSeq in sperm	sperm (semen) and matched peripheral blood	~24–75 years (sperm) 22–83 years (blood)	• Somatic SNVs in sperm accumulate ~1.67 substitutions per year per haploid genome, and indels ~0.10 per year (linear with age) • In blood, mutation rates are much higher than in sperm	• Sperm SNVs are dominated by clock-like SBS1 and SBS5, while blood shows SBS1, SBS5 plus SBS19 • SBS1 and SBS5 accumulate 6-9× faster in blood than in sperm
Shoag et al., 2025a [b]	NanoSeq, deep PCR-free WGS, and targeted NanoSeq to track clonal mutations over time	sperm (semen) and matched blood or saliva as reference	21.5–36.9 years 33.5–67.3 years	• Sperm point mutations and indels increase ~linearly with age; every man has more sperm mutations at the later time point than at the earlier one • Mutation rates are broadly similar across individuals, with one clear outlier with a history of undescended testis	• Sperm mutation spectra look the same at all ages: mainly clock-like SBS1 and SBS5 • Clonal mosaicism in sperm is very stable over ~20 years: most clonal mutations stay at almost the same frequency, with little evidence for big age-driven clonal expansions or losses
Abascal et al., 2021 [b]	NanoSeq	blood, colonic crypt epithelium, frontal cortex neurons, bladder and colonic smooth muscle	Cord blood / neonates; adults from ~20 up to ~80-100 years	• Somatic substitutions and indels increase ~linearly with age in all tissues • Results suggest mutation accumulation is mainly time- rather than division-dependent • Across tissues, per-cell substitution rates are broadly similar	• Mutation spectra in all tissues are dominated by SBS1-like, SBS5-like, SBS16-like signatures, all of which increase with age • Similar rates and spectra in dividing and non-dividing tissues support a major role for division-independent DNA damage and repair processes

[a] Study investigated aging in a mouse model

[b] Study investigated aging in humans

Note. This table summarizes representative studies identified through a targeted narrative literature review and is not intended to represent a systematic review.

Together, these endogenous processes create a constant background of DNA damage from which mutations arise, making mutation accumulation an unavoidable consequence of normal physiology. Because many of these events are rare and occur stochastically in individual cells, accurately characterizing age-associated mutations requires highly sensitive methods for mutation detection.

1.3 Current approaches to measuring mutagenesis

Because mutagenesis is central to genome stability, tissue function, disease risk, reproductive health, and aging, experimental approaches to measure mutations have been developed and refined over several decades. Early methods often relied on reporter-based systems that converted rare mutational events into selectable or screenable phenotypes; sequencing-based approaches now allow mutations to be detected directly in endogenous DNA.

1.3.1 Classical methods

Mutations have predominantly been quantified using a range of classical reporter-based assays in bacteria and cultured cells, as well as *in vivo* transgenic rodent mutation models (Corvi & Madia, 2017; OECD 471, 2020). Historically, most *in vivo* mutation assays rely on individual reporter genes for mutation quantification, with the most widely used method being the transgenic rodent (TGR) assay (OECD 488, 2025). In the TGR assay, rodents carry recoverable bacterial reporter transgenes integrated into their genome; after exposure or aging, genomic DNA is isolated from tissues, the reporter sequences are rescued into bacteriophage or plasmid vectors, and mutations are scored in bacteria under selective conditions (Lambert et al., 2005). These reporter-based assays have several important limitations: (a) they quantify mutations in single bacterial genes that do not accurately represent the mammalian genome; (b) they are poorly suited for high-resolution analysis of mutation spectra and therefore provide limited insight into the mutational processes underlying observed variants (Beal et al., 2020); and (c) they are limited to specific rodent strains which are difficult to integrate with standard test methods that use other non-transgenic models. These limitations are especially problematic for studying spontaneous, age-

related nuclear mutagenesis, where biologically relevant questions depend on detecting rare mutations as they accumulate over time in endogenous mammalian DNA and vary across tissues.

1.3.2 DNA sequencing

Next-generation sequencing (NGS) provides genome-wide sequence information and has become a central tool for studying somatic mutations in aging tissues. It enables the study of mutations at multiple endogenous loci within any cell type, in physiologically relevant contexts. However, the technical error rates of NGS (1 per 10^2–10^3 sequenced nucleotides; Kennedy et al., 2014) are much higher than the spontaneous mutation frequency (MF; mutations per base pair) of normal tissues (1 in 10^7–10^8 nucleotides; Ren et al., 2022), making it challenging to distinguish between true mutations and sequencing artifacts. Thus, conventional NGS is ill-suited for the accurate characterization of rare age-related mutations that are present in only a few cells within a tissue. These limitations have motivated the development of error-corrected NGS methods, which are designed to detect low-frequency mutations with much higher accuracy.

1.3.3 Application of classical methods in aging research

Early studies using reporter-based and targeted mutation-detection approaches examine how MF and mutation spectra changed with age in mammalian tissues (Table 1.1). Overall, these studies report age-associated increases in MF at reporter loci, although the apparent pattern of accumulation varies across assays and tissues, with some studies reporting approximately linear increases and others showing tissue-specific plateaus or weaker age-related trends. Mutation accumulation, therefore, may be tissue- and cell type-specific (Busuttil et al., 2007; Dollé et al., 1997, 2000; Fischer & Stringer, 2008; Kimoto et al., 2017; Ono et al., 2000; Turker et al., 2007). In several studies, tissues with higher mitotic activity appear to show faster mutation accumulation, a pattern often interpreted as reflecting replication-associated mutagenesis (Busuttil et al., 2007; Dollé et al., 1997), although this relationship is not observed

consistently across assays or studies (Abascal et al., 2021; Ono et al., 2000). Moreover, increases in reporter-gene MF may reflect both the formation of new mutations and the persistence or clonal propagation of cells carrying mutations that arose earlier, particularly in assays that score mutant cells or clusters than independently arising sequence variants. Mutation spectra analyses also suggest tissue- and cell type-specific variation; spectra are relatively similar early in life but appear to diverge with age in some tissues (Busuttil et al., 2007; Dollé et al., 1997, 2000; Kimoto et al., 2017; Turker et al., 2007), and C:G>T:A transitions, often at CpG sites, are among the predominant mutation types observed with age (consistent with spontaneous deamination of 5-methylcytosine; Busuttil et al., 2007; Dollé et al., 2000; Ono et al., 2000). Collectively, these findings provide early evidence that mutation burden and spectra can change with age, while also showing that reporter-based studies are constrained by assay sensitivity and by the limited, often non-endogenous loci examined.

1.4 Duplex Sequencing: High-accuracy detection of ultra-low-frequency mutations

1.4.1 Principles of Duplex Sequencing

Error-corrected next-generation sequencing (ecNGS) methods have been developed to overcome the limited ability of conventional NGS to quantify low-frequency mutations. Most ecNGS approaches use molecular barcodes to group reads from the same original DNA molecule and build consensus sequences that filter out technical errors (Menon & Brash, 2023; Yauk et al., 2026). These include single-strand consensus methods, which correct errors on individual strands, and duplex-based methods, which combine information from both strands of a DNA molecule, such as Duplex Sequencing (DS) and NanoSeq. Among these methods, DS achieves an accuracy that is 10,000-fold greater than standard NGS (Kennedy et al., 2014; Menon & Brash, 2023) and is, therefore, well-suited for detecting rare age-related mutations.

DS reduces sequencing error rates to approximately 1 in 10^7–10^8 sequenced nucleotides by exploiting the double-stranded nature of DNA (Salk & Kennedy, 2020; Schmitt et al., 2012). Briefly, fragmented DNA molecules are labelled with unique molecular identifiers so that, following PCR amplification and sequencing, reads originating from each original molecule can be grouped into tag families. Reads within a tag family are first collapsed into single-strand

consensus sequences (SSCS), which are then paired and further collapsed to generate duplex consensus sequences (DCS). As most sequencing and PCR errors arise randomly and independently on individual strands, they are unlikely to occur at the same position on both strands of the same original DNA molecule. True mutations present in the original double-stranded molecule before library preparation will be represented on both strands and, therefore, retained in the DCS. Thus, DS eliminates nearly all sequencing errors, greatly increasing sensitivity for *de novo* mutation detection.

DS and other ecNGS methods enable direct, sensitive measurement of rare mutations *in vivo* across tissues (Yauk et al., 2026). In addition to estimating MF, DS preserves the sequence context of detected variants, allowing mutation spectra to be used to infer the mutational processes contributing to observed variants (Marchetti et al., 2023b; Salk & Kennedy, 2020). Individual variants are defined by their genomic position and sequence change, allowing DS to distinguish unique mutational events from recurrent detection of an identical variant. This recurrence can mark the persistence or expansion of a mutant lineage; however, targeted DS alone cannot determine whether lineage expansion occurred through neutral drift or through selection acting on an unmeasured variant elsewhere in the genome.

ecNGS can be implemented using broader genomic approaches or targeted panels, depending on the biological question and the depth of detection required. Targeted DS panels trade genomic breadth for sequencing depth, allowing rare variants to be measured sensitively at selected endogenous loci. These panels may be custom-designed for specific loci of interest, while standardized commercial panels have also been developed to support reproducible mutagenicity assessment across laboratories. Commercial DS panels are composed of twenty ~2.4 kb endogenous targets distributed across all autosomes, spanning coding and non-coding loci and mirroring the GC content of the mammalian genome. The loci were chosen because they are selectively neutral, perform optimally in hybrid capture, and lack pseudogenes or repetitive elements that could confound alignment or variant calling (Dodge et al., 2023). Panels with this design are commercially available for mutagenesis studies in mouse (LeBlanc et al., 2022), rat (Smith-Roe et al., 2023), and human (Cho et al., 2023) tissues. Because these predefined targets are deliberately distributed across diverse genomic contexts, DS data

generated with such panels can be stratified by genomic features such as GC content, chromatin state, and functional status (e.g., genic vs non-genic) to examine how these properties influence mutation susceptibility (LeBlanc et al., 2022). Overall, this design allows targeted DS panels to support integrated analysis of mutation burden, spectra, genomic context, and recurrent variant detection within the same assay.

The targeted nature of these panels also defines the scope of inference. Their small, predefined target space (approximately 0.002% of the genome) enables highly sensitive detection of rare mutations, but limits analyses that require genome-wide resolution or interrogation of loci outside the panel. Nevertheless, DS has gained traction as a powerful tool for mutation analysis (Marchetti et al., 2023a), particularly when high sensitivity is required to detect rare mutations in normal tissues.

1.4.2 Applications of duplex-based error-corrected sequencing in aging research

Duplex-based ecNGS methods, including DS and NanoSeq, have been applied to aging tissues, but most early work focused on mitochondrial DNA (mtDNA) rather than the nuclear genome. This emphasis is partly technical: the mitochondrial genome is small, approximately 16.5 kb in humans and 16.3 kb in mice, compared with approximately 3.2 Gb and 2.7 Gb for the human and mouse nuclear genomes, respectively, making ultra-deep coverage feasible (Brown, 2002b; NCBI, 2012, 2023). In addition, mtDNA typically carries somatic mutations at much higher frequencies (on the order of 10^{-6}–10^{-5} per base in aged tissues; Kennedy et al., 2013; Sanchez-Contreras et al., 2023), often at least an order of magnitude higher than nuclear DNA (Árnadóttir et al., 2024; Leuthner & Meyer, 2021). The small genome size and elevated per-base MF make rare mtDNA variants easier to detect with DS than equivalently rare mutations dispersed across the nuclear genome.

DS studies in mouse and human tissues (Table 1.2) have reported that mtDNA MF increases with age and that point mutations are strongly transition-biased (predominantly C:G>T:A and T:A>C:G) across tissues (Arbeithuber et al., 2020; Kennedy et al., 2013; Sanchez-Contreras et al., 2023). These spectra, coupled with the absence of age-related increases in

ROS-linked transversions, are consistent with replication errors and spontaneous base deamination contributing substantially to age-related mtDNA mutations. DS studies also indicate that mtDNA mutation accumulation is tissue- and region-specific: low-frequency substitutions and small indels are often more frequent in the non-coding D-loop region than in coding regions, while larger multi-tissue analyses show that age-related mtDNA mutation accumulation and clonal expansion patterns vary substantially among tissues (Arbeithuber et al., 2020; Kennedy et al., 2013; Sanchez-Contreras et al., 2023). However, caution is required when extrapolating these observations to nuclear DNA because mitochondrial and nuclear genomes differ substantially in size, organization, replication, repair, copy number, and cellular context.

Age-focused applications of duplex-based ecNGS to nuclear DNA remain relatively sparse and, to date, have largely used genome-scale methods such as NanoSeq rather than DS. NanoSeq (Abascal et al., 2021), a genome-scale ecNGS method, has been applied to age-related mutation accumulation in human somatic tissues and the male germline (Abascal et al., 2021; Neville et al., 2025; Shoag et al., 2025; Table 1.2). These studies suggest that nuclear mutations accumulate approximately linearly with age. In human somatic tissues, NanoSeq has shown broadly similar per-cell substitution rates across dividing and non-dividing tissues, supporting a substantial contribution of time-dependent mutational processes rather than cell division alone (Abascal et al., 2021). In the male germline, age-associated increases are reported in sperm SNVs and indels, while longitudinal sampling suggests that most detectable clonal mutations remain at relatively stable allele fractions over time rather than showing strong age-related clonal expansion (Neville et al., 2025; Shoag et al., 2025). Somatic mutation spectra can also be decomposed into “mutational signatures,” defined as characteristic patterns of mutation types and sequence contexts that reflect the activity of underlying mutational processes or DNA repair pathways (Alexandrov et al., 2013, 2020). In these datasets, the spectra are dominated by SBS1 and SBS5, two signatures observed across many normal tissues and often described as “clock-like” because their activity increases steadily with chronological age. Overall, ecNGS methods have emerged as powerful tools for ultra-sensitive mutation analysis, with significant potential to advance our understanding of how mutation accumulation contributes to aging.

1.5 Knowledge gaps in age-related mutagenesis across tissues

1.5.1 Biological gaps

Previous studies provide strong evidence that mutation burden can change with age, but the available evidence remains fragmented across methods, species, tissues, and genomic compartments. Reporter-based assays provided early evidence for age-associated increases in MF, but they were limited to specific reporter loci and therefore could not provide a high-resolution picture of endogenous nuclear mutagenesis. DS studies of mtDNA have also shown that mutation burden increases with age; however, mtDNA findings do not resolve how mutations accumulate in nuclear DNA because the two genomes are not interchangeable. mtDNA is present in many copies per cell and is maintained within a distinct mitochondrial environment (Kennedy et al., 2013), whereas the nuclear genome contains the vast majority of protein-coding and regulatory sequence and is organized within nuclear chromatin. Age-related mutational patterns observed in mtDNA therefore cannot be assumed to reflect the processes or consequences of mutation accumulation in nuclear DNA, highlighting the need for high-accuracy nuclear DNA studies.

More recent duplex-based studies of age-related nuclear mutagenesis have largely been conducted in humans, which are the ultimate organism of interest but are difficult to control with respect to genetic background, environmental exposures, diet, disease history, and other life-history variables. Human studies are also constrained by tissue availability and are therefore biased toward accessible samples such as blood and sperm, with more limited information from internal organs. In mice, by contrast, aging studies can be performed in controlled environments and can include multiple tissues collected under standardized conditions across age groups. However, high-accuracy studies of age-associated nuclear mutagenesis in mice remain scarce. As a result, existing data do not yet provide a high-resolution picture of how nuclear mutation burden changes with age in a controlled mammalian system.

Previous studies suggest that mutation accumulation can differ among tissues (Moore et al., 2021), but they do not yet determine how these differences emerge across the lifespan. Age-associated nuclear mutagenesis is unlikely to represent a single uniform process across the

body. Although overall MF provides a measure of mutation burden, understanding age-associated mutagenesis requires determining how that burden is structured across tissues. Age-related mutation burden may reflect independent mutational events, expansion or persistence of mutant cell lineages, shifts in mutation spectra, or broader tissue-specific differences. Characterizing these patterns across tissues can therefore provide insight into the biological environments experienced by aging cells. Without nuclear data across tissues and ages, it remains difficult to determine whether age-related mutagenesis reflects a broadly shared aging process or tissue-specific mutational environments.

Addressing this gap requires integrated, lifespan-scale, cross-tissue nuclear mutation data from a controlled mammalian model to determine which features of mutation accumulation are shared across aging tissues and which reflect trajectories shaped by local tissue biology.

1.5.2 Study-design gaps

A major study-design challenge in age-related nuclear mutagenesis is that spontaneous mutations are rare events. Unlike exposure-based mutagenicity studies, in which the exposure may generate enough mutations to make group differences easier to detect, age-related increases in MF are expected to be small, leading to sparse mutation counts and substantial sampling variability. This makes reliable detection of age-related differences statistically challenging.

DS addresses one major barrier to studying rare mutations by improving confidence that detected variants represent true mutations rather than sequencing or PCR artifacts. However, analytical sensitivity and statistical power are not the same. Analytical sensitivity concerns whether individual rare mutations can be detected accurately, whereas statistical power concerns whether the overall study design can detect a true difference in MF between biological groups. Thus, while DS produces highly accurate mutation calls, a study may still fail to detect age- or tissue-related differences if the design does not capture enough information to support those comparisons.

Even for exposure-based mutagenicity studies, where DS is more commonly applied, there has been limited formal guidance on how to choose sample size and sequencing depth. Power depends on several quantities, including baseline MF, expected effect size, variability between animals, and the number of informative duplex bases generated per sample. These considerations are especially important for aging studies, where expected changes in spontaneous MF may be small. A formal power-analysis framework is therefore needed to evaluate how these parameters influence the ability to detect age-related differences in overall MF and to guide realistic DS study design.

1.6 Experimental model system and tissues

This thesis uses MutaMouse males, a transgenic CD2F1 (BALB/c × DBA/2) strain originally developed for *in vivo* mutagenicity testing. It carries an integrated λgt10–*lacZ* reporter transgene (Meier et al., 2019; White et al., 2019); however, this transgene was not used as an experimental endpoint in the present work, which instead characterized mutations in endogenous nuclear DNA. Whole-genome sequencing has not identified disruptions in pathways relevant to genome maintenance or cellular metabolism in the MutaMouse (Meier et al., 2019), supporting its suitability as a genetically characterized, immunocompetent background for studying age-related mutation accumulation.

MutaMouse was selected primarily because Health Canada maintained long-standing colonies and an extensive archive of tissues and associated historical data from previous mutagenesis studies. These resources, supplemented with a smaller set of animals aged for this project, provided tissues from multiple organs and age groups, allowing us to assemble a cross-sectional series spanning much of the MutaMouse lifespan. This design enables direct comparison of age-related changes in nuclear mutation burden and spectra across tissues, while reducing the number of additional animals required and aligning with 3R principles (Replacement, Reduction, Refinement) through the reuse of existing samples.

The age range of the MutaMouse animals was selected to capture broad biological phases across the postnatal lifespan, from early life at 3 weeks to mature adulthood and late

life at 78–85 weeks. The 3-week time point corresponds to weaning and was selected as the earliest practical postnatal reference point in this study. At this stage, mice are independently viable and provide sufficient tissue for downstream sequencing, while still representing an early-life cohort. However, this time point should not be interpreted as a mutation-free baseline, since mutations detected at 3 weeks may include events that arose during prenatal development or the early postnatal period. The intermediate adult cohorts and ≥78-week aged cohort were included to compare mutation burden across mature and late-life stages. This age structure reflects the fact that aging is a lifelong, biologically non-uniform process, during which mutation accumulation may be shaped by changing patterns. Although previous studies suggest that some mutation classes can accumulate approximately linearly with age (Abascal et al., 2021; Fischer & Stringer, 2008; Neville et al., 2025; Ono et al., 2000; Sanchez-Contreras et al., 2023; Shoag et al., 2025), this thesis does not assume a single linear trajectory across the lifespan. Instead, the selected age groups allow the assessment of whether age-related mutation patterns are broadly consistent across tissues or vary according to biological phase and tissue context.

Tissue selection was designed to capture biological contexts that may shape how nuclear mutations arise and accumulate with age. Features such as proliferative (mitotic) activity, metabolic demand, environmental exposure, regenerative organization, and tissue-specific genome-maintenance or quality-control environments are expected to influence age-associated accumulation patterns, including mutation spectra and opportunity for clonal expansion of mutant cells. Sex may also influence age-related mutagenesis (Turker et al., 2007), meaning that cross-tissue comparisons are most interpretable when conducted within a single experimental sex rather than across mixed-sex groups. In addition, male germ cells provide sufficient material for nuclear DS while oocyte-based DS is constrained by limited oocyte availability and low per-cell DNA input (Arbeithuber et al., 2020), and human pedigree studies show a strong paternal bias in germline *de novo* mutations (Gao et al., 2019; Kong et al., 2012). This thesis therefore focuses on male mice, enabling comparison of somatic tissues and male germ cells within the same sex and genetic background while avoiding sex as an additional

biological variable. To probe tissue-specific biological contexts within this framework, we focused on four mouse tissues that differ markedly in their features:

- **Liver**. The liver is one of the most metabolically active organs in the body and the main site of first-pass detoxification. Most of the tissue volume comes from hepatocytes, which perform the bulk of metabolic and detoxification functions, with additional contributions from non-parenchymal cells such as Kupffer cells, sinusoidal endothelial cells and hepatic stellate cells (Casas-Grajales & Muriel, 2015; Malarkey et al., 2005). The high metabolic demand on these cells, together with the metabolism of endogenous compounds, generates ROS and other reactive intermediates through pathways such as cytochrome P450 metabolism, mitochondrial electron transport, peroxisomal oxidation, and inflammatory cell activity (Casas-Grajales & Muriel, 2015) which can damage DNA if not fully neutralized. At the same time, hepatocyte proliferation is relatively low in adult mice (mitotic index ~0.36%; White et al., 2017), so cells experience repeated metabolic and oxidative stress over long periods. The liver is, therefore, a useful tissue for characterizing age-related mutation accumulation in a high-metabolism, low-proliferation, detoxification-intensive environment.

- **Lung**. The lung is a specialized barrier organ for gas exchange that is constantly exposed to the external environment and high oxygen levels, which impose a moderate but continuous metabolic load on the tissue. It contains a mixture of epithelial cells (which line the airways and alveoli), immune cells, and various stromal and endothelial support cells (Angelidis et al., 2019; Kawasaki et al., 2022; Keenum et al., 2023). The lung epithelium also contains several stem and progenitor cell populations that provide substantial regenerative capacity (Angelidis et al., 2019), despite the tissue as a whole being mostly quiescent and having relatively low mitotic activity (mitotic index ~0.14%; White et al., 2017). Even under normal conditions, breathing exposes lung cells to mechanical stretch, fluctuating oxygen levels, and immune responses to inhaled particles or microbes. With advancing age, lung structure and immune function progressively decline, lung function deteriorates even in otherwise healthy individuals, and susceptibility to chronic lung diseases such as chronic obstructive pulmonary disease (COPD), lung cancer, and interstitial lung disease increases (Angelidis et

al., 2019). Together, these features make the lung a useful tissue for characterizing age-related mutation accumulation in a barrier organ directly exposed to the external environment and the cyclic oxygen changes associated with respiration.

- **Bone marrow (BM)**. BM contains haematopoietic stem cells (HSCs) and many stages of developing blood cells, together with stromal and other support cells. It is one of the most proliferative tissues in the body, producing large numbers of new blood cells throughout life (Jaiswal & Ebert, 2019; Lauridsen et al., 2018) with a mitotic index of ~5% (White et al., 2017). In this setting, DNA damage can be converted into fixed mutations as cells replicate and pass their genomes to daughter cells. Age-related changes in haematopoietic stem cells, including clonal expansions driven by somatic mutations, have been increasingly detected in humans and model organisms (Jaiswal & Ebert, 2019). BM, therefore, is particularly useful for characterizing age-related mutation accumulation in a highly proliferative compartment.

- **Male germ cells**. The seminiferous tubules (ST) of the testes are the site of spermatogenesis and contain spermatogonial stem cells (SSCs), germ cells at different stages of sperm development, as well as a small contribution from supporting Sertoli and other somatic cells (Diao et al., 2022; Ryu et al., 2006). SSCs form a self-renewing population that sustains sperm production throughout life through continued proliferation and differentiation into spermatogonial progenitors and downstream germ cells (mitotic index of approximately 0.94%; Erexson et al., unpublished). Over time, however, this long-term proliferative activity is accompanied by age-related changes in SSCs and a decline in their functional capacity (Kubota & Brinster, 2018; Ryu et al., 2006). Because SSCs and their progeny give rise to sperm, mutations arising in the male germline can be transmitted to the next generation. In humans, large-scale sequencing studies have shown that the number of germline *de novo* mutations in offspring increases with both paternal and maternal age, and that approximately 80% of these mutations originate from the paternal rather than the maternal germline (Gao et al., 2019; Shoag et al., 2025; Shojaeisaadi et al., 2024). This pattern has traditionally been attributed to spermatogenesis, in which male germ cells undergo many more rounds of cell division than oocytes; however, recent

human studies suggest that division-independent mutational processes also play a role (de Manuel et al., 2022; Shojaeisaadi et al., 2024). Moreover, advanced paternal age is associated with increased germline *de novo* mutations and elevated risks of certain genetic and neurodevelopmental disorders, which together account for a sizeable fraction of rare pediatric disease (Kong et al., 2012; Krug et al., 2020; Shoag et al., 2025; Shojaeisaadi et al., 2024). Nonetheless, the male germline appears comparatively resistant to mutation and maintains high genomic integrity, supported by exceptionally efficient DNA repair (Dong et al., 2022; Shoag et al., 2025; Xavier et al., 2018) as well as other quality-control mechanisms such as apoptosis-mediated removal of abnormal or damaged germ cells and antioxidant and metabolic regulation within the seminiferous epithelium that buffer oxidative stress (Asadi et al., 2021; Dong et al., 2022; Dutta et al., 2021). Thus, a mixed population of male germ cells from the ST was collected to characterize age-related mutation accumulation for comparison of somatic and germline compartments within the same organism.

Overall, these tissues span high- and low-proliferation compartments, and somatic and germline cells, providing a framework for comparing age-related nuclear mutagenesis across distinct biological contexts.

1.7 Thesis hypotheses and objectives

The overarching objective of this thesis is to characterize how mutations in nuclear DNA accumulate with age across mouse tissues with different biological characteristics. Based on previous work on age-related mutagenesis (summarized in §1.3.3 and §1.4.2), this thesis tests the hypothesis that nuclear MF and mutation spectra change with age in a tissue-dependent manner, reflecting differences in tissue biology.

To address this hypothesis, the thesis has two specific objectives:

Aim 1 (Chapter 2): Develop and apply a statistical framework for planning Duplex Sequencing mutagenesis studies in mice.

1.1 Evaluate how baseline mutation frequencies, sample variability, expected effect sizes, sample sizes, and sequencing depth influence statistical power to detect differences in mutation frequency, using 14-week-old adult MutaMouse tissues as an empirical reference condition.

1.2 Use this framework to estimate sample size requirements and design considerations for DS-based aging studies under realistic scenarios.

Aim 2 (Chapter 3): Empirically characterize age-related nuclear mutagenesis across multiple tissues in MutaMouse males using Duplex Sequencing.

2.1 Quantify age-related changes in mutation frequency across liver, lung, bone marrow, and germ cells from MutaMouse males.

2.2 Characterize age- and tissue-related differences in mutation spectra.

2.3 Examine how genomic features relate to age-related mutation patterns.

1.8 References

Abascal, F., Harvey, L. M. R., Mitchell, E., Lawson, A. R. J., Lensing, S. V., Ellis, P., Russell, A. J. C., Alcantara, R. E., Baez-Ortega, A., Wang, Y., Kwa, E. J., Lee-Six, H., Cagan, A., Coorens, T. H. H., Chapman, M. S., Olafsson, S., Leonard, S., Jones, D., Machado, H. E., … Martincorena, I. (2021). Somatic mutation landscapes at single-molecule resolution. *Nature*, *593*(7859), 405–410. https://doi.org/10.1038/s41586-021-03477-4

Ageing and health. (2025, October 1). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., Wu, Y., Boot, A., Covington, K. R., Gordenin, D. A., Bergstrom, E. N., Islam, S. M. A., Lopez-Bigas, N., Klimczak, L. J., McPherson, J. R., Morganella, S., Sabarinathan, R., Wheeler, D. A., Mustonen, V., Getz, G., … Stratton, M. R. (2020). The repertoire of mutational signatures in human cancer. *Nature*, *578*(7793), 94–101. https://doi.org/10.1038/s41586-020-1943-3

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A.-L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., … Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature*, *500*(7463), 415–421. https://doi.org/10.1038/nature12477

Angelidis, I., Simon, L. M., Fernandez, I. E., Strunz, M., Mayr, C. H., Greiffo, F. R., Tsitsiridis, G., Ansari, M., Graf, E., Strom, T.-M., Nagendran, M., Desai, T., Eickelberg, O., Mann, M., Theis, F. J., & Schiller, H. B. (2019). An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nature Communications*, *10*(1), 963. https://doi.org/10.1038/s41467-019-08831-9

Arbeithuber, B., Hester, J., Cremona, M. A., Stoler, N., Zaidi, A., Higgins, B., Anthony, K., Chiaromonte, F., Diaz, F. J., & Makova, K. D. (2020). Age-related accumulation of de novo mitochondrial mutations in mammalian oocytes and somatic tissues. *PLOS Biology*, *18*(7), e3000745. https://doi.org/10.1371/journal.pbio.3000745

Árnadóttir, E. R., Moore, K. H. S., Guðmundsdóttir, V. B., Ebenesersdóttir, S. S., Guity, K., Jónsson, H., Stefánsson, K., & Helgason, A. (2024). The rate and nature of mitochondrial DNA mutations in human pedigrees. *Cell*, *187*(15), 3904-3918.e8. https://doi.org/10.1016/j.cell.2024.05.022

Asadi, A., Ghahremani, R., Abdolmaleki, A., & Rajaei, F. (2021). Role of sperm apoptosis and oxidative stress in male infertility: A narrative review. *International Journal of Reproductive Biomedicine*, *19*(6), 493–504. https://doi.org/10.18502/ijrm.v19i6.9371

Beal, M. A., Meier, M. J., LeBlanc, D. P., Maurice, C., O'Brien, J. M., Yauk, C. L., & Marchetti, F. (2020). Chemically induced mutations in a MutaMouse reporter gene inform

mechanisms underlying human cancer mutational signatures. *Communications Biology*, *3*(1), 1–11. https://doi.org/10.1038/s42003-020-01174-y

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., … Rakyan, V. K. (2019). DNA methylation aging clocks: Challenges and recommendations. *Genome Biology*, *20*(1), 249. https://doi.org/10.1186/s13059-019-1824-y

Brown, T. A. (2002a). Mutation, Repair and Recombination. In *Genomes. 2nd edition*. Wiley-Liss. https://www.ncbi.nlm.nih.gov/books/NBK21114/

Brown, T. A. (2002b). The Human Genome. In *Genomes. 2nd edition*. Wiley-Liss. https://www.ncbi.nlm.nih.gov/books/NBK21134/

Busuttil, R. A., Garcia, A. M., Reddick, R. L., Dollé, M. E. T., Calder, R. B., Nelson, J. F., & Vijg, J. (2007). Intra-Organ Variation in Age-Related Mutation Accumulation in the Mouse. *PLOS ONE*, *2*(9), e876. https://doi.org/10.1371/journal.pone.0000876

Casas-Grajales, S., & Muriel, P. (2015). Antioxidants in liver health. *World Journal of Gastrointestinal Pharmacology and Therapeutics*, *6*(3), 59–72. https://doi.org/10.4292/wjgpt.v6.i3.59

Chen, R., Wang, Y., Zhang, S., Bulloch, G., Zhang, J., Liao, H., Shang, X., Clark, M., Peng, Q., Ge, Z., Cheng, C.-Y., Gao, Y., He, M., & Zhu, Z. (2023). Biomarkers of ageing: Current state-of-art, challenges, and opportunities. *MedComm – Future Medicine*, *2*(2), e50. https://doi.org/10.1002/mef2.50

Cho, E., Swartz, C. D., Williams, A., Rivas, M., Recio, L., Witt, K. L., Schmidt, E. K., Yaplee, J., Smith, T. H., Van, P., Lo, F. Y., Valentine, C. C., Salk, J. J., Marchetti, F., Smith-Roe, S. L., & Yauk, C. L. (2023). *Error-corrected Duplex Sequencing enables direct detection and quantification of mutations in human TK6 cells with remarkable inter-laboratory consistency* (p. 2023.02.22.529418). bioRxiv. https://doi.org/10.1101/2023.02.22.529418

Corvi, R., & Madia, F. (2017). *In vitro* genotoxicity testing–Can the performance be enhanced? *Food and Chemical Toxicology, Strategies in Genotoxicity Testing*, *106*, 600–608. https://doi.org/10.1016/j.fct.2016.08.024

de Manuel, M., Wu, F. L., & Przeworski, M. (2022). A paternal bias in germline mutation is widespread in amniotes and can arise independently of cell division numbers. *eLife*, *11*, e80008. https://doi.org/10.7554/eLife.80008

Diao, L., Turek, P. J., John, C. M., Fang, F., & Reijo Pera, R. A. (2022). Roles of Spermatogonial Stem Cells in Spermatogenesis and Fertility Restoration. *Frontiers in Endocrinology*, *13*. https://doi.org/10.3389/fendo.2022.895528

Dodge, A. E., LeBlanc, D. P. M., Zhou, G., Williams, A., Meier, M. J., Van, P., Lo, F. Y., Valentine Iii, C. C., Salk, J. J., Yauk, C. L., & Marchetti, F. (2023). Duplex sequencing provides detailed characterization of mutation frequencies and spectra in the bone marrow of MutaMouse males exposed to procarbazine hydrochloride. *Archives of Toxicology*, *97*(8), 2245–2259. https://doi.org/10.1007/s00204-023-03527-y

Dollé, M. E. T., Giese, H., Hopkins, C. L., Martus, H.-J., Hausdorff, J. M., & Vijg, J. (1997). Rapid accumulation of genome rearrangements in liver but not in brain of old mice. *Nature Genetics*, *17*(4), Article 4. https://doi.org/10.1038/ng1297-431

Dollé, M. E. T., Snyder, W. K., Gossen, J. A., Lohman, P. H. M., & Vijg, J. (2000). Distinct spectra of somatic mutations accumulated with age in mouse heart and small intestine. *Proceedings of the National Academy of Sciences*, *97*(15), 8403–8408. https://doi.org/10.1073/pnas.97.15.8403

Dong, S., Chen, C., Zhang, J., Gao, Y., Zeng, X., & Zhang, X. (2022). Testicular aging, male fertility and beyond. *Frontiers in Endocrinology*, *13*. https://doi.org/10.3389/fendo.2022.1012119

Durland, J., & Ahmadian-Moghadam, H. (2023). Genetics, Mutagenesis. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK560519/

Dutta, S., Sengupta, P., Slama, P., & Roychoudhury, S. (2021). Oxidative Stress, Testicular Inflammatory Pathways, and Male Reproduction. *International Journal of Molecular Sciences*, *22*(18), 10043. https://doi.org/10.3390/ijms221810043

Erickson, R. P. (2003). Somatic gene mutation and human disease other than cancer. *Mutation Research/Reviews in Mutation Research, Special Issue in Honor of James V. Neel*, *543*(2), 125–136. https://doi.org/10.1016/S1383-5742(03)00010-3

Fischer, J. M., & Stringer, J. R. (2008). Mutation in aging mice occurs in diverse cell types that proliferate postmutation. *Aging Cell*, *7*(5), 667–680. https://doi.org/10.1111/j.1474-9726.2008.00416.x

Gao, Z., Moorjani, P., Sasani, T. A., Pedersen, B. S., Quinlan, A. R., Jorde, L. B., Amster, G., & Przeworski, M. (2019). Overlooked roles of DNA damage and maternal age in generating human germline mutations. *Proceedings of the National Academy of Sciences*, *116*(19), 9491–9500. https://doi.org/10.1073/pnas.1901259116

Gladyshev, V. N., Anderson, B., Barlit, H., Barré, B., Beck, S., Behrouz, B., Belsky, D. W., Chaix, A., Chamoli, M., Chen, B. H., Cheng, K., Chuprin, J., Churchill, G. A., Cipriano, A., Colville, A., Deelen, J., Deigin, Y., Edmonds, K. K., English, B. W., … Zhavoronkov, A. (2024). Disagreement on foundational principles of biological aging. *PNAS Nexus*, *3*(12), pgae499. https://doi.org/10.1093/pnasnexus/pgae499

Jackson, M., Marks, L., May, G. H. W., & Wilson, J. B. (2018). The genetic basis of disease. *Essays in Biochemistry*, *62*(5), 643–723. https://doi.org/10.1042/EBC20170053

Jaiswal, S., & Ebert, B. L. (2019). Clonal hematopoiesis in human aging and disease. *Science (New York, N.Y.)*, *366*(6465), eaan4673. https://doi.org/10.1126/science.aan4673

Jones, M. J., Goodman, S. J., & Kobor, M. S. (2015). DNA methylation and healthy human aging. *Aging Cell*, *14*(6), 924–932. https://doi.org/10.1111/acel.12349

Kawasaki, T., Ikegawa, M., & Kawai, T. (2022). Antigen Presentation in the Lung. *Frontiers in Immunology*, *13*. https://doi.org/10.3389/fimmu.2022.860915

Keenum, M. C., Chatterjee, P., Atalis, A., Pandey, B., Jimenez, A., & Roy, K. (2023). Single-cell epitope-transcriptomics reveal lung stromal and immune cell response kinetics to nanoparticle-delivered RIG-I and TLR4 agonists. *Biomaterials*, *297*, 122097. https://doi.org/10.1016/j.biomaterials.2023.122097

Kennedy, S. R., Salk, J. J., Schmitt, M. W., & Loeb, L. A. (2013). Ultra-Sensitive Sequencing Reveals an Age-Related Increase in Somatic Mitochondrial Mutations That Are Inconsistent with Oxidative Damage. *PLOS Genetics*, *9*(9), e1003794. https://doi.org/10.1371/journal.pgen.1003794

Kennedy, S. R., Schmitt, M. W., Fox, E. J., Kohrn, B. F., Salk, J. J., Ahn, E. H., Prindle, M. J., Kuong, K. J., Shen, J.-C., Risques, R.-A., & Loeb, L. A. (2014). Detecting ultralow-frequency mutations by Duplex Sequencing. *Nature Protocols*, *9*(11), 2586–2606. https://doi.org/10.1038/nprot.2014.170

Kimoto, T., Kay, J. E., Li, N., & Engelward, B. P. (2017). Recombinant cells in the lung increase with age via de novo recombination events and clonal expansion. *Environmental and Molecular Mutagenesis*, *58*(3), 135–145. https://doi.org/10.1002/em.22082

Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S. A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., Wong, W. S. W., Sigurdsson, G., Walters, G. B., Steinberg, S., Helgason, H., Thorleifsson, G., Gudbjartsson, D. F., Helgason, A., Magnusson, O. T., … Stefansson, K. (2012). Rate of de novo mutations and the importance of father's age to disease risk. *Nature*, *488*(7412), 471–475. https://doi.org/10.1038/nature11396

Krug, A., Wöhr, M., Seffer, D., Rippberger, H., Sungur, A. Ö., Dietsche, B., Stein, F., Sivalingam, S., Forstner, A. J., Witt, S. H., Dukal, H., Streit, F., Maaser, A., Heilmann-Heimbach, S., Andlauer, T. F. M., Herms, S., Hoffmann, P., Rietschel, M., Nöthen, M. M., … Kircher, T. (2020). Advanced paternal age as a risk factor for neurodevelopmental disorders: A translational study. *Molecular Autism*, *11*, 54. https://doi.org/10.1186/s13229-020-00345-2

Kubota, H., & Brinster, R. L. (2018). Spermatogonial stem cells†. *Biology of Reproduction*, *99*(1), 52–74. https://doi.org/10.1093/biolre/ioy077

Lambert, I. B., Singer, T. M., Boucher, S. E., & Douglas, G. R. (2005). Detailed review of transgenic rodent mutation assays. *Mutation Research/Reviews in Mutation Research*, *590*(1), 1–280. https://doi.org/10.1016/j.mrrev.2005.04.002

Lauridsen, F. K. B., Jensen, T. L., Rapin, N., Aslan, D., Wilhelmson, A. S., Pundhir, S., Rehn, M., Paul, F., Giladi, A., Hasemann, M. S., Serup, P., Amit, I., & Porse, B. T. (2018). Differences in Cell Cycle Status Underlie Transcriptional Heterogeneity in the HSC Compartment. *Cell Reports*, *24*(3), 766–780. https://doi.org/10.1016/j.celrep.2018.06.057

LeBlanc, D. P. M., Meier, M., Lo, F. Y., Schmidt, E., Valentine, C., Williams, A., Salk, J. J., Yauk, C. L., & Marchetti, F. (2022). Duplex sequencing identifies genomic features that determine susceptibility to benzo(a)pyrene-induced in vivo mutations. *BMC Genomics*, *23*(1), 542. https://doi.org/10.1186/s12864-022-08752-w

Leuthner, T. C., & Meyer, J. N. (2021). Mitochondrial DNA Mutagenesis: Feature of and Biomarker for Environmental Exposures and Aging. *Current Environmental Health Reports*, *8*(4), 294–308. https://doi.org/10.1007/s40572-021-00329-1

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The Hallmarks of Aging. *Cell*, *153*(6), 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039

Malarkey, D. E., Johnson, K., Ryan, L., Boorman, G., & Maronpot, R. R. (2005). New Insights into Functional Aspects of Liver Morphology. *Toxicologic Pathology*, *33*(1), 27–34. https://doi.org/10.1080/01926230590881826

Marchetti, F., Cardoso, R., Chen, C. L., Douglas, G. R., Elloway, J., Escobar, P. A., Harper Jr, T., Heflich, R. H., Kidd, D., Lynch, A. M., Myers, M. B., Parsons, B. L., Salk, J. J., Settivari, R. S., Smith-Roe, S. L., Witt, K. L., Yauk, C., Young, R. R., Zhang, S., & Minocherhomji, S. (2023). Error-corrected next-generation sequencing to advance nonclinical genotoxicity and carcinogenicity testing. *Nature Reviews Drug Discovery*, *22*(3), 165–166. https://doi.org/10.1038/d41573-023-00014-y

Marchetti, F., Cardoso, R., Chen, C. L., Douglas, G. R., Elloway, J., Escobar, P. A., Harper, T., Heflich, R. H., Kidd, D., Lynch, A. M., Myers, M. B., Parsons, B. L., Salk, J. J., Settivari, R. S., Smith-Roe, S. L., Witt, K. L., Yauk, C. L., Young, R., Zhang, S., & Minocherhomji, S. (2023). Error-corrected next generation sequencing – Promises and challenges for genotoxicity and cancer risk assessment. *Mutation Research/Reviews in Mutation Research*, *792*, 108466. https://doi.org/10.1016/j.mrrev.2023.108466

Martincorena, I. (2019). Somatic mutation and clonal expansions in human tissues. *Genome Medicine*, *11*, 35. https://doi.org/10.1186/s13073-019-0648-4

Meier, M. J., Beal, M. A., Schoenrock, A., Yauk, C. L., & Marchetti, F. (2019). Whole Genome Sequencing of the Mutamouse Model Reveals Strain- and Colony-Level Variation, and

Genomic Features of the Transgene Integration Site. *Scientific Reports*, *9*(1), 13775. https://doi.org/10.1038/s41598-019-50302-0

Melzer, D., Pilling, L. C., & Ferrucci, L. (2020). The genetics of human ageing. *Nature Reviews. Genetics*, *21*(2), 88–101. https://doi.org/10.1038/s41576-019-0183-6

Menon, V., & Brash, D. E. (2023). Next-generation sequencing methodologies to detect low-frequency mutations: "Catch me if you can." *Mutation Research/Reviews in Mutation Research*, *792*, 108471. https://doi.org/10.1016/j.mrrev.2023.108471

Moore, L., Cagan, A., Coorens, T. H. H., Neville, M. D. C., Sanghvi, R., Sanders, M. A., Oliver, T. R. W., Leongamornlert, 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., … Rahbari, R. (2021). The mutational landscape of human somatic and germline cells. *Nature*, *597*(7876), 381–386. https://doi.org/10.1038/s41586-021-03822-7

NCBI. (2012). *Mus musculus genome assembly GRCm38*. NCBI. https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001635.20/

NCBI. (2023). *Mus musculus mitochondrion, complete genome* [Dataset]. NCBI Nucleotide Database (34538597). http://www.ncbi.nlm.nih.gov/nuccore/NC_005089.1

Neville, M. D. C., Lawson, A. R. J., Sanghvi, R., Abascal, F., Pham, M. H., Cagan, A., Nicola, P. A., Bayzetinova, T., Baez-Ortega, A., Roberts, K., Lensing, S. V., Widaa, S., Alcantara, R. E., García, M. P., Wadge, S., Stratton, M. R., Campbell, P. J., Small, K., Martincorena, I., … Rahbari, R. (2025). Sperm sequencing reveals extensive positive selection in the male germline. *Nature*, *647*(8089), 421–428. https://doi.org/10.1038/s41586-025-09448-3

OECD. (2020). *Test No. 471: Bacterial Reverse Mutation Test*. Organisation for Economic Co-operation and Development. https://doi.org/10.1787/9789264071247-en

OECD. (2025). *Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays*. Organisation for Economic Co-operation and Development. https://doi.org/10.1787/9789264203907-en

Oh, H. S.-H., Rutledge, J., Nachun, D., Pálovics, R., Abiose, O., Moran-Losada, P., Channappa, D., Urey, D. Y., Kim, K., Sung, Y. J., Wang, L., Timsina, J., Western, D., Liu, M., Kohlfeld, P., Budde, J., Wilson, E. N., Guen, Y., Maurer, T. M., … Wyss-Coray, T. (2023). Organ aging signatures in the plasma proteome track health and disease. *Nature*, *624*(7990), Article 7990. https://doi.org/10.1038/s41586-023-06802-1

Ono, T., Ikehata, H., Nakamura, S., Saito, Y., Hosoi, Y., Takai, Y., Yamada, S., Onodera, J., & Yamamoto, K. (2000). Age-associated increase of spontaneous mutant frequency and molecular nature of mutation in newborn and old lacZ-transgenic mouse. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *447*(2), 165–177. https://doi.org/10.1016/S0027-5107(99)00200-6

Partridge, L., Deelen, J., & Slagboom, P. E. (2018). Facing up to the global challenges of ageing. *Nature*, *561*(7721), Article 7721. https://doi.org/10.1038/s41586-018-0457-8

Ren, P., Dong, X., & Vijg, J. (2022). Age-related somatic mutation burden in human tissues. *Frontiers in Aging*, *3*. https://www.frontiersin.org/articles/10.3389/fragi.2022.1018119

Ricke, R. M., & van Deursen, J. M. (2013). Aneuploidy in health, disease, and aging. *Journal of Cell Biology*, *201*(1), 11–21. https://doi.org/10.1083/jcb.201301061

Ryu, B.-Y., Orwig, K. E., Oatley, J. M., Avarbock, M. R., & Brinster, R. L. (2006). Effects of Aging and Niche Microenvironment on Spermatogonial Stem Cell Self-Renewal. *STEM CELLS*, *24*(6), 1505–1511. https://doi.org/10.1634/stemcells.2005-0580

Salk, J. J., & Kennedy, S. R. (2020). Next-Generation Genotoxicology: Using Modern Sequencing Technologies to Assess Somatic Mutagenesis and Cancer Risk. *Environmental and Molecular Mutagenesis*, *61*(1), 135–151. https://doi.org/10.1002/em.22342

Sanchez-Contreras, M., Sweetwyne, M. T., Tsantilas, K. A., Whitson, J. A., Campbell, M. D., Kohrn, B. F., Kim, H. J., Hipp, M. J., Fredrickson, J., Nguyen, M. M., Hurley, J. B., Marcinek, D. J., Rabinovitch, P. S., & Kennedy, S. R. (2023). The multi-tissue landscape of somatic mtDNA mutations indicates tissue-specific accumulation and removal in aging. *eLife*, *12*, e83395. https://doi.org/10.7554/eLife.83395

Schmitt, M. W., Kennedy, S. R., Salk, J. J., Fox, E. J., Hiatt, J. B., & Loeb, L. A. (2012). Detection of ultra-rare mutations by next-generation sequencing. *Proceedings of the National Academy of Sciences*, *109*(36), 14508–14513. https://doi.org/10.1073/pnas.1208715109

Shelby, M. D. (1994). Human germ cell mutagens. *Environmental and Molecular Mutagenesis*, *23*(S2), 30–34. https://doi.org/10.1002/em.2850230609

Shoag, J. E., Srinivasa, A., Loh, C. A., Liu, M. H., Lassen, E., Melanaphy, S., Costa, B. M., Grońska-Pęski, M., Jabara, N. T., Picciotto, S., Choi, U., Bohorquez, A. D., Barbieri, C. E., Callum, P., Skytte, A.-B., & Evrony, G. D. (2025). Direct measurement of the male germline mutation rate in individuals using sequential sperm samples. *Nature Communications*, *16*(1), 2546. https://doi.org/10.1038/s41467-025-57507-0

Shojaeisaadi, H., Schoenrock, A., Meier, M. J., Williams, A., Norris, J. M., Palmer, N. D., Yauk, C. L., & Marchetti, F. (2024). Mutational signature analyses in multi-child families reveal sources of age-related increases in human germline mutations. *Communications Biology*, *7*(1), 1451. https://doi.org/10.1038/s42003-024-07140-2

Smith-Roe, S. L., Hobbs, C. A., Hull, V., Todd Auman, J., Recio, L., Streicker, M. A., Rivas, M. V., Pratt, G. A., Lo, F. Y., Higgins, J. E., Schmidt, E. K., Williams, L. N., Nachmanson, D., Valentine III, C. C., Salk, J. J., & Witt, K. L. (2023). Adopting duplex sequencing technology for genetic toxicity testing: A proof-of-concept mutagenesis experiment with N-ethyl-N-nitrosourea

(ENU)-exposed rats. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *891*, 503669. https://doi.org/10.1016/j.mrgentox.2023.503669

Turker, M. S., Lasarev, M., Connolly, L., Kasameyer, E., & Roessler, D. (2007). Age-related accumulation of autosomal mutations in solid tissues of the mouse is gender and cell type specific. *Aging Cell*, *6*(1), 73–86. https://doi.org/10.1111/j.1474-9726.2006.00264.x

Vijg, J. (2021). From DNA damage to mutations: All roads lead to aging. *Ageing Research Reviews*, *68*, 101316. https://doi.org/10.1016/j.arr.2021.101316

White, P. A., Douglas, G. R., Phillips, D. H., & Arlt, V. M. (2017). Quantitative relationships between lacZ mutant frequency and DNA adduct frequency in Muta^{TM}Mouse tissues and cultured cells exposed to 3-nitrobenzanthrone. *Mutagenesis*, *32*(2), 299–312. https://doi.org/10.1093/mutage/gew067

White, P. A., Luijten, M., Mishima, M., Cox, J. A., Hanna, J. N., Maertens, R. M., & Zwart, E. P. (2019). In vitro mammalian cell mutation assays based on transgenic reporters: A report of the International Workshop on Genotoxicity Testing (IWGT). *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *847*, 403039. https://doi.org/10.1016/j.mrgentox.2019.04.002

Xavier, M. J., Mitchell, L. A., McEwan, K. E., Scott, R. J., & Aitken, R. J. (2018). Genomic integrity in the male germ line: Evidence in support of the disposable soma hypothesis. *Reproduction*, *156*(3), 269–282. https://doi.org/10.1530/REP-18-0202

Yauk, C. L., Lynch, A. M., Dobrovolsky, V. N., Schuler, M., Smith-Roe, S. L., Fitzgerald, D., Higgins, J., Honarvar, N., Le Curieux, F., Matsumura, S., Minocherhomji, S., Recio, L., Salk, J. J., Sugiyama, K.-I., Suzuki, T., Wills, J. W., & Marchetti, F. (2026). Application of error-corrected sequencing technologies for in vivo regulatory mutagenicity assessment. *Regulatory Toxicology and Pharmacology: RTP*, *164*, 105985. https://doi.org/10.1016/j.yrtph.2025.105985

Yizhak, K., Aguet, F., Kim, J., Hess, J. M., Kübler, K., Grimsby, J., Frazer, R., Zhang, H., Haradhvala, N. J., Rosebrock, D., Livitz, D., Li, X., Arich-Landkof, E., Shoresh, N., Stewart, C., Segrè, A. V., Branton, P. A., Polak, P., Ardlie, K. G., & Getz, G. (2019). RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science*, *364*(6444), eaaw0726. https://doi.org/10.1126/science.aaw0726

Zhu, J., Tsai, H.-J., Gordon, M. R., & Li, R. (2018). Cellular Stress Associated with Aneuploidy. *Developmental Cell*, *44*(4), 420–431. https://doi.org/10.1016/j.devcel.2018.02.002

CHAPTER TWO

DATA CHAPTER
Elena Esina

Box 2.A: Publication & Permissions

This chapter adapts:

> Esina, E., Dodge, A. E., Williams, A., Schuster, D. M., LeBlanc, D. P. M., Marchetti, F., & Yauk, C. L. (2024). Power analyses to inform Duplex Sequencing study designs for MutaMouse liver and bone marrow. Environmental and Molecular Mutagenesis, 65(8), 234–242. https://doi.org/10.1002/em.22619

Changes from the published version: condensed general DS background to avoid duplication; thesis formatting; renumbered figures/tables; minor caption edits.

For author contributions, see the front-matter "Statement of Contributions" (p. xii).

Funding & acknowledgments (paper):

Supported by the NSERC Discovery Grant (RGPIN-2021-02806) with support from the CFI John Evans Leaders Fund (#233109), the Ontario Research Fund (#40569) and uOttawa; Canada Research Chairs Program (CRC-2020-00060); Burroughs Wellcome Innovations in Regulatory Science Award (#1021737); Health Canada Genomics R&D Initiative and Chemicals Management Plan. Thesis-level personal acknowledgments appear in the front matter (p. iv).

Competing interests:

The authors declare no competing interests.

2.1 Abstract

Regulatory genetic toxicology testing is essential for identifying potentially mutagenic hazards. Duplex Sequencing (DS) is an error-corrected next-generation sequencing technology

that provides substantial advantages for mutation analysis over conventional mutagenicity assays including: improved accuracy of mutation detection, ability to measure changes in mutation spectrum, and applicability across diverse biological models. To apply DS for regulatory toxicology testing, power analyses are required to determine suitable sample sizes and study designs. In this study, we explored study designs to achieve sufficient power for various effect sizes in chemical mutagenicity assessment. We collected data from MutaMouse bone marrow and liver samples that were analyzed by DS using TwinStrand's Mouse Mutagenesis Panel. Average duplex reads achieved in two separates studies on liver and bone marrow were 8.4×10^8 ($\pm 7.4 \times 10^7$) and 9.5×10^8 ($\pm 1.0 \times 10^8$), respectively. Baseline mean mutation frequencies (MF) were 4.6×10^{-8} ($\pm 6.7 \times 10^{-9}$) and 4.6×10^{-8} ($\pm 1.1 \times 10^{-8}$), with estimated standard deviations for the animal-to-animal random effect of 0.15 and 0.20, for liver and bone marrow, respectively. We conducted simulation analyses based on these empirically derived parameters. We found that a sample size of four animals per group is sufficient to obtain over 80% power to detect a two-fold change in MF relative to baseline. In addition, we estimated the minimal total number of informative duplex bases sequenced with different sample sizes required to retain power for various effect sizes. Our work provides foundational data for establishing suitable study designs for mutagenicity testing using DS.

2.2 Introduction

Conventional *in vivo* mutagenicity tests rely on individual reporter genes for mutation quantification. The most widely used mutagenicity test is the transgenic rodent (TGR) assay (OECD, 2022), which relies on bacterial reporter genes for mutation quantification by using viral phages to encapsulate these genes and infect bacteria; mutations are identified by plaque formation in bacterial lawns grown under selective conditions (Lambert et al., 2005). These reporter-based assays have several important limitations (see Chapter 1, §1.3.1), so development of methods to circumvent these limitations would allow for better quantification and understanding of mutagenesis to protect human health.

Duplex Sequencing (DS), an error-corrected next-generation DNA sequencing technology, provides an alternative method to quantify mutations that addresses the limitations of reporter-based assays (Valentine et al., 2020). General DS background is condensed here to avoid duplication; see Chapter 1, §1.4 for the consolidated overview.

DS has been effective at measuring dose-dependent increases in MF following exposure to well-established mutagens. Prior research has analyzed benzo[a]pyrene (LeBlanc et al., 2022), procarbazine hydrochloride (Dodge et al., 2023), N-nitrosodiethylamine (Bercu et al., 2023), and N-ethyl-N-nitrosourea (Smith-Roe et al., 2023) exposures in rodent models. These studies demonstrate the potential for DS in mutagenicity testing. However, every new assay must be evaluated to determine the minimum detectable "signal" and the optimal approach to detect it, given such a signal exists. Power analyses are necessary to establish acceptable study design parameters to advance the use of DS in mutagenicity assessment.

Power is the probability that the null hypothesis can be correctly rejected given the sample size and effect size of the study. An underpowered study is more likely to miss true effects resulting in misleading conclusions, whereas an overpowered study may use more animals than necessary raising ethical concerns. Typically, 80% power is considered standard and what researchers aim for in their studies (Festing, 2018); it indicates an 80% chance of detecting a true effect if it exists and a 20% chance of missing it, balancing ethical and practical considerations with the risk of overlooking true effects.

We conducted simulation studies to evaluate the sample size and required number of informative duplex bases sequenced per sample for different effect sizes relevant to analyses of mouse bone marrow (BM) and liver. These tissues are widely used in regulatory mutagenicity testing and have very different characteristics. BM is highly proliferative, allowing DNA damage to be quickly converted into mutations and fixed in the genome. In contrast, liver is only moderately proliferative, but is the most metabolically active tissue and the first pass organ for detoxication. We simulated different study designs to evaluate how power changes with different parameters and draw conclusions on feasible sample size and depth of sequencing.

2.3 Methods

2.3.1 Baseline data

The DS data used in this study are from two studies on the livers and BM of MutaMouse males exposed daily for 28 days by oral gavage to olive oil (vehicle control: VC) or toxicants of interest (Schuster et al., 2024; LeBlanc et al., in preparation). These studies will henceforth be referred to as "Study A" and "Study B." The methodology for DNA extraction, DS library preparation, and DS data analysis was as described previously (Dodge et al., 2023), except that DNA was enzymatically fragmented, not ultrasonically sheared, during library preparation. The computational analysis of raw DS data was handled by TwinStrand via their DNAnexus® pipeline. The protocol for this pipeline is detailed by Kennedy et al. (2014) and the underlying code published (Kennedy-Lab-UW/Duplex-Seq-Pipeline, 2020/2023).

The data from Study A and Study B were used to identify a reasonable estimate of baseline MF and the number of informative duplex bases retrieved for mouse BM and liver DNA analyzed with a 500 ng DNA input by DS. The standard deviations used for the power analysis cover a range of values observed in the literature and in our own studies. In BM, these values were 0.06 (Dodge et al., 2023; LeBlanc et al., 2022), 0.10 (Study A), and 0.27 (Study B). In liver, standard deviations were 0.14 (Study A and Study B). There are three other papers that have used DS for mutagenicity assessment in chemically treated rodents (Bercu et al., 2023; Smith-Roe et al., 2023; Valentine et al., 2020). Unfortunately, none of these papers report the standard deviation on their MF. Thus, we use the values from Dodge, LeBlanc and those presented herein as reasonable estimates of possible standard variations to be used for the calculations. To calculate standard deviation, a generalized linear mixed model (GLMM) with a binomial error distribution was used to fit the data, incorporating fixed effects (treatment groups, targets, and their interaction) and random effects (experiment and animal ID) to account for overdispersion and estimate animal-to-animal variation using variance component analysis (Dertinger et al., 2023). These variance component estimates served as the basis for the range of standard deviations used in the simulation experiments for power analyses.

2.3.2 DS data interpretation

Mutation scoring in the studies used was done under the assumption that all identical mutations within a sample are the result of clonal expansion from a single mutational event and so are counted once (Dodge et al., 2023). Germline mutations (with variant allele frequency (VAF) >0.01) were filtered out and only true *de novo* mutations were retained. Additionally, the DNAnexus® pipeline removes missing values based on whether a base can be confidently identified for each position.

Sequencing results also give a measurement for the total number of informative duplex bases sequenced for each sample. These values can be converted into a target number of required raw reads. The number of unprocessed reads roughly making up one duplex read was empirically derived by Kennedy et al. (2014). These authors determined that six raw reads are sufficient to make an SSCS and that an average of six SSCSs is needed to make one DCS. Therefore, ideally, we need 36 raw reads to make one DCS. Four more raw reads were added to account for error and low-quality reads that fail to form duplex sequences, making a total of 40 raw reads to make a DCS. Thus, the following formula can be used to convert duplex bases into the number of raw reads needed to achieve a certain power:

$$\mathrm{N} = \frac{40\mathrm{DG}}{\mathrm{R}} \tag{1}$$

where N is the number of raw paired-end reads per sample, D is the average depth of coverage, G is the genome target size in base pairs, and R is the post-analysis read length in bases (300 bp for all of the datasets). The multiplication of DG is the number of informative duplex bases sequenced. Thus, this formula can be simplified to:

$$\mathrm{N} = \frac{40\mathrm{S}}{\mathrm{R}} \tag{2}$$

where S is the total number of informative duplex bases sequenced per sample.

2.3.3 Power analysis

Overdispersion occurs when the observed variance is greater than the variance expected under a standard binomial distribution, often due to heterogeneity among the

subjects (animals). GLMM provides a suitable model to estimate this variability, as an input to the simulation for the power analysis. Gelman and Hill (2007) described an approach similar to the one we took for our power analyses, albeit in a more complex multivariate context. We applied a simplified version to simulate univariate overdispersed binomial data using a random intercept model. The first step of simulation-based power analysis was to simulate a large number of data sets. When simulating responses from an overdispersed binomial distribution, we only needed to provide an estimate of the MF and an estimate of the animal-to-animal variability; these were derived from the results of the GLMM from studies A and B or from previously published studies (Dodge et al., 2023; LeBlanc et al., 2022). Estimates for these parameters, or a range of plausible values, are often available from previous studies or from a pilot study. If no data are available, a plausible range of parameter values is justified based on prior knowledge of the study system.

To generate random samples from an overdispersed binomial distribution in R (R: The R Project for Statistical Computing, n.d.; version 4.3.1) using the rbinom() function, three arguments were needed: the individual MF for each simulated observation adjusted for tissue variability (p_i), the number of informative duplex bases (or trials), and the number of random samples (n). Since we were generating only one observation for each sample, i would run from i = 1, 2, … n. The individual pi's were obtained using the following model:

$$\mathrm{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \mu + Z_i \tag{3}$$

where $\mu = \ln(\mathrm{MF})$ and Z_i was the overdispersion mechanism with the following distribution $Z_i \sim N(0, \sigma^2)$. Thus, Z_i modified the MF based on the sample standard deviation for a particular tissue. Equation (3) could be re-written to closely approximate p_i:

$$p_i = \frac{\exp(\mu + Z_i)}{1 + \exp(\mu + Z_i)} \tag{4}$$

As $1 + \exp(\mu + Z_i) \approx 1$ for small MFs, the equation could be further simplified to:

$$p_i \approx \exp(\mu + Z_i) = \mathrm{MF} \times \exp(Z_i) \tag{5}$$

To simulate mutation data for the treatment group, these p_i's were multiplied by a specified fold change:

$$p_i \approx FC \times MF \times \exp(Z_i) \quad (6)$$

with FC representing the desired fold change or effect size.

The rnorm() function was used to randomly generate realizations for the Z_i random variable. The rnorm() function required three parameters: the required sample size n which was set to 1, the mean which was set to zero, and the standard deviation which was set to be the animal-to-animal variation parameter determined by the simulation scenario.

We performed a power analysis at varying sample sizes and total number of informative duplex bases and degree of animal-to-animal variation. Simulated data for hypothetical treatment groups were randomly generated to produce defined increases in MF (e.g., 1.2-fold, 1.5-fold, and 2-fold). A generalized linear model (GLM) was used to calculate significance using the Wald statistic estimated using the doBy R library (Halekoh & Højsgaard, 2024; version 4.6.19). Experiments were simulated at least 1000 times for each scenario and the power was estimated as the frequency of experiments resulting in a significant difference between simulated treated and control groups at a significance level of 0.05.

2.4 Results and Discussion

The experimental DS data used for the power analysis are summarized in Table 2.1. The four datasets are described by the following parameters:

- **The number of test animals used (sample size) per dose group.** Study A and Study B used 4 animals per dose group. The data from these experiments were combined and simulated to produce different sample sizes for the power analysis (described below).
- **Background MF.** The VC MF was 5.33×10^{-8} ($\pm 7.35 \times 10^{-9}$) and 3.79×10^{-8} ($\pm 3.35 \times 10^{-9}$) mutations per bp for liver in Study A and Study B, respectively, and 4.13×10^{-8} ($\pm 1.35 \times 10^{-9}$) and 5.02×10^{-8} ($\pm 1.69 \times 10^{-8}$) mutations per bp for BM in Study A and Study B, respectively. The average pooled VC MF (i.e., combining both studies; n = 8 per tissue)

for liver and BM was 4.56×10^{-8} ($\pm 6.75 \times 10^{-9}$) and 4.58×10^{-8} ($\pm 1.14 \times 10^{-8}$) mutations per bp, respectively. As the averages for both tissues are very similar, an overall average of 4.57×10^{-8} mutations per bp was used in the subsequent power analyses.

- **Informative duplex bases sequenced per sample.** The number of raw reads for each dataset was 3.57×10^8, 2.47×10^8, 3.48×10^8, and 3.36×10^8 for liver and BM in Study A and B, respectively. The average total number of informative duplex bases achieved for each sample was 8.34×10^8 ($\pm 9.21 \times 10^7$) and 8.51×10^8 ($\pm 6.05 \times 10^7$) duplex bases in the liver studies, and 8.32×10^8 ($\pm 7.51 \times 10^7$) and 1.12×10^9 ($\pm 7.39 \times 10^7$) duplex bases in the BM studies. Note that these numbers are lower that what is expected based on the conversion formula Equation (2) because the studies achieved peak tag family sizes that were larger than the optimal study described therein; peak tag family sizes in these experiments was a minimum of 9.

Table 2.1: Overview of the four DS datasets that served as the foundation for the simulation studies.

Tissue	Study	Sample size per dose group[a]	Average background MF[b] mutations per bp (*± SE*)	Average informative duplex bases per sample[c] (*± SE*)
Liver	A	4	5.33 x 10^{-8} *(± 7.35 x 10^{-9})*	8.51 x 10^8 *(± 6.05 x 10^7)*
	B	4	3.79 x 10^{-8} *(± 3.53 x 10^{-9})*	8.34 x 10^8 *(± 9.21 x 10^7)*
Bone marrow	A	4	4.13 x 10^{-8} *(± 1.35 x 10^{-9})*	8.32 x 10^8 *(± 7.51 x 10^7)*
	B	4	5.02 x 10^{-8} *(± 1.69 x 10^{-8})*	1.12 x 10^9 *(± 7.39 x 10^7)*

[a] Male mice were exposed for 28 days by oral gavage to olive oil (vehicle control) or toxicants of interest (Study A and Study B).

[b] MF of the vehicle controls in each study.

[c] Number of informative duplex bases is averaged across all samples used in the experiments (i.e., both exposed and vehicle controls).

These empirical parameters were used to generate a range of hypothetical characteristics for potentially feasible DS experiments: sample sizes of 3–8 per dose group; a background MF of 4.57×10^{-8} mutations per bp consistent across liver and BM; an average total

informative duplex bases sequenced per animal of up to 2×10^9 duplex bases; a fold change of 1.5 from the VC group; and sample standard deviations of 0.1, 0.15, 0.2, and 0.3. The standard deviation in this context represents the sample variability of MF on the ln (natural logarithm) scale as the logit link function was used in the GLMM to model the probability of observing a mutation. For instance, an MF of 4.57×10^{-8} is −16.90 on the ln scale; with a standard deviation of 0.3, the generated adjusted MF for each observation (p_i) would then randomly fall anywhere within the range of (−16.90 × 0.97 = −16.39) to (−16.90 × 1.03 = −17.41) based on the density distribution provided in Figure 2.1. This interval translates back to an MF range of 2.75×10^{-8} to 7.59×10^{-8}. Figure 2.1 also shows the distribution of the animal-to-animal random effect for the other standard deviation values as well. As the chosen test parameters are based on empirical data, they enhance the biological relevance of the analysis.

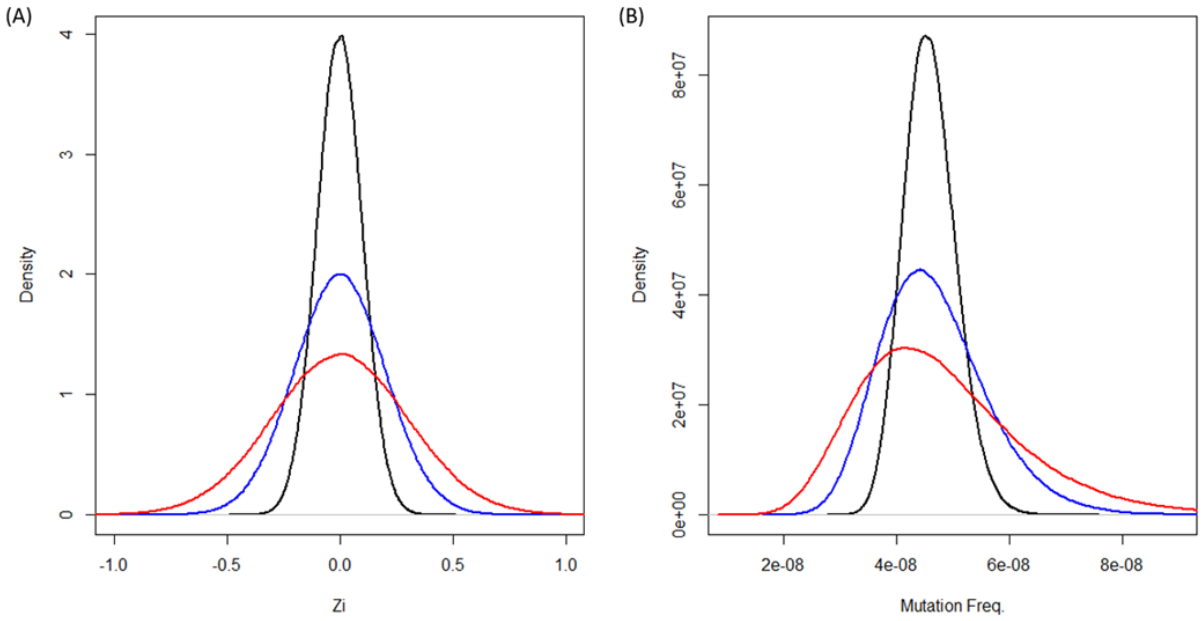

Figure 2.1: The distribution of the animal-to-animal variability for standard deviations 0.1 (black), 0.2 (blue) and 0.3 (red), or $\mathbf{N(0, 0.1^2)}$, $\mathbf{N(0, 0.2^2)}$, and $\mathbf{N(0, 0.3^2)}$ (A), adjusted to the individual MFs (i.e., $\mathbf{MF \times \exp(Z_i)}$) (B). Kernel density estimation was conducted using default Gaussian kernel and bandwidth for univariate observations. Z_i is the overdispersion mechanism with the following distribution $Z_i \sim N(0, \sigma^2)$.

Our study investigated MFs and sequencing depths produced in the analyses of BM and liver in MutaMouse males. These values may vary for other models, tissues, and laboratories. For example, Dodge et al. (2023) and LeBlanc et al. (2022) used sonication for DNA fragmentation and reported higher control MFs (e.g., 1.31×10^{-7} mutations per bp reported by Dodge et al. (2023) versus our 4.57×10^{-8} mutations per bp). Dodge et al. (2023) and LeBlanc et al. (2022) sequenced averages of 7.98×10^{8} and 8.50×10^{8} duplex bases, respectively, which are within our typical range. However, Smith-Roe et al. (2023) sequenced 5.36×10^{8} duplex bases and 4.89×10^{8} duplex bases in their study, which is below our typical range. Laboratories that achieve lower sequencing depths may need more samples to achieve sufficient power, whereas those with higher depths could require fewer samples. The general power analysis approach presented herein is extendable to other parameters, and other researchers can readily apply the code to determine sample sizes to maintain 80% power for their systems.

Our study focused solely on biological variability measured in-house. Comprehensive studies on analytical and technical variability are forthcoming; preliminary data suggest that this variability is negligible with DS. For instance, results showed remarkable reproducibility across two laboratories that sequenced samples from two identically-designed in vitro studies at the same facility (Cho et al., 2023) and in the same samples prepared and sequenced at two different locations (LeBlanc et al., 2022). In both cases, DS produced highly similar results, indicating very low technical variability.

In our study, we first simulated data to determine power to detect a 1.5-fold change in MF between control and exposed groups as the test parameters vary (Figure 2.2). We found that when the sample standard deviation is 0.1, 3 animals per dose group is insufficient to achieve 80% power with 2×10^{9} informative duplex bases sequenced per animal—the maximum investigated in this study. However, increasing sample size to four achieves over 95% power with this number of informative duplex bases and 80% power with 1×10^{9} informative duplex bases. Increasing the sample size further decreases the required number of duplex bases for any given power value. When the sample standard deviation is 0.15, 80% power is achieved with a minimum sample size of five animals per dose group and 1×10^{9} informative duplex bases. With a standard deviation of 0.2, 80% power is achieved with a minimum sample

size of 6 animals per dose group and 1.5×10^9 informative duplex bases; lower samples sizes yield <71% power. When the sample standard deviation is 0.3, 80% power is not achieved with any combination of sample size and numbers of duplex bases investigated here. Overall, increasing either the sample size or the number of informative duplex bases sequenced per animal increases the power of the study design. However, although larger sample sizes notably enhance power, increasing the number of informative duplex bases beyond about 1.2×10^9 duplex bases yields minimal additional power. This convergence occurs when you have sequenced deeply enough to have an accurate measure of changes in MF. Thus, increasing the number of samples has more impact on power than duplex depth by this point. Our analysis sheds light on the dynamics of statistical power as sample size and sequencing depth vary.

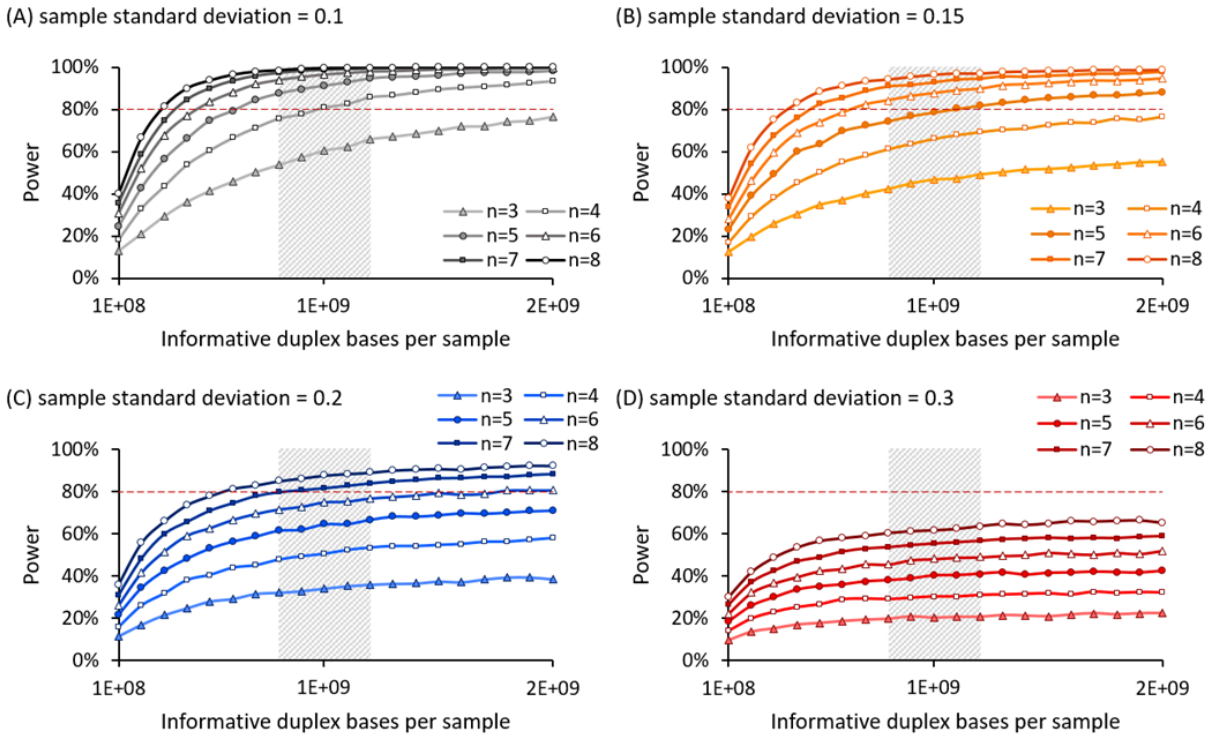

Figure 2.2: The estimated power to detect a 1.5-fold change in MF between control and exposed groups, with sample sizes of 3-8 animals per group, and sample standard deviations of 0.1 (A), 0.15 (B), 0.2 (C), and 0.3 (D). The grey highlighted zone shows the typical number of informative duplex bases sequenced per sample for our laboratory (between 8.0×10^8 and 1.2×10^9 duplex bases).

We ran additional modeling experiments to determine the minimum sequencing depth and sample size needed to observe a significant increase in MF with a power of 80% and a p-value of <0.05 for different effect sizes and variances. For this analysis, the investigated parameter range was expanded: sample sizes were varied from 2 to 20 animals per dose group; background MF was kept constant at 4.57×10^{-8} mutations per bp; the average total number of informative duplex bases sequenced per animal was modeled up to 2×10^9 duplex bases; sample standard deviations ranged from 0.05 to 0.3, with increments of 0.05; and expected fold changes of 1.2, 1.5, and 2 were considered from the VC group to cover a wide range of responses. For each combination of sample size, variance, and expected fold change, we varied the number of informative duplex bases sequenced per animal until we found the minimum value needed to observe a significant increase in MF with a power of 80% (± 0.01%). This analysis provides a foundation to establish study designs required to achieve statistically significant results in diverse mutagenicity studies.

Our results demonstrate that several feasible study designs are suitable across the above range of parameters (Figure 2.3). A 1.2-fold change is only detectible with a minimum of 8 animals per dose group and an average of 1.6×10^9 informative duplex bases sequenced per sample if the sample standard deviation is extremely low—0.05. Further sample size increases lower this required number of informative duplex bases; however, even with as many as 20 animals per dose group, study designs with sample standard deviation >0.15 cannot detect 1.2-fold changes, given that the required number of duplex bases modeled did not exceed 2×10^9 informative duplex bases. Detecting a 1.5-fold change is more feasible. Studies with a sample standard deviation of 0.05 are sufficiently powered with at least 3 animals per dose group, requiring a corresponding 1.2×10^9 informative duplex bases per sample. Studies where the sample standard deviation is 0.3—the highest investigated variance—are only sufficiently powered with 12 animals per dose group and 8.9×10^8 informative duplex bases. Finally, detecting a 2-fold change requires a sample size of 2 and 1.2×10^9 informative duplex bases when the sample standard deviation is 0.05, and a sample size of 5 and 7.0×10^8 informative duplex bases when the sample standard deviation is 0.3. Overall, our analysis reveals study

designs that are amenable to detecting different fold changes using different sample sizes and sequencing depths.

The feasibility of various parameters in rodent DS studies needs to be considered when conducting power analyses. Some designs require more samples per dose group than ethically permissible to achieve sufficient power. One way to reduce sample size requirements is to increase the number of informative duplex bases sequenced per sample. However, with a fixed sample size, power tends to plateau with an increasing number of informative duplex bases (Figure 2.2). Where this happens depends on the characteristics of the dataset. Therefore, increasing the number of informative duplex bases to use smaller sample sizes might not

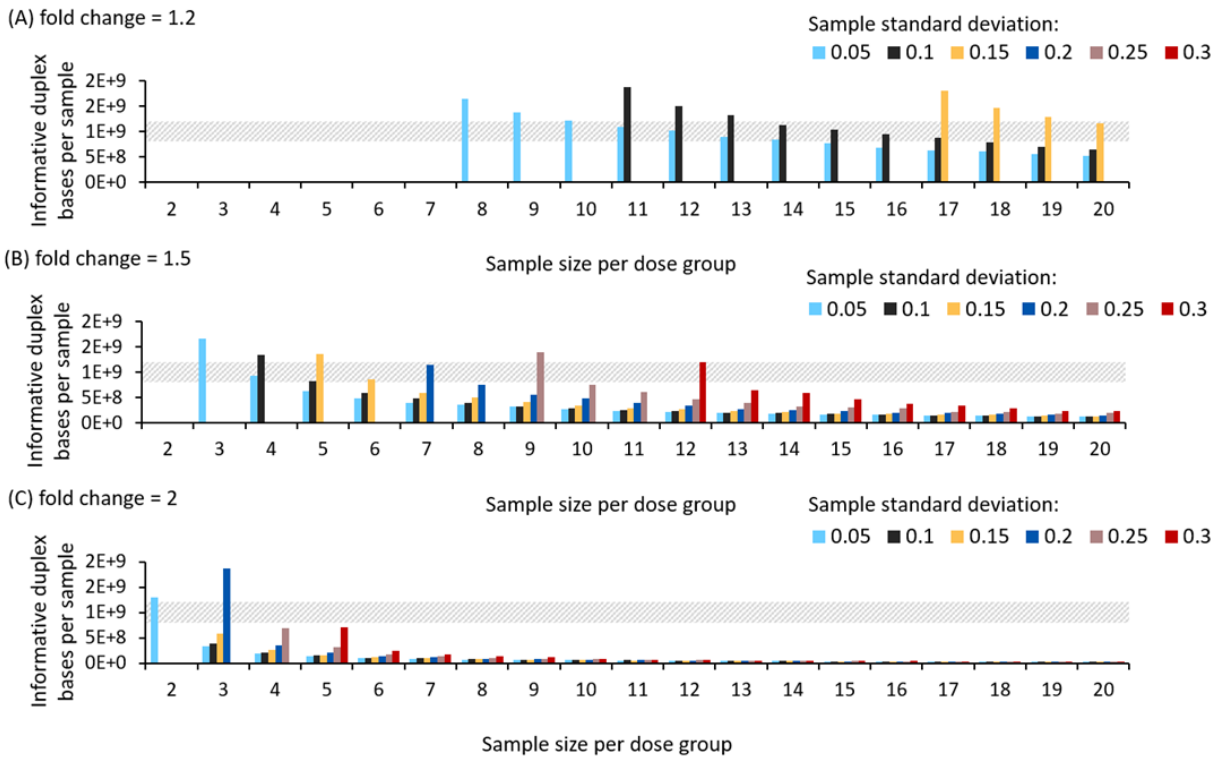

Figure 2.3: Minimum sequencing depth needed to detect a fold change of 1.2 (A), 1.5 (B), or 2 (C) given a background mutation frequency of 4.57 x 10^{-8} mutations per bp and a range of sample standard deviation. The grey highlighted zone shows the typical number of informative duplex bases sequenced per sample for our laboratory (between 8.0 x 10^8 and 1.2 x 10^9 duplex bases). No bar is shown when the required sequencing depth exceeds 2 x 10^9 duplex bases per sample.

always be possible. Thus, the trade-off between increasing sample size or the number of informative duplex bases must be carefully considered when designing studies to accommodate high sample deviations and small effect sizes.

Another consideration when increasing the number of informative duplex bases is the biological limit of how deeply we can sequence our samples. There is a limited availability of unique DNA molecules in a given amount of input DNA—known as DNA library complexity. Complexity must be balanced against the number of informative duplex bases achieved. Starting with a more complex sample (i.e., with a larger amount of input DNA) would allow us to increase the number of informative duplex bases. This area presents an opportunity for future research exploration to determine optimal DNA inputs and the number of informative duplex bases for specific DS applications.

Our analysis enabled us to consider acceptable study designs based on two main factors: the minimum sample size per dose group and the most feasible number of informative duplex bases sequenced per sample. In our laboratory, we typically sequence to a depth of 1.2×10^9 informative duplex bases per sample, which is economically feasible at present. Since the required number of informative duplex bases per sample and sample size are inversely correlated, several study designs that minimize sample size are outside the scope of what we can realistically sequence. Given this restriction, we considered study designs for DS that require $\leq 2.0 \times 10^9$ informative duplex bases (Figure 2.3). Using this restriction, detecting a 1.2-fold change in MF is challenging; large sample sizes are required to compensate for power loss due the lower number of informative duplex bases per sample and the low expected fold change in MF. In contrast, the detection of a 2-fold change is well within our scope. With a sample size of 4, we achieve sufficient power to analyze samples across most variances and at an easily attainable number of informative duplex bases. Increasing the sample size to 5 extends this capability to include samples with a standard deviation of 0.3. Overall, our results suggest that samples sizes of 4–5 are sufficient to detect at least a 1.5-fold change with a feasible number of informative duplex bases even with high sample deviation using the present enzymatic fragmentation protocol in MutaMouse liver and BM samples. The TGR test guideline indicates that a minimum of five mice per dose group should be used to ensure sufficient

statistical power to detect at least a 2-fold change in MF (OECD, 2022). Although most OECD mutagenicity tests are designed to be able to detect this level of change, DS can feasibly detect changes as small as 1.5-fold, making it the more sensitive technology. Thus, DS has a higher power to detect the required effect sizes recommended in mutation analyses than the TGR assay.

The parameter range investigated in this study did not encompass all possible values. Study parameters in other experiments may differ, potentially altering the recommended study designs. Sample standard deviation, background MF, and effect size are not under the control of the researcher; they are determined by the study compound characteristics and the biology of the organism. However, researchers can adjust sample size and sequencing depth to optimize the power of the study design. In addition, we calculated power for experiments with two experimental groups: control and treatment. However, the strategy we used is also applicable to studies with multiple treatment groups. The main difference in the analysis would be the application of a Bonferroni correction to adjust the significance level for multiple testing. Our results provide a foundation for strategic optimization of study designs for mutation analysis.

2.5 Conclusions

Study designs should be based on the biological characteristics of the test system and the objective of the mutagenicity assessment. With DS, ethical considerations about animal use must be balanced against sequencing (cost) feasibility, anticipated effect size, and group variances. Our power analysis provides a basis for study design selection for liver and BM tissues in MutaMouse animals. We show that a variety of study designs can work to detect fold changes of 1.5 and 2. However, the routine detection of a 1.2-fold change is likely beyond our present capabilities. Curation of historical controls will be critical for DS analyses across different tissues. This will enable us to achieve greater accuracy in our estimates of background MF and sample deviation. Once enough data are collected, a similar power analysis could be done for mutation spectra. Overall, our study informs sample size and the required number of

informative duplex bases needed per sample for DS study designs in MutaMouse BM and liver tissue to promote regulatory testing using this powerful new technology.

Box 2.B: Thesis Integration Note (non-peer-reviewed)

See Chapter 3, §3.3.1 for how these design rules set target depth and sample size.

2.6 References

Bercu, J. P., Zhang, S., Sobol, Z., Escobar, P. A., Van, P., & Schuler, M. (2023). Comparison of the transgenic rodent mutation assay, error corrected next generation duplex sequencing, and the alkaline comet assay to detect dose-related mutations following exposure to N-nitrosodiethylamine. Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 891, 503685. https://doi.org/10.1016/j.mrgentox.2023.503685

Cho, E., Swartz, C. D., Williams, A., Rivas, M., Recio, L., Witt, K. L., Schmidt, E. K., Yaplee, J., Smith, T. H., Van, P., Lo, F. Y., Valentine, C. C., Salk, J. J., Marchetti, F., Smith-Roe, S. L., & Yauk, C. L. (2023). Error-corrected Duplex Sequencing enables direct detection and quantification of mutations in human TK6 cells with remarkable inter-laboratory consistency (p. 2023.02.22.529418). bioRxiv. https://doi.org/10.1101/2023.02.22.529418

Dertinger, S. D., Li, D., Beevers, C., Douglas, G. R., Heflich, R. H., Lovell, D. P., Roberts, D. J., Smith, R., Uno, Y., Williams, A., Witt, K. L., Zeller, A., & Zhou, C. (2023). Assessing the quality and making appropriate use of historical negative control data: A report of the International Workshop on Genotoxicity Testing (IWGT). Environmental and Molecular Mutagenesis, n/a(n/a). https://doi.org/10.1002/em.22541

Dodge, A. E., LeBlanc, D. P. M., Zhou, G., Williams, A., Meier, M. J., Van, P., Lo, F. Y., Valentine Iii, C. C., Salk, J. J., Yauk, C. L., & Marchetti, F. (2023). Duplex sequencing provides detailed characterization of mutation frequencies and spectra in the bone marrow of MutaMouse males exposed to procarbazine hydrochloride. Archives of Toxicology, 97(8), 2245–2259. https://doi.org/10.1007/s00204-023-03527-y

Esina, E., Dodge, A. E., Williams, A., Schuster, D. M., LeBlanc, D. P. M., Marchetti, F., & Yauk, C. L. (2024). Power analyses to inform Duplex Sequencing study designs for MutaMouse liver and bone marrow. Environmental and Molecular Mutagenesis, 65(8), 234–242. https://doi.org/10.1002/em.22619

Festing, M. F. (2018). On determining sample size in experiments involving laboratory animals. Laboratory Animals, 52(4), 341–350. https://doi.org/10.1177/0023677217738268

Gelman, A., & Hill, J. (2007). Sample size and power calculations. In Data Analysis using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Halekoh, U., & Højsgaard, S. (2024). doBy: Groupwise Statistics, LSmeans, Linear Estimates, Utilities (Version 4.6.21) [Computer software]. https://cran.r-project.org/web/packages/doBy/index.html

Kennedy, S. R., Schmitt, M. W., Fox, E. J., Kohrn, B. F., Salk, J. J., Ahn, E. H., Prindle, M. J., Kuong, K. J., Shen, J.-C., Risques, R.-A., & Loeb, L. A. (2014). Detecting ultralow-frequency mutations by Duplex Sequencing. Nature Protocols, 9(11), 2586–2606. https://doi.org/10.1038/nprot.2014.170

Kennedy-Lab-UW/Duplex-Seq-Pipeline. (2023). [HTML]. Kennedy-Lab-UW. https://github.com/Kennedy-Lab-UW/Duplex-Seq-Pipeline (Original work published 2020)

Lambert, I. B., Singer, T. M., Boucher, S. E., & Douglas, G. R. (2005). Detailed review of transgenic rodent mutation assays. Mutation Research/Reviews in Mutation Research, 590(1), 1–280. https://doi.org/10.1016/j.mrrev.2005.04.002

LeBlanc, D. P. M., Meier, M., Lo, F. Y., Schmidt, E., Valentine, C., Williams, A., Salk, J. J., Yauk, C. L., & Marchetti, F. (2022). Duplex sequencing identifies genomic features that determine susceptibility to benzo(a)pyrene-induced in vivo mutations. BMC Genomics, 23(1), 542. https://doi.org/10.1186/s12864-022-08752-w

OECD. (2022). Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/environment/test-no-488-transgenic-rodent-somatic-and-germ-cell-gene-mutation-assays_9789264203907-en

R: The R Project for Statistical Computing. (n.d.). Retrieved May 7, 2024, from https://www.r-project.org/

Schuster, D. M., LeBlanc, D. P. M., Zhou, G., Meier, M. J., Dodge, A. E., White, P. A., Long, A. S., Williams, A., Hobbs, C., Diesing, A., Smith-Roe, S. L., Salk, J. J., Marchetti, F., & Yauk, C. L. (2024). Dose-related Mutagenic and Clastogenic Effects of Benzo[b]fluoranthene in Mouse Somatic Tissues Detected by Duplex Sequencing and the Micronucleus Assay (p. 2024.07.26.605228). bioRxiv. https://doi.org/10.1101/2024.07.26.605228

Smith-Roe, S. L., Hobbs, C. A., Hull, V., Todd Auman, J., Recio, L., Streicker, M. A., Rivas, M. V., Pratt, G. A., Lo, F. Y., Higgins, J. E., Schmidt, E. K., Williams, L. N., Nachmanson, D., Valentine III, C. C., Salk, J. J., & Witt, K. L. (2023). Adopting duplex sequencing technology for genetic toxicity testing: A proof-of-concept mutagenesis experiment with N-ethyl-N-nitrosourea (ENU)-exposed rats. Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 891, 503669. https://doi.org/10.1016/j.mrgentox.2023.503669

Valentine, C. C., Young, R. R., Fielden, M. R., Kulkarni, R., Williams, L. N., Li, T., Minocherhomji, S., & Salk, J. J. (2020). Direct quantification of in vivo mutagenesis and carcinogenesis using duplex sequencing. Proceedings of the National Academy of Sciences, 117(52), 33414–33425. https://doi.org/10.1073/pnas.2013724117

CHAPTER THREE

DATA CHAPTER
Elena Esina

3.1 Abstract

Age-related mutation accumulation is often treated as a general feature of aging, yet the mechanisms driving this process may differ substantially across tissues. Nuclear mutations can affect tissue function and, in germ cells, contribute to heritable genetic risk, making it important to understand how they arise across biological contexts. However, much of the existing evidence comes from systems that do not directly resolve endogenous nuclear mutagenesis across multiple tissues under controlled *in vivo* conditions. Here, we used Duplex Sequencing to quantify endogenous nuclear mutation frequency (MF), mutational spectra, and genomic distribution in liver, bone marrow, lung, and germ cells from MutaMouse males across age. MF was relatively similar across tissues early in life, with the unique-mutation estimate, MF_{Min}, remaining within ~3.6–4.6 × 10^{-8} mutations/base pair at 3, 14, and 23 weeks. At older ages, divergence was driven primarily by a soma-germline separation: liver, bone marrow, and lung showed broadly similar age-related increases in MF_{Min}, rising 1.5–2.0-fold between 3 and ≥78 weeks, whereas male germ cells did not show a significant age-related increase. Stratification by genic context preserved this compartmental pattern, with somatic tissues generally showing higher intergenic than genic MF overall (driven mainly by bone marrow and lung) while germ cells showed the opposite relationship. In contrast, stratification by mutation spectrum revealed tissue-specific subtype patterns even within somatic tissues, with liver showing the clearest subtype-level shifts and lung showing a more uniform increase across mutation classes. Collectively, age-related nuclear mutagenesis seems to be a tissue-dependent process shaped by the biological context in which mutations arise and persist rather than a uniform consequence of aging alone.

3.2 Introduction

Biological aging involves a range of cellular and molecular changes, but one of the most direct readouts of the cumulative history of DNA damage and repair is the accumulation of somatic mutations (Navarro et al., 2006; Yousefzadeh et al., 2021). Somatic mutation burden is thought to contribute to aging in at least two ways: by increasing transcriptional noise and progressively impairing cellular function in non-germline tissues, and by driving the clonal

expansion of mutant cells, which is associated with increased disease risk (Vijg, 2021). In germ cells, DNA damage and mutations can result in birth defects and heritable diseases in offspring (Shelby, 1994): mutations present in the zygote are propagated to essentially all somatic and germline cells of the offspring, where they can affect the individual's health and, if transmitted further, contribute to the spread of these variants in the population over successive generations. Thus, age-associated mutation accumulation can influence both the functional decline of tissues within an individual and the health of future generations. This is particularly concerning in the context of population aging and the growing trend toward delayed parenthood, as increasing numbers of people are reaching reproductive ages associated with poorer health and a higher burden of chronic disease. A central goal of aging research is therefore to understand how these mutations arise and accumulate over the lifespan and contribute to age-related functional decline.

Reporter-based aging studies provided the first experimental evidence that somatic mutation burden can increase across the lifespan, but they also highlighted the difficulty of interpreting MF from reporter systems alone. These studies show that reporter-gene MF tend to increase linearly with age and suggest that the rate and extent of mutation accumulation are tissue- and cell type-specific (Busuttil et al., 2007; Dollé et al., 1997, 2000; Fischer & Stringer, 2008; Kimoto et al., 2017; Ono et al., 2000; Turker et al., 2007) with mutations arising from both new mutation events and clonal expansion of pre-existing mutants (Fischer & Stringer, 2008; Kimoto et al., 2017; Ono et al., 2000; Turker et al., 2007). Tissues with higher mitotic activity typically show faster mutation accumulation (Busuttil et al., 2007; Dollé et al., 1997), although this relationship is not universal across all studies (Ono et al., 2000). Mutation spectra also show tissue- and cell type-specific variation (Busuttil et al., 2007; Dollé et al., 1997, 2000; Kimoto et al., 2017; Turker et al., 2007), and C:G>T:A transitions, often at CpG sites, are the predominant mutation type seen with aging (Busuttil et al., 2007; Dollé et al., 2000; Ono et al., 2000). However, these assays interrogate only a narrow, often exogenous portion of the genome, and mutation-spectrum analysis requires substantial follow-up work while recovering relatively few mutations. As a result, studies on reporter genes offer only a partial view of how

age-related mutations are distributed across endogenous nuclear loci and motivate sequencing-based approaches that directly interrogate native genomes.

Duplex Sequencing (DS) and related error-corrected sequencing (ECS) methods have been applied to both mitochondrial and nuclear genomes in aging tissues, although most prior work to date has focused on mitochondrial DNA (mtDNA). DS studies in mouse and human tissues show that mtDNA MF increases with age and that mutation burden varies across genomic regions, with non-coding regions such as the D-loop carrying higher MFs than coding regions (Arbeithuber et al., 2020; Kennedy et al., 2013; Sanchez-Contreras et al., 2023). However, it remains unclear to what extent age-related mutational patterns in mtDNA mirror those in nuclear DNA, given the profound biological and organizational differences between mitochondrial and nuclear genomes. Thus, findings from mtDNA cannot substitute for direct analyses of endogenous nuclear mutagenesis.

Recent ECS studies in humans have provided the first genome-scale insights into age-related nuclear mutation burden, signatures, and, in some tissues, clonal dynamics, but they have not systematically compared these features across multiple tissues under controlled *in vivo* conditions. NanoSeq (Abascal et al., 2021), a genome-scale ECS method, has been applied to quantify age-related mutation accumulation across human somatic tissues and the male germline, and these studies suggest that nuclear mutations accumulate approximately linearly with age (Abascal et al., 2021; Neville et al., 2025; Shoag et al., 2025). Decomposition of the resulting somatic mutation spectra into COSMIC "mutational signatures"—characteristic patterns of base substitutions in specific sequence contexts that reflect the activity of underlying mutational processes or DNA repair pathways (Alexandrov et al., 2013)—shows that age-associated spectra are dominated by SBS1 and SBS5, two clock-like signatures whose contributions increase with age. However, because human studies are inherently observational and shaped by inter-individual differences in genetics, environment, lifestyle, and exposure history, they are less able to isolate age as the primary variable. These studies provide important insight into age-related mutagenesis in humans, but they leave unresolved how aging shapes endogenous nuclear mutagenesis across tissues under controlled conditions.

What remains largely missing are comparative multi-tissue *in vivo* studies of endogenous nuclear mutagenesis in controlled mammalian models. Here, we used DS in MutaMouse males to compare age-related nuclear mutation burden, spectra, and genomic distribution across liver, bone marrow (BM), lung, and male germ cells. These tissues span high- and low-proliferation contexts, and somatic and germline compartments, representing distinct biological contexts in which age-related nuclear mutations may arise and persist. This multi-tissue design thus enables us to test how aging and tissue biology jointly shape nuclear mutation accumulation.

3.3 Methods

3.3.1 Animal handling and tissue collection

MutaMouse male mice, a CD2F1 (BALB/c × DBA/2) transgenic mutation-reporter strain, were used in this study. Animals from the Health Canada MutaMouse colony were sampled at five representative ages: 3 weeks (weaning; "young"), 14, 23, and 36 weeks ("mature"), and ≥78 weeks ("old"). All animal procedures were approved by the Health Canada Ottawa Animal Care Committee and conducted in accordance with the guidelines of the Canadian Council on Animal Care. At euthanasia, liver, femur BM, lung, and germ cells from the seminiferous tubules (ST) were harvested. Germ cell suspensions were prepared from seminiferous tubules as previously described (O'Brien et al., 2014). All tissues were immediately flash frozen in liquid nitrogen and stored at −80 °C until use.

Samples from pre-existing DS studies were included in the dataset to supplement the age series for each tissue. Specifically, liver and BM libraries from intermediate ages reported by Schuster et al. (2024), lung libraries from an ongoing study (Smith-Roe et al., submitted), and germ-cell libraries from an independent study (Stewart et al., 2025) were combined with newly generated libraries from the present work to assemble the final age series for liver, BM, lung, and germ cells. Sample size per tissue and age group is summarised in Table 3.1.

Animals in the 14-, 23-, and 36-week cohorts were vehicle-control groups from previous studies (Schuster et al., 2024; Smith-Roe et al., submitted; Stewart et al., 2025) and received

daily oral gavage with olive oil for 28 days (14-week cohort), 90 days (23-week cohort), or 180 days (36-week cohorts). The 3-week cohort consisted of pups born to dams that had received olive oil vehicle by oral gavage for 10 days as part of another study (unpublished). In contrast, the ≥78-week cohort comprised untreated colony animals that were maintained until they reached predetermined humane endpoints; because these animals did not all survive to the planned 2-year time point, their ages at euthanasia ranged from 78 weeks to 85 weeks, and the group is thus referred to as "≥78 weeks".

Sample size for this study was informed by a previous power analysis of DS experiments in MutaMouse liver and BM (Chapter 2), which indicated that n=8 animals per age group provided adequate power to detect a 1.5-fold change with a feasible sequencing depth. We therefore aimed to sequence up to eight male mice per age group for each tissue. Final sample sizes per tissue and age group ranged from 4 to 8 animals because some samples did not meet inclusion criteria, had poor DNA quality, or failed DS library preparation or downstream quality-control checks (Summarized in Table 3.1).

Table 3.1: Number of MutaMouse male per tissue and age group.

Age groups (weeks)	**Sample sizes**			
	Liver	BM	Lung	Germ cell
3	8	8	7	8
14	8 [a]	8 [a]	7 [d]	8
23	8 [a]	8 [a]	7 [d]	6 [b]
36	8 [a]	8 [a]	7 [d]	7 [c]
≥78	8	6	4	8

[a] *n=4 from (Schuster et al., 2024)*
[b] *n=6 from (Stewart et al., 2025)*
[c] *n=5 from (Stewart et al., 2025)*
[d] *n=4 from (Smith-Roe et al., submitted)*

3.3.2 Duplex Sequencing

DNA was extracted from tissue samples using the DNeasy Blood & Tissue Kit (Cat. #69504, Qiagen, Hilden, Germany) according to the manufacturer's instructions. DS libraries were prepared from 500 ng of genomic DNA per sample following the TwinStrand Biosciences DS protocol (version 1.1 described by Valentine et al., 2020) on a panel of 20 x 2.4 kb loci

spread across the mouse autosomes (described previously in LeBlanc et al., 2022). Briefly, DNA was enzymatically fragmented to a mean fragment size of 300 base pairs (bp), end-polished, A-tailed, ligated to DS sequencing Adapters™(TwinStrand Biosciences, Seattle WA, USA), and PCR-amplified. Libraries were pooled and sequenced on a NovaSeq 6000 system (Illumina, San Diego, CA, USA) by Psomagen Inc. (Rockville, MD, USA) or by the Ottawa Health Research Institute StemCore Laboratory (Ottawa, ON, Canada). All libraries met the target of approximately 500 million duplex bases per sample and a minimum duplex coverage of 10,000x over the target loci. The sequencing FASTQ files used in the current study are available in the NCBI Sequence Read Archive: PRJNA1131313 (Schuster et al., 2024), PRJNA1228006 (Stewart et al., 2025), and PRJNA1358256 (Smith-Roe et al., submitted).

Per-library FASTQ files were analyzed using the TwinStrand DuplexSeq Mutagenesis App (version 4.7.0; TwinStrand Biosciences, Seattle, WA, USA), as described in Valentine et al. (2020). For each library, the App reports sequencing quality metrics and generates a variant call file (VCF) and sample summary tables of MF and base substitution spectra. In brief, the pipeline identifies and extracts unique molecular identifier (UMI) sequences, corrects likely errors in UMI sequences to improve read grouping, and aligns raw reads to the mouse reference genome (mm10). Reads are then grouped by UMI and strand-defining sequence, trimmed for base quality, and collapsed into double-stranded consensus reads to suppress PCR and sequencing errors. Consensus reads undergo post-processing and re-alignment, followed by variant calling to produce the per-library .mut files for downstream analyses.

3.3.3 Mutation data interpretation and statistical analysis

All analyses were performed in R (version 4.4.1), except when specified otherwise, using the MutSeqR (Dodge et al., 2025; version 0.99.0) package for mutation analysis.

Quality control and variant filtering

Raw variant calls were subjected to several quality and biological filters before MFs were computed.

First, MutSeqR's filter_mut() function was used to identify variants within the Mouse Mutagenesis Panel (TwinStrand Biosciences, Seattle, WA, USA) that have variant allele fraction (VAF) ≥ 1% and pass MutSeqR's internal quality flags, including checks for end-repair artefacts. No variants in this dataset were flagged as end-repair artefacts, consistent with the use of enzymatic fragmentation rather than sonication in the updated DS library-preparation pipeline. Single nucleotide variants (SNVs) overlapping germline multi-nucleotide variants (MNVs; VAF ≥ 1%) were excluded, as recommended by the duplex sequencing platform provider (TwinStrand Biosciences). In addition, we excluded a recurrent homopolymer hotspot on chromosome 11 (chr11:37936714–37936764, mm10/GRCm38) that showed clear evidence of sequencing artefact (see §3.4.3 for details).

To ensure that our analyses focused on somatic mutations rather than inherited (germline) variation, we combined a VAF-based classifier with dbSNP annotation to identify constitutional (germline-like) sites. Each variant was annotated against dbSNP by querying a tabix-indexed dbSNP VCF (GRCm38) and recording exact coordinate matches. For each sample, variants with >1% VAF and/or overlapping with a dbSNP site were classified as constitutional, while the remaining variants were retained for somatic analysis. We then applied this classification across the cohort: if a given position was classified as constitutional in at least one mouse, it was treated as constitutional in all mice and excluded from the somatic analysis set. All downstream analyses were restricted to variants that passed all filters; these are hereafter referred to as somatic mutations.

Mutation frequency calculation

MF was calculated under two alternative assumptions: (1) all identical somatic mutations within a sample were assumed to be the result of clonal expansion from a single mutational event and were counted once (MF_{Min}), preventing potential overestimation of MF due to artificial increases from clonal expansion; or (2) all identical somatic mutations were assumed to occur as independent mutational events and were counted separately (MF_{Max}), protecting against underestimating the frequency of true mutations. Because in practice it is not possible to know which assumption is correct (and the truth may be a mixture of both), we computed and analysed MF_{Min} and MF_{Max} in parallel.

The MutSeqR function calculate_mf()was used to create sample-level somatic mutation counts. MF_{Min} and MF_{Max} for each group were computed as the number of mutations divided by the total interrogated depth for that group. As part of this calculation, calculate_mf() was used for position-level depth correction. The sample's age was retained as metadata and used as a predictor in downstream statistical modelling. Note that when the same genomic position appeared on multiple rows for a given sample (e.g., due to how variant records were aggregated), only one row was allowed to contribute coverage for that position and the depth for all additional rows was set to zero to prevent double-counting of coverage at the same site.

Statistical modelling framework of age effects on mutation frequency

MFs were analysed using binomial-family generalized linear models with a logit link, with the number of mutated reads modelled relative to total duplex depth. For fixed-effect models, overdispersion was assessed using the ratio of residual deviance to residual degrees of freedom, and quasibinomial models were used where appropriate. Where repeated observations from the same sample were included, mixed-effects models with a random intercept for sample were used. Previous work has suggested that mutation burden increases linearly with age (Kennedy et al., 2013; Ono et al., 2000; Sanchez-Contreras et al., 2023), but this assumption may not hold in all tissues. Accordingly, age was modeled either as a continuous variable (weeks) or as a categorical factor. Age was treated as continuous when a linear term provided an adequate fit, and as categorical when the categorical model provided a significantly better fit ($p < 0.05$). When age was modeled as a categorical factor, contrasts were used to compare each age group against the 3-week reference group, with p-values adjusted for multiple testing using the Holm-Šidák method.

For all age-effect analyses, mutation data were first summarized for each sample using calculate_mf(), and the resulting MF estimates were then analysed using MutSeqR::model_mf(). For overall mutation burden, both MF_{Min} and MF_{Max} were modeled as a function of age. For mutation spectra, MF_{Min} was analyzed by mutation subtype, including major mutation classes with SNVs further resolved into the six canonical substitution classes, to test whether age-related patterns differed., MF_{Min} was also analyzed by genic context (genic vs

intergenic) to test whether age effects differed across contexts. Where applicable, models included interactions between age and the grouping variable of interest.

Trinucleotide mutational signature analysis

For exploratory trinucleotide-context analysis, SNVs were summarized as SBS96 mutation spectra using SigProfilerMatrixGenerator (mm10) within each age group, and mutational signatures were evaluated using SigProfilerExtractor with comparison to COSMIC v3.4 signatures.

Post hoc power analysis

To characterize the sensitivity of the study design to age-related changes in MF_{Min}, we performed a simulation-based power analysis following the approach described in Chapter 2. Briefly, we used the observed sample sizes, baseline MF_{Min} at the youngest age (3 weeks), the average number of informative duplex bases per sample, and overdispersion estimates based on the between-animal variance estimated from a quasibinomial mixed-effects model to evaluate the power to detect observed fold-changes in MF between the youngest and oldest ages. Power was defined as the proportion of simulations yielding a p-value below a Bonferroni-adjusted threshold ($\alpha\_adj = 0.05/(\text{number of age groups}-1)$).

Positional patterns and mutation clustering

We examined positional patterns of mutation and potential clustering along the target loci under the MF_{Min} assumption. For these analyses, we first recentered coordinates within each target so that the first interrogated base corresponded to position 1, yielding relative positions from 1 to 2400.

For each locus and relative position, we aggregated mutation counts across all samples and ages. For each position, we defined a binary indicator per sample: a sample was counted as mutated (1) at that position if it had ≥1 non-reference allele calls at that site, and as not mutated (0) otherwise. If multiple distinct alternate alleles were called at the same position in the same sample, this was coded as 1. Thus, the mutation-based count at a position corresponds to the number of samples in which that site was mutated, irrespective of the specific alternate allele.

These counts were visualised as one-dimensional “spike” or stacked-segment plots to identify positional concentrations of mutations (i.e., hotspots) and mutation-sparse regions (cold spots) along each target. These positional analyses were descriptive rather than inferential and were used to identify visually apparent hotspots. Loci that showed evidence of potential hotspots across tissues on these plots were flagged for more detailed inspection. For these, we constructed composite plots combining four tracks: (i) age-stratified mutation spikes, (ii) median total depth across samples at each position, (iii) local GC content computed in a 101 bp sliding window with the mean GC content of the entire locus, and (iv) short-repeat context. Homopolymer tracts (≥4 identical bases), di-nucleotide repeats (≥3 units), and tri-nucleotide repeats (≥3 units) were identified directly from the reference sequence and displayed as separate tracks.

Tissue comparisons

Tissue comparisons were performed to assess differences in mutation burden, mutation spectra, and genic context across liver, BM, lung, and germ cells. Analyses were conducted at three levels: (i) across all four tissues individually, (ii) across the three somatic tissues only (liver, BM, and lung), and (iii) between germ cells and the somatic tissues. This framework was used to determine whether tissue effects reflected general differences among tissues, were specific to germ cells, or persisted within somatic tissues after excluding germ cells.

For overall mutation burden, MF_{Min} was compared among tissues within each age group. We also tested whether the age-related rate of mutation accumulation differed among tissues. In addition, the age-related rate of mutation accumulation in germ cells was compared with the mean rate across the three somatic tissues.

Mutation spectra were compared between groups within each age using MutSeqR::spectra_comparison(), restricting the analysis to MF_{Min}. To assess broader differences in age-related spectral change across tissues, mutation types were additionally grouped into five major variant classes: SNVs, indels (insertions and deletions), MNVs, complex variants, and structural variants (SVs). These classes were analyzed across tissue, age, and mutation class, with sample included as a random intercept.

For genic context analyses, MF_{Min} was modeled across tissue, age, and genic context, with sample included as a random intercept. Pairwise comparisons were first performed within combinations of age and genic context, with particular focus on contrasts between germ cells and each somatic tissue. A grouped analysis comparing germ cells with the somatic tissues combined was then used to test whether germline-somatic differences accounted for the observed context-specific effects.

3.4 Results

3.4.1 Overall mutation frequencies diverge between germ cells and somatic tissues with age.

We quantified mean MF in liver, BM, lung, and germ cells from MutaMouse males using DS; Figure 3.1 shows MF_{Min} and MF_{Max} across tissue and age. At 3, 14, and 23 weeks of age, MF_{Min} in all tissues remained within a narrow range (~3.6–4.6 × 10^{-8} mutations/base pair). Pairwise comparisons among tissues at each age and across these early age groups within each tissue were non-significant. MF_{Max} showed a similarly narrow range of ~4.2–5.9 × 10^{-8} mutations per base pair at 3 and 14 weeks.

Despite this early-life stability, age-related changes in MF_{Min} differed among tissues over the full age range. When age was modeled as a continuous variable, the increase in MF_{Min} with age differed among tissues ($p \approx 1.7 \times 10^{-15}$). Pairwise comparisons showed that the germ-cell slope differed significantly from those of liver, lung, and BM, as well as from the average somatic slope (all $p < 0.0001$). By contrast, liver, BM, and lung all showed positive age-related slopes, with no significant differences among their overall linear trends. Thus, the overall difference among tissue-specific age trends was driven primarily by germ cells, whereas the somatic tissues showed broadly similar age related increases in MF_{Min}.

Across the three somatic tissues, both MF_{Min} and MF_{Max} increased with age. At 3 weeks, MF_{Min} ranged 3.88–4.37 × 10^{-8} mutations per base pair in liver, lung, and BM. By ≥78 weeks, MF_{Min} increased to 6.53–7.87 × 10^{-8} mutations per base pair, corresponding to a 1.5–2.0-fold increase relative to 3 weeks. MF_{Max} showed an analogous 1.7–1.8-fold increase from 4.46–5.18 × 10^{-8} to 7.84–8.81 × 10^{-8} mutations per base pair between the youngest and oldest age group.

At the oldest age group, the highest MF_{Min} was observed in liver, whereas the highest MF_{Max} was observed in BM. In lung, MF_{Max} also showed an early increase between 3 and 23 weeks; however, this was not maintained at later ages and was driven by only two animals. Despite this localized deviation in MF_{Max}, the overall age-related increases in somatic MF were broadly similar.

When MF_{Min} was examined separately from MF_{Max}, the fitted shape of the age effect differed somewhat among the somatic tissues. In liver and lung, MF_{Min} was adequately described by a linear age effect on the logit scale, consistent with a constant proportional increase in MF with age. Conversely, in BM, MF_{Min} changed little at earlier ages and increased more noticeably only at ≥78 weeks; this pattern was better captured by a step-wise categorical age model than by a single linear age term. Overall, somatic tissues shared an age-related increase in MF, although the fitted shape of that increase differs across tissues.

In contrast, MF in germ cells did not show significant age-related changes. In fact, MF_{Min} was $3.83 \times 10^{-8} \pm 3.31 \times 10^{-9}$ mutations per base pair at 3 weeks and $2.84 \times 10^{-8} \pm 3.38 \times 10^{-9}$ mutations per base pair at ≥78 weeks (0.74-fold), although this difference was not statistically significant. MF_{Max} showed a similar pattern, changing from $4.79 \times 10^{-8} \pm 5.03 \times 10^{-9}$ to $3.86 \times 10^{-8} \pm 6.09 \times 10^{-9}$ mutations per base pair between the youngest and oldest age groups (0.81-fold). Thus, these results distinguish germ cells from the progressive age-related accumulation of nuclear MF in somatic tissues over the examined ages.

A post hoc power analysis was performed to evaluate the interpretability of the age-related MF_{Min} comparisons. The somatic tissue analyses had high estimated power to detect the observed age-related increases in MF_{Min}, exceeding 90% in liver, BM, and lung. In contrast, estimated power was substantially lower for germ cells (29.16%), below the conventional 80% threshold, reflecting both higher germ-cell variability and the smaller observed fold-change between 3 and ≥78 weeks.

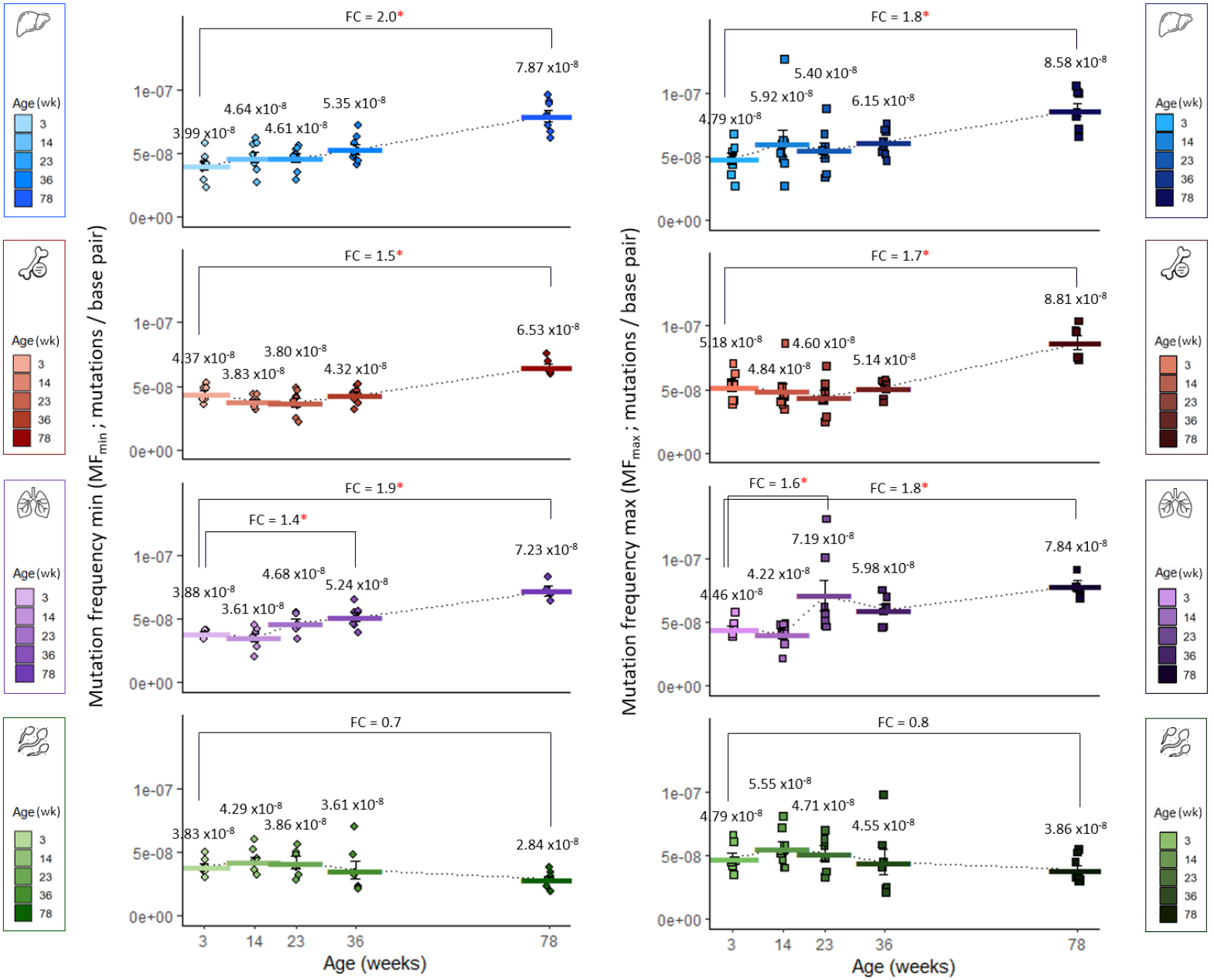

Figure 3.1: Duplex Sequencing mean mutation frequencies (MF_{Min} on the left and MF_{Max} on the right) ± SEM in the liver, bone marrow, lung, and germ cells (top to bottom) of MutaMouse males at 3, 14, 23, 36, ≥78 weeks. Each point represents the MF (mutations per bp) for each animal. Data labels indicate the mean MF (mutations per bp) per age group (N = 4–8 animals; see Table 3.1). Asterisks indicate significant differences ($p < 0.05$) in the fold change (FC), relative to the control group in the average MF across animals, based on model-based pairwise contrasts with Sidak-adjusted p-values from a quasibinomial generalized linear model.

3.4.2 Genic and intergenic mutation frequencies differ in a tissue-specific manner.

To determine whether age-related MF patterns differed between genic and intergenic regions, MF_{Min} was examined separately by genic context across tissues and ages (Figure 3.2). Aggregating MF_{Min} by genic context did not alter the observed overall age-related patterns: MF_{Min} in somatic tissues generally increased with age, whereas germ cells did not show the same age-related increase. Instead, analysis by genic context revealed differences in the relative magnitude of genic and intergenic MF_{Min} between somatic and germline compartments, with intergenic MF_{Min} generally higher in somatic tissues and genic MF_{Min} higher in germ cells ($p < 2.2 \times 10^{-16}$). The effect of age on genic and intergenic MF_{Min} also differed between somatic tissues and germ cells ($p = 0.026$).

All somatic tissues showed increases in MF_{Min} across age groups, but only BM and lung showed significant overall differences between genic and intergenic regions. In liver, MF_{Min} did not differ significantly between contexts, and the age-related increase was similar in genic and intergenic regions. By contrast, in BM and lung, intergenic MF_{Min} was higher overall than genic MF_{Min} when averaged across age groups. In BM, genic and intergenic MF_{Min} increased by similar amounts between the youngest and oldest age groups, from $3.76 \times 10^{-8} \pm 3.24 \times 10^{-9}$ to $5.46 \times 10^{-8} \pm 4.43 \times 10^{-9}$ and from $4.91 \times 10^{-8} \pm 3.46 \times 10^{-9}$ to $7.47 \times 10^{-8} \pm 4.85 \times 10^{-9}$ respectively (both 1.5-fold). In lung, both contexts increased with age, but the estimated increase was larger in intergenic regions, rising from $4.03 \times 10^{-8} \pm 3.48 \times 10^{-9}$ to $9.02 \times 10^{-8} \pm 6.89 \times 10^{-9}$ (2.2-fold) versus $3.70 \times 10^{-8} \pm 3.68 \times 10^{-9}$ to $5.04 \times 10^{-8} \pm 5.71 \times 10^{-9}$ (1.4-fold) in genic regions, but the difference in age-related increase between genic and intergenic regions was not statistically significant. Overall, stratifying MF_{Min} by genic context did not alter the shared age-related increase observed across somatic tissues, but it revealed tissue-specific differences in genic versus intergenic MF_{Min}, with significantly higher intergenic MF_{Min} in BM and lung but not liver. However, there was no evidence that genic and intergenic regions accumulated mutations at significantly different rates with age in any somatic tissue.

In germ cells, MF_{Min} also differed significantly between genic and intergenic regions, with genic MF_{Min} being higher overall than intergenic MF_{Min} when averaged across age groups. Neither context changed significantly with age: genic MF_{Min} was $4.35 \times 10^{-8} \pm 4.06 \times 10^{-9}$ at 3

weeks and $3.45 \times 10^{-8} \pm 3.81 \times 10^{-9}$ at ≥78 weeks (0.79-fold), while intergenic MF_{Min} was $3.37 \times 10^{-8} \pm 3.21 \times 10^{-9}$ and $2.31 \times 10^{-8} \pm 2.72 \times 10^{-9}$ at the same ages (0.68-fold). Despite this overall difference, age-related patterns were similar in both genic and intergenic contexts.

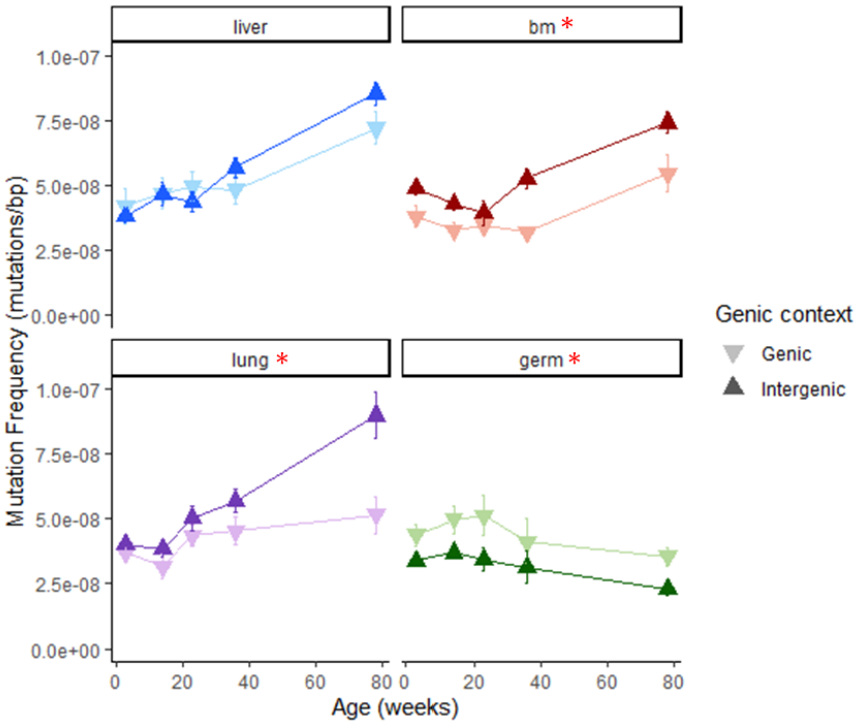

Figure 3.2: Duplex Sequencing mean mutation frequencies (MF_{Min}; mutations per bp) ± SEM across age and genic context in the liver, bone marrow, germ cells, and lung of MutaMouse males at 3, 14, 23, 36, and ≥78 weeks. Upward triangles denote intergenic mutation frequencies (the average of 11 targets) and downward triangles denote genic mutation frequencies (the average of 9 targets). Asterisks in facet labels indicate a significant main effect of genic context in tissue-specific Type II analysis-of-deviance tests ($p < 0.05$).

3.4.3 Mutations are unevenly distributed across targets and genomic positions

To assess how mutations were distributed across the loci, we mapped the number of samples with a mutation at each nucleotide position across all 20 targets in each tissue (Figure 3.3). Most positions showed no recurrence or only isolated events, whereas a small number of positions showed clear recurrent spikes. Several of these spikes were visible across tissues.

The most prominent recurrent signal occurred near the end of the intergenic target on chromosome 11 and was visible across all four tissues. Closer inspection of this region using a composite positional plot (Figure 3.4) showed that the apparent spike comprised adjacent 1-bp deletions at positions 2351 and 2356, observed in 2–7 samples per tissue (5–22%) and 7–18 samples per tissue (18–56%), respectively. These two deletion events were co-observed in 13 samples, with 13 of 18 samples (72%) harboring the deletion at position 2351 also carrying the deletion at position 2356, suggesting that they may reflect related local events rather than independent occurrences. A greater than tenfold decrease in the number of informative duplex bases was also observed at position 2363 in all tissues. This interval coincides with several adjacent homopolymer tracts. The cross-tissue recurrence of the signal, the clustering of adjacent 1-bp deletions and their frequent co-occurrence in the same samples, the location near the end of the target, the presence of nearby homopolymer tracts, and the abrupt drop in informative duplex depth at a nearby position are more consistent with a local technical artifact than with a biologically meaningful mutational hotspot. On this basis, positions 2350–2400 in the terminal region of the chromosome 11 target were excluded from downstream analyses to avoid inflating MF estimates.

Additional recurrent spikes were also observed in the genic targets on chromosomes 13, 1.2, 15, and 12 (Table S3.1). Several occurred in or near homopolymer runs, although no single shared sequence feature was apparent across loci. These observations were not investigated further in the present study.

Overall, positional mapping showed that recurrent mutation signals (or technical noise) were concentrated at a small number of sites rather than distributed uniformly across the panel.

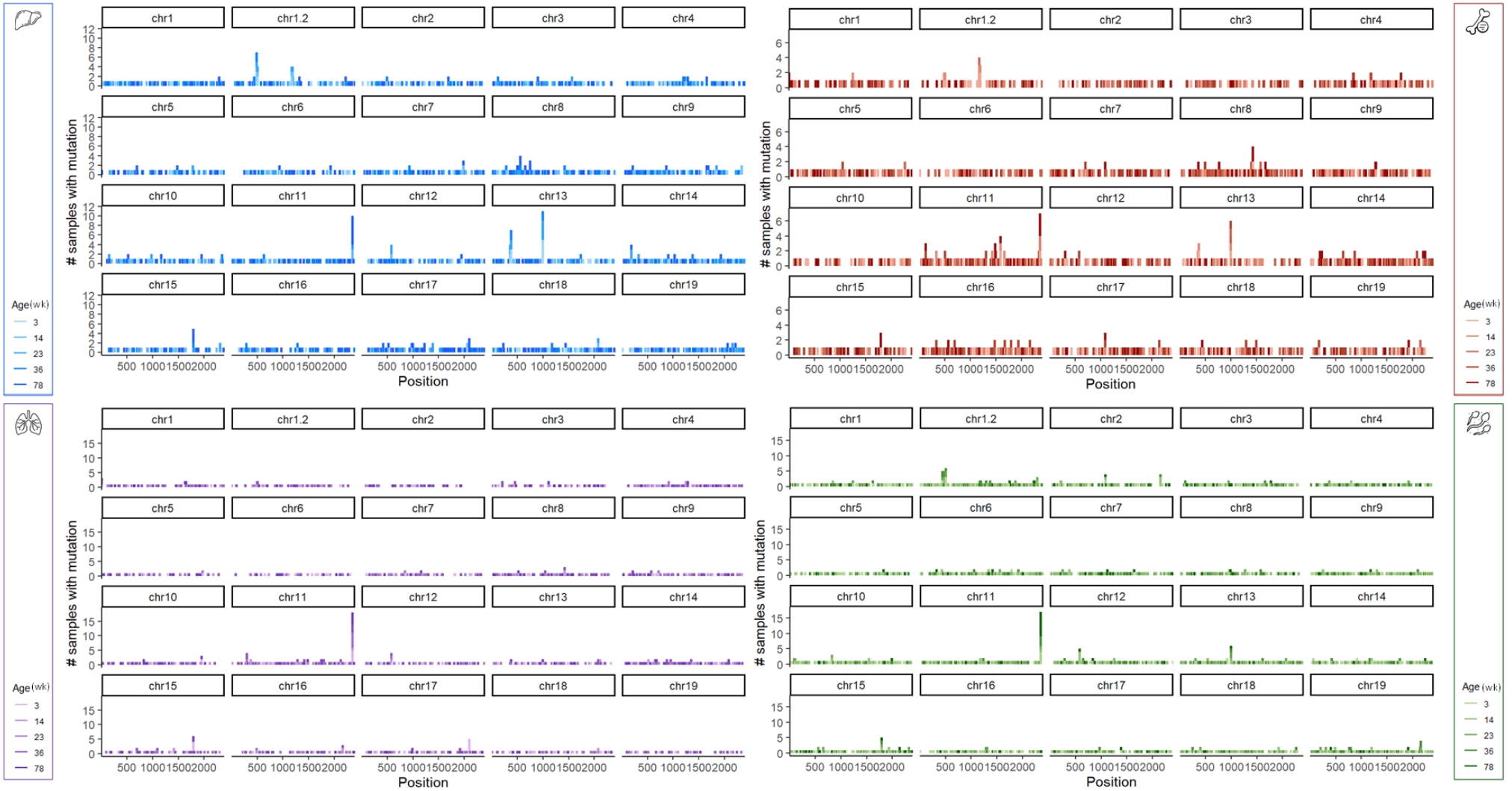

Figure 3.3: Global position-wise mutation maps across 20 targets in liver, bone marrow, lung, and germ cells. Within each tissue, mutation occurrence was aggregated across all samples and ages and displayed by relative nucleotide position for each target (facets; 2400 nt per target). For a given position, each sample contributed a value of 1 if at least one non-reference allele was detected at that site, irrespective of the specific alternate allele, and 0 otherwise. Accordingly, bar height represents the number of samples mutated at each position rather than the total number of mutation calls. Stacked bars are coloured by age group within each tissue, with lighter shades indicating younger age groups and darker shades indicating older age groups.

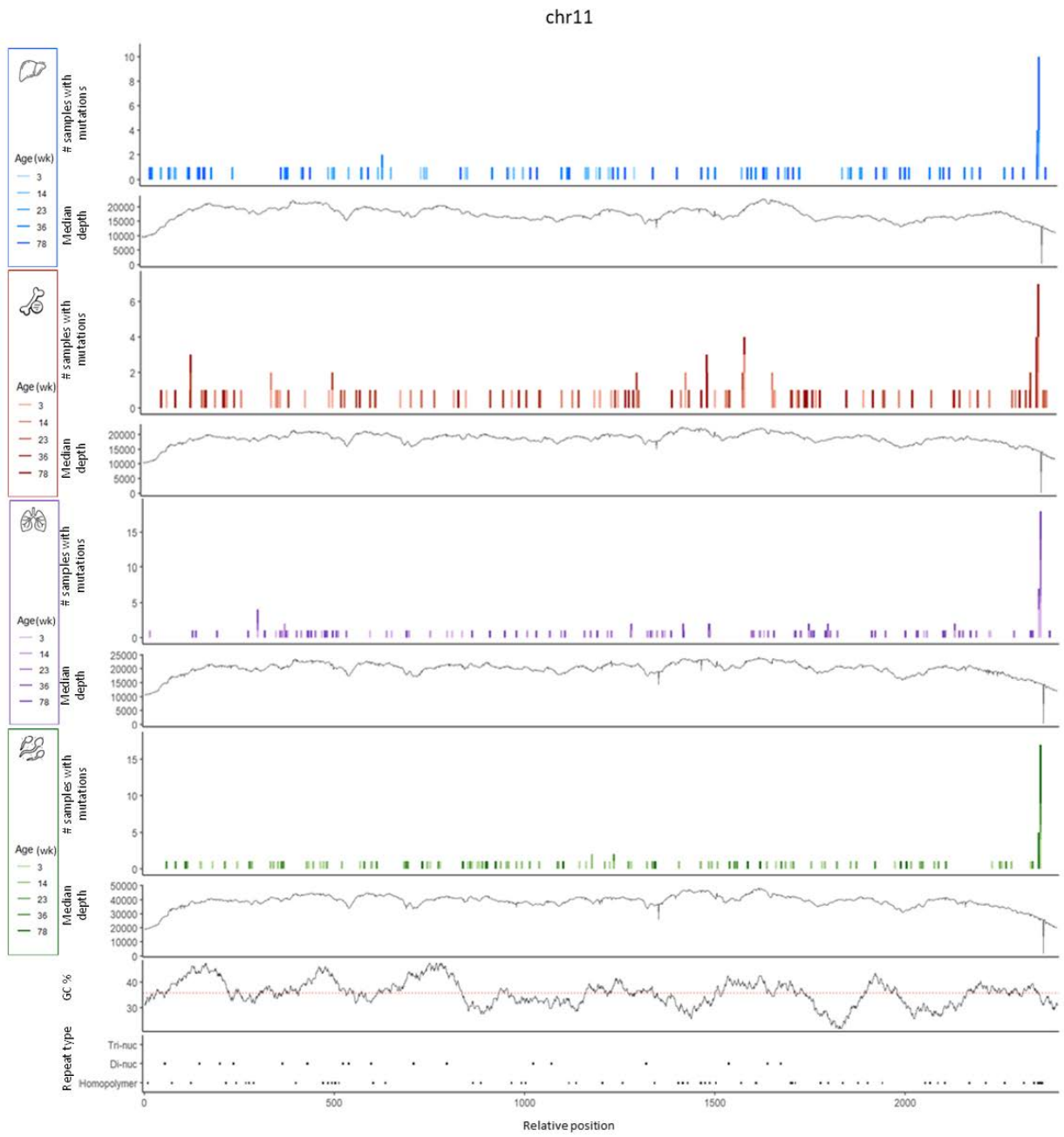

Figure 3.4: Composite positional map of the chromosome 11 target selected for detailed hotspot inspection. The top panel shows age-stratified mutation occurrence across relative nucleotide position within the target. At each site, bar height indicates the number of samples with at least one non-reference allele call, with each sample counted once per position regardless of the number of alternate alleles observed. Stacked colours denote age group. Lower tracks show median total sequencing depth, local GC content (101 bp sliding window; full-target mean indicated with red line), and short-repeat context, including homopolymer, dinucleotide, and trinucleotide repeat tracts.

3.4.4 Mutation spectra show tissue-specific age-related patterns.

To assess whether mutation spectra differed across tissues and age, we first compared overall mutation patterns among tissue-by-age groups and then modeled spectra after grouping mutations into five major variant classes—SNVs, indels, MNVs, SVs, and complex mutations—to reduce the multiple-comparison burden. Across all tissues, mutation spectra differed by tissue and changed with age. The specific variant classes contributing to age-related spectral changes also differed by tissue. Pairwise comparisons further showed that germ-cell spectra differed from those of liver, BM, and lung at every age examined (all adjusted $p < 0.05$), whereas somatic tissues were generally more similar to one another, although BM and lung differed at 3 weeks and liver and lung differed at 23 weeks. When the analysis was restricted to the three somatic tissues, spectra still differed among tissues with age ($p \approx 0.008$), but the overall effect of age was similar when averaged across major variant classes. Nevertheless, the mutation classes contributing to these age-related changes differed among somatic tissues (summarized below). Thus, the clearest broad-scale spectral distinction remained between germline and somatic tissues, while somatic tissues also showed tissue-specific age-related spectral patterns.

To examine these patterns at higher resolution, we next summarized the simple mutation spectra by separating SNVs into the six pyrimidine-referenced base-substitution classes (SBS-6) and indels into insertions and deletions (Figure 3.5). Because this stratification divided rare mutations across tissue, age, and subtype categories, subtype-level patterns should be interpreted cautiously. Across the three somatic tissues, the simple spectra were broadly similar, and MF_{Min} of individual classes increased with age, consistent with the parallel age-related patterns observed at the level of major variant classes. Lung showed the most uniform pattern, with no clear evidence that any one subtype changed disproportionately with age (age-by-subtype: $p = 0.2233$). By contrast, liver showed the clearest subtype-level changes, with significant increases in C:G>T:A, T:A>A:T, T:A>C:G, and T:A>G:C between the youngest and oldest age groups (2.6-, 2.7-, 4.7-, and 3.4-fold, respectively; age-by-subtype: liver, $p \approx 1.1 \times 10^{-8}$). Because C:G>T:A mutations can include CpG-associated C:G>T:A events, we further stratified this subtype by CpG context. CpG-context C:G>T:A mutations represented a minority of

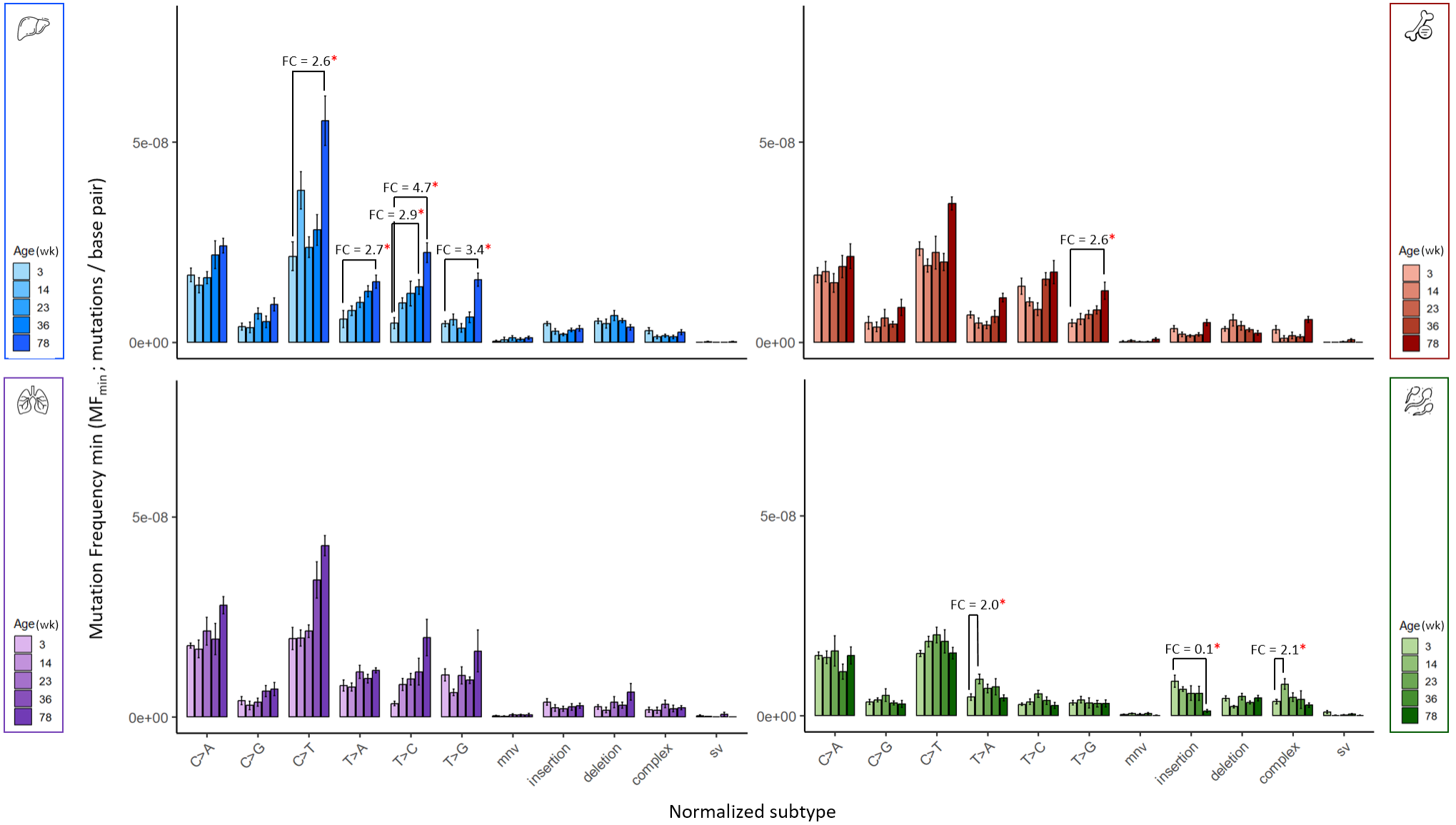

Figure 3.5: Duplex Sequencing mutation frequencies (MF_{Min}) by normalized subtype in the liver, bone marrow, lung, and germ cells of MutaMouse males at various ages. Mutation subtypes are presented as the average MF ± SEM (N = 4–8 animals; see Table 3.1). Mutation subtypes include the six single nucleotide variants using a pyrimidine reference, multi-nucleotide variants (MNV), insertions, deletions, complex mutations (where an indel and a substitution occurred in the same mutational event), and structural variants (SV). Asterisks indicate significant differences ($p < 0.05$) relative to the control group, based on model-based pairwise contrasts from a binomial generalized linear mixed-effects model including age, normalized subtype, and their interaction, with sample as a random intercept; p-values were adjusted using the Sidak method.

observed C:G>T:A mutations in each tissue (ranging from 18.7% in germ cells to 34.1% in lung) and the CpG fraction differed by tissue but did not show a significant overall age-related increase (tissue: $p = 4.3 \times 10^{-7}$; age: $p = 0.16$). BM showed a more limited shift in the simple spectra, with T:A>G:C showing the clearest age-related increase (2.6-fold), whereas most other subtypes changed less markedly (age-by-subtype: $p = 0.0022$). Overall, somatic tissues shared a broad age-related increase in MF, but the distribution of that increase across mutation classes was more uneven in liver and BM than in lung.

Consistent with the lack of age-related change in overall germ-cell MF, the germ-cell simple spectra did not show an age-related shift in composition in contrast to the somatic tissues. Instead, changes were subtype-specific and often transient. T:A>A:T increased 2.0-fold between 3 and 14 weeks, and complex mutations increased 2.1-fold over the same interval. Several other subtypes, including C:G>G:C, C:G>T:A, T:A>C:G, and possibly T:A>G:C, also showed trends of early-life increases that were not maintained at later ages, producing bell-shaped age patterns in the spectra. Insertions, by contrast, decreased significantly with age, falling to 0.14-fold of the 3-week value by ≥78 weeks. Thus, unlike the broadly increasing somatic spectra, germ-cell spectra were characterized by more transient and subtype-specific age-related changes, without a consistent overall shift across age.

Trinucleotide signature fitting was also explored by decomposing tissue-by-age spectra into COSMIC-based reference signatures. However, most spectra showed low cosine similarity between observed and reconstructed profiles, likely reflecting sparse mutation counts in these normal tissues. In these exploratory fits, much of the reconstructed signal was often assigned to SBS5- and SBS40a-like signature contributions, although the low reconstruction quality limited confident assignment of specific signatures (Figure S3.5 in Supplementary Materials).

3.5 Discussion

We used DS to examine how aging and tissue biology jointly shape nuclear mutation accumulation in MutaMouse males. Overall, our results show that age-related mutation accumulation was strongly tissue-dependent: somatic tissues exhibited broadly similar

increases in MF with age, whereas male germ cells showed no age-related change. Additional differences in mutation spectra and genomic distribution across tissues suggest that tissue context may also shape the types and genomic locations of mutations recovered, although spectral patterns should be interpreted cautiously because mutation counts became sparse after stratification by class.

Tissue-specific divergence observed with age emerged from a shared low baseline mutation burden early in life, as MFs were similar across liver, BM, lung, and germ cells at the youngest ages examined. Prior transgenic reporter studies similarly found broadly comparable mutant frequencies across tissues early in life, with tissue-specific differences becoming more apparent at later ages (Dollé et al., 2002; Ono et al., 2000). Although the combined age series included cohorts with different vehicle-exposure histories, olive oil is routinely used as a vehicle control in MutaMouse mutagenicity studies (Gingerich et al., 2014) and was not expected to induce mutations under the exposure conditions used here. Consistent with this expectation, no significant differences in MF were observed between the 3-week group, which was not directly gavaged, and the olive-oil-gavaged intermediate-age groups, suggesting that vehicle exposure was unlikely to materially influence the observed age-related patterns. This pattern suggests that mutation burden is shaped by broadly shared processes early in life, with tissue-specific biology exerting greater influence as animals age.

Among the somatic tissues, the age-related increase in MF was broadly similar but not identical in form. In liver and lung, MF_{Min} increased more than MF_{Max}, consistent with the gradual accumulation of many unique mutations rather than a strong contribution from large clonal expansions at the neutral loci examined. Under baseline conditions, liver and lung are largely quiescent, with most cell division occurring in bursts during localized maintenance or repair (Eenjes et al., 2022; Stanger, 2015). Such events may generate small lineage expansions, but may be less likely to produce large, persistent clones that are repeatedly sampled across many independent DNA molecules. The liver and lung patterns are, therefore, more consistent with the progressive accumulation of many unique mutational events than with sustained dominance by a small number of expanded mutant clones.

BM differed from liver and lung in showing a more delayed increase in MF, rather than a gradual accumulation across adulthood; there was less change in MF earlier in life and a more pronounced rise at ≥78 weeks. This late-life step suggests an old-age shift in BM biology. Unlike liver and lung, where mutations in resident stem or progenitor cells can be retained locally through long-lived tissue lineages, many differentiated BM-derived cells are transient; therefore, age-related MF in BM is likely to more strongly reflect mutation accumulation and clonal dynamics within the hematopoietic stem and progenitor compartment. A plausible explanation is inflammaging in old marrow (Zhao et al., 2021), which may alter niche signaling and increase overall cellular stress. Consistent with this, aged hematopoietic stem cells up-regulate inflammation- and stress-response genes and down-regulate chromatin and DNA-repair programs (Chambers et al., 2007; Zhao et al., 2021), changes that could support both higher mutation burden and clonal skew. BM also showed a larger increase in MF_{Max} than MF_{Min}, consistent with a stronger late-life contribution from expanded mutant clones. Age-related mutation accumulation in BM may therefore reflect both increased mutation burden within hematopoietic stem and progenitor cells and a strong contribution from late-life clonal expansion.

Unlike the somatic tissues, the male germline remained relatively stable across age and did not show an age-related increase in nuclear mutation burden. This differs from human studies, in which sperm mutation burden and paternally derived *de novo* mutations in offspring increase with paternal age (Neville et al., 2025; Shojaeisaadi et al., 2024). The apparent stability of germ-cell MF may therefore reflect, at least in part, the relatively young germ-cell age of even our oldest mice. As a heuristic cross-species proxy, comparative germline-mutation models quantify post-pubertal male germline aging in terms of ongoing spermatogonial stem-cell divisions, which constitute the dominant source of male germline cell divisions after puberty (Lindsay et al., 2019). Under the assumptions used in that framework (puberty at 1 month in mouse and 15 years in human, with 42 SSC divisions/year in mouse and 23/year in human) an 85-week-old mouse corresponds to only about 2.8 human post-pubertal germline years, or roughly a 17.8-year-old human in germ-cell terms. A spermatogenesis-duration-based comparison gives a similar estimate: using the approximate duration from spermatogonial stem

cell to mature sperm in mouse (~49 days; Hagio et al., 2021) and from spermatogenesis through epididymal transit/maturation in humans (~81–88 days; ~74 days of spermatogenesis plus 1–2 weeks of epididymal transit/maturation; Dacheux & Dacheux, 2014; Neto et al., 2016), the oldest mice again correspond to only about 2.6–2.8 human post-pubertal germline years. Therefore, even our oldest mice may still be relatively young from a germline-aging perspective and may precede the ages at which a detectable increase in mutation burden emerges (although this comparison should be interpreted as a rough cross-species estimate that does not mechanistically explain the absence of an age effect in our data).

The relative stability of germline mutation burden is also broadly consistent with previous rodent germline studies. In MutaMouse itself, *lacZ*-based analyses of spontaneous mutant frequency in germ cells from 92 control males, aged 13-28 weeks, reported no significant age effect, suggesting a relatively stable early-adult germline burden (Zhou et al., 2025). Similarly, Martin et al. (2001) reported no evidence for age-related accumulation of *lacZ* mutations in testicular DNA in a *lacZ* transgenic mouse. In a different transgenic mouse model, Walter et al. (1998) used a *lacI* reporter and observed a decline in mutant frequency across spermatogenesis in young adults (~60 days; ~9 weeks), implying strong intra-testicular quality control. At the same time, spermatogenic cells from old mice (28 months; ~120 weeks) showed significantly higher mutant frequencies than in young (60-day; ~9-week) or middle-aged (15-month; ~65-week) mice. Our MutaMouse animals span from 3 weeks to late adulthood (78-85 weeks, ~18-20 months) and do not show any evidence of a late-life increase in nuclear MF within this interval. That said, our oldest animals are only modestly older than the “middle-aged” 15-month (~65-week) group studied by Walter et al. (1998) and are substantially younger than their 28-month (~120-week) cohort. We therefore cannot exclude a later mutational breakdown that emerges beyond the ages examined here. Thus, existing work and our results support a model in which the mouse male germline remains low-burden through much of life and may only show a clear mutational breakdown, if at all, at very advanced ages. This suggests that, although mice remain useful for studying germline mutagenesis mechanisms, standard aging studies in mice may not fully recapitulate the gradual paternal-age-associated increase in germline mutation burden observed in humans.

However, interpretation of the germ-cell findings should be cautious, as power to detect a modest age-related effect was limited. Based on the design calculations in Esina et al. (2024; Chapter 2), the effect size and sample standard deviation observed here would likely require a substantially larger germ-cell sample size to achieve adequate power (between 9 and ≥20 animals per age group, with the true requirement likely falling closer to the upper than the lower end of this range). Thus, while we did not detect an age-related increase in germ-cell nuclear MF, and did not observe changes comparable to those seen in somatic tissues, we cannot exclude smaller changes that this study was underpowered to detect. This limitation may also be relevant when interpreting prior rodent reporter-gene studies, where age-related germline effects may have been difficult to detect because of smaller reporter targets and assay-specific sensitivity limits compared with DS.

All these burden-level patterns suggest that age-related nuclear mutagenesis in MutaMouse males does not follow the same temporal trajectory across tissues. Earlier reporter studies often described mutation accumulation as broadly progressive across life, sometimes at relatively linear, albeit tissue-specific, rates (Fischer & Stringer, 2008; Ono et al., 2000). However, prior work also suggested that this pattern was not universal. In mice, age-related mutation accumulation was also found to plateau in brain tissue (Dollé et al., 1997; Stuart et al., 2000). Our results support this more heterogeneous view. Liver and lung were most consistent with approximately exponential increases across age, BM showed a delayed late-life rise, and germ cells showed no increase over the ages examined. Therefore, these findings argue against one common age-related pattern across tissues and instead suggest that mutation accumulation unfolds with tissue-specific temporal patterns.

Analysis of mutation spectra added a further dimension to these tissue differences by suggesting that age-related mutation accumulation may not be distributed identically across mutation classes, though interpretation should be cautious due to the low mutation counts after stratification. Although liver, lung, and BM all showed broadly increasing somatic mutation burden with age, the spectral changes accompanying those increases were not identical: liver and BM showed stronger statistical evidence for subtype-specific age-related changes, whereas lung showed no clear evidence that individual subtypes changed

disproportionately with age. Germ cells remained distinct, showing no comparable progressive spectral shift and instead displaying transient early-life changes in several subtypes together with a decline in insertions across age. Thus, even when somatic tissues showed broadly similar age-related increases in mutation burden, the underlying spectral changes were not uniform, suggesting that tissues may reach similar burden-level outcomes through partly distinct contributions from different mutational processes.

Stratification by genic context added a further layer of tissue specificity by showing that the genomic distribution of mutation burden also differed across tissues. In BM and lung, MFs were lower in genic than intergenic regions, whereas liver showed no clear separation between contexts. Germ cells were distinct again, showing the opposite relationship, with higher genic than intergenic mutation frequencies overall, despite showing no comparable age-related increase in either context. Thus, the association between genic context and mutation burden was itself tissue-dependent: somatic tissues were broadly more similar to one another than to germ cells, but were not identical, with liver showing a less differentiated genic-intergenic pattern than BM and lung. These findings indicate that the relative enrichment of mutations in genic regions cannot be explained by a single shared model across tissues.

Together, these burden-, spectrum-, and genic-context patterns are challenging to reconcile with a model in which age-related nuclear mutagenesis is driven primarily by proliferation or any other single global process. Although proliferative status has often been invoked to explain tissue differences in mutation accumulation (Busuttil et al., 2007; Dollé et al., 2000; Stuart et al., 2000), our results do not follow the predictions of a replication-dominated model, consistent with other studies showing that age-related mutation accumulation can reflect multiple processes beyond replication alone (Abascal et al., 2021; Ono et al., 2000; Shojaeisaadi et al., 2024). The strongest age-related increases in mutation burden were not observed in the most proliferative tissues; the relatively quiescent liver and lung showed clearer age-related accumulation than BM and especially germ cells. Likewise, C:G>T:A mutations did not provide strong evidence for a shared CpG/SBS1-like aging process: liver showed the clearest increase in this subtype, BM and lung showed only modest trends, and germ cells did not show a progressive increase. Because C:G>T:A transitions can arise through

multiple mechanisms, the presence of C:G>T:A mutations alone does not imply replication-associated mutagenesis. However, age-associated C:G>T:A mutations in humans are often enriched at CpG sites, where deamination of methylated cytosines contributes to SBS1-like aging signatures. Thus, a strong CpG-driven C:G>T:A signal would have supported a canonical cell-division-associated aging process. In our data, however, C:G>T:A mutations were mostly non-CpG and did not show a consistent age-related increase in CpG fraction, suggesting that this subtype is unlikely to reflect a uniform CpG-specific aging process alone. Furthermore, trinucleotide signature profiles were dominated by clock-like SBS5/SBS40a-type contributions, which track age but are not tightly linked to replication alone, rather than SBS1, the better-established cell-division-associated clock-like signature. The genic-context results point in the same direction. BM and lung were consistent with lower mutation burden in genic than intergenic regions, a pattern that has been interpreted in prior DS studies as potentially reflecting transcription-associated protection, including transcription-coupled repair (Dodge et al., 2023; Schuster et al., 2024). However, neither liver nor germ cells followed this pattern. Although transcription-associated repair can shape mutation accumulation patterns, including genic/intergenic asymmetry and transcriptional strand bias (Luzadder et al., 2025; Minko et al., 2024), this protection does not seem to operate uniformly across tissues. Therefore, these findings support a model in which age-related nuclear mutagenesis in MutaMouse males reflects tissue-specific differences in damage, repair, turnover, clonal dynamics, and regional protection, with proliferation and transcription-associated repair contributing in some contexts but not serving as sufficient general explanations.

In addition to the insights into the biology of aging in mice, this study also provided empirical validation of the targeted Mouse Mutagenesis Panel. Positional mapping showed that most mutations were isolated across the panel, but a small number of positions exhibited recurrent signals across tissues, despite the tissue-specific nature of age-related mutation accumulation. Because the same genomic positions are repeatedly interrogated across samples in a fixed-panel assay, recurrent positional signals can reflect either true local vulnerability or target-specific technical artifacts, and therefore require empirical evaluation before biological interpretation. In the present study, the strongest recurrent signal, near the end of the

chromosome 11 intergenic target, was more consistent with a local technical artifact than with a true biological hotspot and was therefore excluded from downstream analyses. Additional recurrent sites were observed at other loci but were not resolved further. Accordingly, these sites are best treated as flagged recurrent positional signals rather than confirmed biological hotspots. These findings highlight both a limitation and an opportunity: systematic positional mapping can help distinguish technical recurrence from genuinely fragile genomic sites within targeted Duplex Sequencing panels and provide empirical single-base validation of the Mouse Mutagenesis Panel to reduce inflation of MF estimates and misattribution of recurrent signals in future studies.

3.6 Conclusion

Overall, this comparative multi-tissue analysis in MutaMouse males shows that age-related nuclear mutagenesis is not governed by a single global pattern. Instead, mutation accumulation unfolds in tissue-specific ways across liver, BM, lung, and male germ cells, with differences evident in burden, mutational spectrum, signature composition, and genomic distribution. These findings provide a framework for understanding aging-associated mutagenesis as a tissue-dependent process shaped by local biology rather than a uniform consequence of age alone. At the same time, the relative stability of MF from 3 to 23 weeks suggests that a broad range of young-adult starting ages may provide comparable baseline MFs in MutaMouse male studies across these tissues, informing future study designs.

3.7 References

Abascal, F., Harvey, L. M. R., Mitchell, E., Lawson, A. R. J., Lensing, S. V., Ellis, P., Russell, A. J. C., Alcantara, R. E., Baez-Ortega, A., Wang, Y., Kwa, E. J., Lee-Six, H., Cagan, A., Coorens, T. H. H., Chapman, M. S., Olafsson, S., Leonard, S., Jones, D., Machado, H. E., … Martincorena, I. (2021). Somatic mutation landscapes at single-molecule resolution. *Nature*, *593*(7859), 405–410. https://doi.org/10.1038/s41586-021-03477-4

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A.-L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., … Stratton, M. R. (2013). Signatures of mutational processes in human cancer. *Nature*, *500*(7463), 415–421. https://doi.org/10.1038/nature12477

Arbeithuber, B., Hester, J., Cremona, M. A., Stoler, N., Zaidi, A., Higgins, B., Anthony, K., Chiaromonte, F., Diaz, F. J., & Makova, K. D. (2020). Age-related accumulation of de novo mitochondrial mutations in mammalian oocytes and somatic tissues. *PLOS Biology*, *18*(7), e3000745. https://doi.org/10.1371/journal.pbio.3000745

Busuttil, R. A., Garcia, A. M., Reddick, R. L., Dollé, M. E. T., Calder, R. B., Nelson, J. F., & Vijg, J. (2007). Intra-Organ Variation in Age-Related Mutation Accumulation in the Mouse. *PLOS ONE*, *2*(9), e876. https://doi.org/10.1371/journal.pone.0000876

Chambers, S. M., Shaw, C. A., Gatza, C., Fisk, C. J., Donehower, L. A., & Goodell, M. A. (2007). Aging Hematopoietic Stem Cells Decline in Function and Exhibit Epigenetic Dysregulation. *PLOS Biology*, *5*(8), e201. https://doi.org/10.1371/journal.pbio.0050201

Dacheux, J.-L., & Dacheux, F. (2014). New insights into epididymal function in relation to sperm maturation. *Reproduction*, *147*(2), R27–R42. https://doi.org/10.1530/REP-13-0420

Dodge, A. E., LeBlanc, D. P. M., Zhou, G., Williams, A., Meier, M. J., Van, P., Lo, F. Y., Valentine Iii, C. C., Salk, J. J., Yauk, C. L., & Marchetti, F. (2023). Duplex sequencing provides detailed characterization of mutation frequencies and spectra in the bone marrow of MutaMouse males exposed to procarbazine hydrochloride. *Archives of Toxicology*, *97*(8), 2245–2259. https://doi.org/10.1007/s00204-023-03527-y

Dodge, A. E., Williams, A., LeBlanc, D. P. M., Schuster, D. M., Esina, E., Valentine, C. C., Salk, J. J., Maslov, A. Y., Bradley, C., Yauk, C. L., Marchetti, F., & Meier, M. J. (2025). MutSeqR: An Open-Source R Package for Standardized Analysis of Error-Corrected Next-Generation Sequencing Data in Genetic Toxicology. *Bioinformatics Advances*, vbaf265. https://doi.org/10.1093/bioadv/vbaf265

Dollé, M. E. T., Giese, H., Hopkins, C. L., Martus, H.-J., Hausdorff, J. M., & Vijg, J. (1997). Rapid accumulation of genome rearrangements in liver but not in brain of old mice. *Nature Genetics*, *17*(4), Article 4. https://doi.org/10.1038/ng1297-431

Dollé, M. E. T., Snyder, W. K., Dunson, D. B., & Vijg, J. (2002). Mutational fingerprints of aging. *Nucleic Acids Research*, *30*(2), 545–549. https://doi.org/10.1093/nar/30.2.545

Dollé, M. E. T., Snyder, W. K., Gossen, J. A., Lohman, P. H. M., & Vijg, J. (2000). Distinct spectra of somatic mutations accumulated with age in mouse heart and small intestine. *Proceedings of the National Academy of Sciences*, *97*(15), 8403–8408. https://doi.org/10.1073/pnas.97.15.8403

Eenjes, E., Tibboel, D., Wijnen, R. M. H., & Rottier, R. J. (2022). Lung epithelium development and airway regeneration. *Frontiers in Cell and Developmental Biology*, *10*. https://doi.org/10.3389/fcell.2022.1022457

Esina, E., Dodge, A. E., Williams, A., Schuster, D. M., LeBlanc, D. P. M., Marchetti, F., & Yauk, C. L. (2024). Power analyses to inform Duplex Sequencing study designs for MutaMouse liver and bone marrow. *Environmental and Molecular Mutagenesis*, *65*(8), 234–242. https://doi.org/10.1002/em.22619

Fischer, J. M., & Stringer, J. R. (2008). Mutation in aging mice occurs in diverse cell types that proliferate postmutation. *Aging Cell*, *7*(5), 667–680. https://doi.org/10.1111/j.1474-9726.2008.00416.x

Gingerich, J., Soper, L., Lemieux, C., Marchetti, F., & Douglas, G. (2014). *Transgenic Rodent Gene Mutation Assay in Somatic Tissues* (pp. 305–321). https://doi.org/10.1007/978-1-4939-1068-7_18

Hagio, S., Tsuji, N., Furukawa, S., Takeuchi, K., Hayashi, S., Kuroda, Y., Honma, M., & Masumura, K. (2021). Effect of sampling time on somatic and germ cell mutations induced by acrylamide in gpt delta mice. *Genes and Environment*, *43*(1), 4. https://doi.org/10.1186/s41021-021-00175-5

Kennedy, S. R., Salk, J. J., Schmitt, M. W., & Loeb, L. A. (2013). Ultra-Sensitive Sequencing Reveals an Age-Related Increase in Somatic Mitochondrial Mutations That Are Inconsistent with Oxidative Damage. *PLOS Genetics*, *9*(9), e1003794. https://doi.org/10.1371/journal.pgen.1003794

Kimoto, T., Kay, J. E., Li, N., & Engelward, B. P. (2017). Recombinant cells in the lung increase with age via de novo recombination events and clonal expansion. *Environmental and Molecular Mutagenesis*, *58*(3), 135–145. https://doi.org/10.1002/em.22082

LeBlanc, D. P. M., Meier, M., Lo, F. Y., Schmidt, E., Valentine, C., Williams, A., Salk, J. J., Yauk, C. L., & Marchetti, F. (2022). Duplex sequencing identifies genomic features that determine susceptibility to benzo(a)pyrene-induced in vivo mutations. *BMC Genomics*, *23*(1), 542. https://doi.org/10.1186/s12864-022-08752-w

Lindsay, S. J., Rahbari, R., Kaplanis, J., Keane, T., & Hurles, M. E. (2019). Similarities and differences in patterns of germline mutation between mice and humans. *Nature Communications*, *10*(1), 4053. https://doi.org/10.1038/s41467-019-12023-w

Luzadder, M. M., Minko, I. G., Vartanian, V. L., Davenport, M., Fedorov, L. M., McCullough, A. K., & Lloyd, R. S. (2025). The Distinct Roles of NEIL1 and XPA in Limiting Aflatoxin B1–Induced Mutagenesis in Mice. *Molecular Cancer Research*, *23*(1), 46–58. https://doi.org/10.1158/1541-7786.MCR-24-0577

Martin, S. L., Hopkins, C. L., Naumer, A., Dollé, M. E. T., & Vijg, J. (2001). Mutation frequency and type during ageing in mouse seminiferous tubules. *Mechanisms of Ageing and Development*, *122*(12), 1321–1331. https://doi.org/10.1016/S0047-6374(01)00267-6

Minko, I. G., Luzadder, M. M., Vartanian, V. L., Rice, S. P. M., Nguyen, M. M., Sanchez-Contreras, M., Van, P., Kennedy, S. R., McCullough, A. K., & Lloyd, R. S. (2024). Frequencies and spectra of aflatoxin B1-induced mutations in liver genomes of NEIL1-deficient mice as revealed by duplex sequencing. *NAR Molecular Medicine*, *1*(2), ugae006. https://doi.org/10.1093/narmme/ugae006

Navarro, C. L., Cau, P., & Lévy, N. (2006). Molecular bases of progeroid syndromes. *Human Molecular Genetics*, *15*(suppl_2), R151–R161. https://doi.org/10.1093/hmg/ddl214

Neto, F. T. L., Bach, P. V., Najari, B. B., Li, P. S., & Goldstein, M. (2016). Spermatogenesis in humans and its affecting factors. *Seminars in Cell & Developmental Biology, Molecular Mechanisms in Spermatogenesis*, *59*, 10–26. https://doi.org/10.1016/j.semcdb.2016.04.009

Neville, M. D. C., Lawson, A. R. J., Sanghvi, R., Abascal, F., Pham, M. H., Cagan, A., Nicola, P. A., Bayzetinova, T., Baez-Ortega, A., Roberts, K., Lensing, S. V., Widaa, S., Alcantara, R. E., García, M. P., Wadge, S., Stratton, M. R., Campbell, P. J., Small, K., Martincorena, I., …

Rahbari, R. (2025). Sperm sequencing reveals extensive positive selection in the male germline. *Nature*, *647*(8089), 421–428. https://doi.org/10.1038/s41586-025-09448-3

O'Brien, J. M., Beal, M. A., Gingerich, J. D., Soper, L., Douglas, G. R., Yauk, C. L., & Marchetti, F. (2014). Transgenic Rodent Assay for Quantifying Male Germ Cell Mutant Frequency. *Journal of Visualized Experiments : JoVE*, (90), 51576. https://doi.org/10.3791/51576

Ono, T., Ikehata, H., Nakamura, S., Saito, Y., Hosoi, Y., Takai, Y., Yamada, S., Onodera, J., & Yamamoto, K. (2000). Age-associated increase of spontaneous mutant frequency and molecular nature of mutation in newborn and old lacZ-transgenic mouse. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *447*(2), 165–177. https://doi.org/10.1016/S0027-5107(99)00200-6

Sanchez-Contreras, M., Sweetwyne, M. T., Tsantilas, K. A., Whitson, J. A., Campbell, M. D., Kohrn, B. F., Kim, H. J., Hipp, M. J., Fredrickson, J., Nguyen, M. M., Hurley, J. B., Marcinek, D. J., Rabinovitch, P. S., & Kennedy, S. R. (2023). The multi-tissue landscape of somatic mtDNA mutations indicates tissue-specific accumulation and removal in aging. *eLife*, *12*, e83395. https://doi.org/10.7554/eLife.83395

Schuster, D. M., LeBlanc, D. P. M., Zhou, G., Meier, M. J., Dodge, A. E., White, P. A., Long, A. S., Williams, A., Hobbs, C., Diesing, A., Smith-Roe, S. L., Salk, J. J., Marchetti, F., & Yauk, C. L. (2024). Dose-Related Mutagenic and Clastogenic Effects of Benzo[b]fluoranthene in Mouse Somatic Tissues Detected by Duplex Sequencing and the Micronucleus Assay. *Environmental Science & Technology*, *58*(49), 21450–21463. https://doi.org/10.1021/acs.est.4c07236

Shelby, M. D. (1994). Human germ cell mutagens. *Environmental and Molecular Mutagenesis*, *23*(S2), 30–34. https://doi.org/10.1002/em.2850230609

Shoag, J. E., Srinivasa, A., Loh, C. A., Liu, M. H., Lassen, E., Melanaphy, S., Costa, B. M., Grońska-Pęski, M., Jabara, N. T., Picciotto, S., Choi, U., Bohorquez, A. D., Barbieri, C. E., Callum, P., Skytte, A.-B., & Evrony, G. D. (2025). Direct measurement of the male germline mutation rate in individuals using sequential sperm samples. *Nature Communications*, *16*(1), 2546. https://doi.org/10.1038/s41467-025-57507-0

Shojaeisaadi, H., Schoenrock, A., Meier, M. J., Williams, A., Norris, J. M., Palmer, N. D., Yauk, C. L., & Marchetti, F. (2024). Mutational signature analyses in multi-child families reveal sources of age-related increases in human germline mutations. *Communications Biology*, *7*(1), 1451. https://doi.org/10.1038/s42003-024-07140-2

Stanger, B. Z. (2015). Cellular Homeostasis and Repair in the Mammalian Liver. *Annual Review of Physiology*, *77*, 179–200. https://doi.org/10.1146/annurev-physiol-021113-170255

Stewart, M. T., Zhou, G., LeBlanc, D. P. M., Dodge, A. E., Meier, M. J., Williams, A., Long, A. S., White, P. A., Yauk, C. L., & Marchetti, F. (2025). Long term exposure to benzo[*b*]fluoranthene does not induce mutations in Mutamouse male germ cells. *Reproductive Toxicology*, *137*, 108985. https://doi.org/10.1016/j.reprotox.2025.108985

Stuart, G. R., Oda, Y., de Boer, J. G., & Glickman, B. W. (2000). Mutation frequency and specificity with age in liver, bladder and brain of lacI transgenic mice. *Genetics*, *154*(3), 1291–1300. https://doi.org/10.1093/genetics/154.3.1291

Turker, M. S., Lasarev, M., Connolly, L., Kasameyer, E., & Roessler, D. (2007). Age-related accumulation of autosomal mutations in solid tissues of the mouse is gender and cell type specific. *Aging Cell*, *6*(1), 73–86. https://doi.org/10.1111/j.1474-9726.2006.00264.x

Valentine, C. C., Young, R. R., Fielden, M. R., Kulkarni, R., Williams, L. N., Li, T., Minocherhomji, S., & Salk, J. J. (2020). Direct quantification of in vivo mutagenesis and carcinogenesis using duplex sequencing. *Proceedings of the National Academy of Sciences*, *117*(52), 33414–33425. https://doi.org/10.1073/pnas.2013724117

Vijg, J. (2021). From DNA damage to mutations: All roads lead to aging. *Ageing Research Reviews*, *68*, 101316. https://doi.org/10.1016/j.arr.2021.101316

Walter, C. A., Intano, G. W., McCarrey, J. R., McMahan, C. A., & Walter, R. B. (1998). Mutation frequency declines during spermatogenesis in young mice but increases in old mice. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(17), 10015–10019. https://doi.org/10.1073/pnas.95.17.10015

Yousefzadeh, M., Henpita, C., Vyas, R., Soto-Palma, C., Robbins, P., & Niedernhofer, L. (2021). DNA damage—How and why we age? *eLife*, *10*, e62852. https://doi.org/10.7554/eLife.62852

Zhao, J., Ghimire, A., & Liesveld, J. (2021). Marrow failure and aging: The role of "Inflammaging." *Best Practice & Research Clinical Haematology, Bone Marrow Failure Syndromes*, *34*(2), 101283. https://doi.org/10.1016/j.beha.2021.101283

Zhou, G., Williams, A., LeBlanc, D. P. M., Douglas, G. R., Yauk, C. L., & Marchetti, F. (2025). Temporal analyses of germ cell mutations using the MutaMouse model support the recommended design in OECD test guideline 488. *Archives of Toxicology*, *99*(9), 3611–3624. https://doi.org/10.1007/s00204-025-04085-1

CHAPTER FOUR

DISCUSSION
Elena Esina

4.1 Summary of study outcomes

Age-related declines in tissue function are accompanied by molecular changes that accumulate over the lifespan and may contribute to chronic disease (López-Otín et al., 2013; Partridge et al., 2018). Among these changes, nuclear mutations are useful markers of lifelong genomic damage, replication and repair errors, and cellular population dynamics within tissues (Martincorena, 2019; Ren et al., 2022; Vijg, 2021). However, studying age-associated nuclear mutagenesis is challenging because rare mutations are difficult to measure accurately. Existing data also provide an incomplete picture of how nuclear mutation burden changes across tissues with age, leaving current insights fragmented across assays, species, genomic compartments, and tissue contexts. Addressing this gap requires a high-accuracy approach that can assess age-related nuclear mutation accumulation across multiple tissues within a controlled mammalian model.

In this thesis, I investigated whether nuclear mutation frequency (MF) and mutation spectra change with age across male MutaMouse tissues and whether these changes differ by tissue. This question was addressed through two complementary aims: first, to develop and apply a statistical framework for planning Duplex Sequencing (DS)-based mutagenesis studies in mice, and second, to empirically characterize age-related nuclear mutagenesis across multiple tissues in MutaMouse males using DS.

In Aim 1, I applied a simulation-based power analysis to evaluate how DS study-design parameters affect the ability to detect differences in MF. Using baseline MF, sequencing depth, and animal-to-animal variability estimated from previous MutaMouse DS studies, I determined that a sample size of 4–5 animals per group provided greater than 80% power to detect a 1.5-fold or greater difference in MF at typical informative duplex sequencing depths. This analysis also showed that biological variability, expected effect size, sample size, and informative duplex depth jointly determine power in DS experiments, with sample size becoming increasingly important once sequencing depth is sufficient. I used this power analysis to inform the study design for Aim 2, including the selection of sample size targets for the multi-tissue aging study.

In Aim 2, I used DS to conduct a multi-tissue analysis of age-related nuclear mutagenesis in MutaMouse males. I characterized mutation burden, mutation spectra, and genomic-context patterns across liver, bone marrow (BM), lung, and male germ cells. I found that age-related mutation accumulation was strongly tissue-dependent: liver, lung, and BM showed increased MF with age, whereas male germ cells showed no detectable age-related increase over the ages examined. The temporal patterns also differed among somatic tissues, with liver and lung showing more progressive age-related accumulation and BM showing a delayed late-life increase. Differences in mutation spectra and genic-context patterns further suggested that mutation classes and genomic distribution of detected mutations differed across tissues, although these patterns should be interpreted cautiously given the low mutation counts after stratification. These findings support a model in which age-related mutation accumulation is shaped by local tissue biology as well as time.

Together, my analyses and methodologies identify key study-design requirements for applying DS to rare-mutation detection in both basic biology and regulatory genetic toxicology contexts. The aging study supported my overarching hypothesis that age-related nuclear mutagenesis is tissue-dependent, with somatic tissues showing increased MF with age, male germ cells remaining comparatively stable over the ages examined, and mutational patterns varying across tissues.

4.2 Fulfillment of thesis objectives

Objective 1.1: Evaluate how baseline mutation frequencies, sample variability, expected effect sizes, sample sizes, and sequencing depth influence statistical power to detect differences in mutation frequency, using 14-week-old adult MutaMouse tissues as an empirical reference condition.

I developed and applied a simulation-based framework to evaluate how key study-design parameters influence the statistical power of DS-based MF analyses. Using empirical liver and BM data from 14-week-old adult MutaMouse males as reference conditions, I modeled the effects of baseline MF, animal-to-animal variation, expected effect size, sample size, and informative duplex sequencing depth. I found that statistical power increased with

larger sample sizes, greater sequencing depth, and larger expected effect sizes, but decreased as inter-animal variability increased, making true differences in MF harder to distinguish from biological noise. I also found that increasing sequencing depth produced diminishing returns. Within the modeled conditions, increasing informative duplex depth improved power up to a point, after which additional sequencing produced minimal gains and sample size became the more influential design parameter. These analyses provided a quantitative basis for planning DS study designs in adult MutaMouse tissues and informed the sample-size targets used in the multi-tissue aging study. Because this simulation-based approach can be parameterized with user-defined assumptions about MF, variability, effect size, sample size, and sequencing depth, it also provides a framework that other researchers can apply to estimate suitable DS study designs in other model systems and experimental contexts.

Objective 1.2: Use this framework to identify feasible sample-size and sequencing-depth combinations for DS-based mutagenesis studies under realistic MutaMouse experimental conditions, and apply these estimates to inform the design of the empirical aging study.

I used the simulation-based framework to identify practical combinations of sample size and informative duplex sequencing depth for DS-based mutagenesis studies under realistic MutaMouse experimental conditions. My analyses showed that sample sizes of 4–5 animals per group were sufficient to detect at least a 1.5-fold change in MF with a feasible number of informative duplex bases, even under conditions of relatively high inter-animal variability. In contrast, routine detection of smaller changes, particularly 1.2-fold changes in MF, remained challenging because larger sample sizes would be required to compensate for the lower expected effect size and practical limits on sequencing depth. These analyses and tools provide practical guidance for selecting suitable study designs that should be used for MF analysis using DS, and I used these estimates to inform the Chapter 3 aging study design, where the target sample size was set at up to 8 male mice per age group for each tissue.

Objective 2.1: Quantify age-related changes in mutation frequency across tissues.

I used DS to quantify nuclear MF in liver, BM, lung, and germ cells from MutaMouse males across multiple ages. At early ages (3, 14, or 23 weeks), MF was similar across all four tissues. Over the full age range, however, tissue-specific age-related patterns emerged, driven largely by divergence between germ cells and somatic tissues. MF increased with age in all three somatic tissues, with MF_{Min} rising approximately 1.5–2.0-fold between 3 weeks and ≥78 weeks. The shape of this increase differed among tissues: liver and lung showed gradual age-related accumulation, whereas BM showed a delayed increase that became most apparent at ≥78 weeks. In contrast, germ cells did not show a significant age-related increase in either MF_{Min} or MF_{Max}. This lack of a detectable germ cell age effect should be viewed cautiously, as the germ cell analysis had limited power to detect modest age-related changes in MF. These findings show that age-related nuclear mutation accumulation in MutaMouse males is tissue-dependent, with broadly similar accumulation across somatic tissues but no detectable increase in germ cells over the ages examined.

Objective 2.2: Characterize age- and tissue-related differences in mutation spectra.

I characterized mutation spectra across tissues and age groups using broad mutation classes, simple mutation spectra, and exploratory trinucleotide-context signature fitting. These analyses showed that spectra varied across tissues and age groups, with the clearest distinction occurring between germ cells and somatic tissues. Subtype-level analyses suggested possible tissue-specific variation, but sparse mutation counts limited strong conclusions about age-related spectral shifts within individual somatic tissues. Consistent with the absence of an age-related increase in germ cell MF, I did not identify a progressive age-related shift in germ cell spectra. Trinucleotide signature fitting produced low cosine similarity for most reconstructed spectra, limiting confidence in the assignment of specific mutational signatures.

Objective 2.3: Examine how genomic features relate to mutation patterns.

I examined mutation patterns in relation to genic context and positional distribution across the targeted panel. I found that genic context was associated with MF in a tissue-

dependent manner, with higher intergenic MF in BM and lung, no clear genic-intergenic separation in liver, and higher genic MF in germ cells. Within each tissue, however, genic and intergenic regions followed broadly similar age-related patterns, suggesting that genomic context influenced MF in addition to age. Positional mutation mapping identified recurrent signals across the targeted panel, including a prominent chromosome 11 intergenic signal that was more consistent with a local technical artifact than a biological hotspot and was excluded from downstream analyses. Other recurrent sites represent candidate hotspots for future validation. Together, these analyses showed that genomic context influenced mutation patterns in a tissue-dependent manner and identified positional signals that can guide future studies.

4.3 Contribution to scientific knowledge

This thesis contributes to scientific knowledge in two main areas: (1) by providing a quantitative framework for designing DS mutagenesis studies; and (2) by generating controlled multi-tissue nuclear DNA data showing that age-related mutation accumulation is tissue-dependent in MutaMouse males and should be interpreted in relation to local biological context and broader mutational patterns.

4.3.1 Establishing an *a priori* study-design framework for Duplex Sequencing mutagenesis studies

A major contribution of this thesis is the development of an *a priori* approach for designing DS-based mutagenesis studies. Prior to this work, limited guidance was available for prospectively selecting sample sizes and sequencing-depth targets before data generation. DS experiments require substantial funding, personnel time, DNA input, library preparation reagents, sequencing capacity, and downstream analysis. Inefficient study designs therefore carry an opportunity cost, because resources spent on unnecessary sequencing or poorly chosen sample sizes cannot be redirected toward additional experimental questions or validation work. Animal use adds a separate ethical concern. Underpowered studies may be unable to detect the expected effect size, making negative results difficult to interpret and reducing the value of the animals already used. Conversely, unnecessarily large studies increase animal use without improving the scientific conclusion. By evaluating sample size and

informative duplex sequencing depth prospectively, this approach supports more ethically justified DS experimental designs. It also defines feasible study designs across combinations of expected effect size, inter-animal variability, and achievable sequencing depth, providing a rational basis for DS-based mutation studies that use available resources efficiently.

DS is increasingly being evaluated as an alternative approach for *in vivo* mutagenicity assessment (Dodge et al., 2023; LeBlanc et al., 2022, 2025), including exposure-based studies where defined effect sizes must be detected with sufficient confidence. The Chapter 2 power-analysis outputs and associated study-design tools support ongoing international efforts to evaluate error-corrected sequencing technologies for regulatory toxicology, including efforts to develop harmonized approaches for study design and potential application in OECD genetic toxicology test guidelines (Yauk et al., 2026). The scenario-based results allow researchers to estimate the number of animals per group required under different assumptions about the system being studied, including cases where tissue-level variability is high, sequencing depth is constrained, or the minimum effect size of interest is modest. I also provide a worked example of this approach: the empirical aging study in Chapter 3 used the Chapter 2 power analysis to target up to eight male mice per age group for each tissue. This design supported detection of age-related MF changes in the three somatic tissues and clarified why smaller age-related effects in germ cells could not be excluded. Overall, this work contributes prospective design guidance for DS-based mutagenesis studies and demonstrates how an *a priori* power analysis can be incorporated before data generation.

4.3.2 Providing controlled evidence for tissue-dependent age-related nuclear mutagenesis

Age-related mutagenesis has been studied across many systems, but the existing literature makes it difficult to determine how nuclear mutations accumulate across tissues during aging. Prior work has shown that mutation burden can change with age, yet these studies differ in several key ways (as reviewed in Chapter 1): (1) reporter loci versus endogenous DNA; (2) mitochondrial versus nuclear genomes; (3) human data versus controlled animal studies; (4) accessible tissues, such as blood and sperm, versus internal organs that are harder to sample systematically; and (5) different sequencing and assay approaches. These

differences do not make the existing literature uninformative, but they limit the ability to distinguish general features of age-related mutagenesis from patterns shaped by assay design, species, genomic compartment, or tissue context. I reduced some of this ambiguity by applying the same high-accuracy DS method to study age-related nuclear DNA mutation accumulation across multiple tissues in a controlled MutaMouse model. By holding several major sources of biological and technical variation constant, I could evaluate differences among tissues more directly.

Using this integrated design, I showed in Chapter 3 that age-related nuclear mutation accumulation was structured by tissue context at the level of overall mutation burden. Early-life MFs were broadly similar across liver, lung, BM, and germ cells, suggesting a shared low baseline mutation burden. Across age, however, the tissues diverged. Liver and lung showed patterns consistent with gradual accumulation, BM showed a delayed increase that became most apparent in later life, and germ cells showed no detectable age-related increase over the ages examined. Analyses of mutation spectra and genomic context further suggested that tissue-specificity was evident in the composition and genomic distribution of mutation burden, as well as in overall MF. Therefore, age-related nuclear mutagenesis in MutaMouse seems best understood as a set of tissue-specific trajectories emerging from a shared early-life baseline, with the clearest divergence occurring between somatic tissues and the male germline and more modest differences among somatic tissues. The persistence of tissue-specific patterns under a standardized experimental design supports tissue context as an important determinant of how nuclear mutations accumulate with age.

Age-related mutation burden should be considered in relation to the tissue in which it is measured. An MF estimate from one tissue cannot serve as a whole-organism summary, because similar burden-level patterns do not necessarily indicate that the same age-related processes are operating across tissues. Tissue selection should therefore be aligned explicitly with the biological question, and MF should be evaluated alongside broader mutational patterns, including sequence context and genomic distribution, to clarify how chronological age and tissue context jointly shape nuclear mutagenesis.

4.4 Future directions

This thesis establishes DS as a high-accuracy approach for measuring rare nuclear mutations across aging MutaMouse tissues and shows that age-related mutagenesis varies by tissue context. Building on these findings, future studies should refine DS study designs for specific biological questions, resolve mutation patterns at finer tissue and cell-population resolution, identify mechanisms underlying tissue-specific trajectories, clarify how mouse age should be compared with human age, and validate recurrent positional signals in the targeted panel. These directions would strengthen the design and interpretation of DS-based aging studies and help define if and how MutaMouse can be used most effectively to study age-related nuclear mutagenesis.

4.4.1 Moving from general-purpose mutation assessment toward question-driven genomic approaches

The Mouse Mutagenesis Panel used in this thesis was designed to sample putatively neutral regions that are broadly representative of the mouse genome (LeBlanc et al., 2022). This made it well suited for estimating spontaneous MF and comparing overall age-related mutation accumulation across tissues. The present work therefore provides proof-of-concept evidence that targeted DS can be used to study spontaneous age-related nuclear mutagenesis in control animals, where expected effect sizes may be modest and mutation counts are low.

A key future direction is to move from general estimates of mutation burden toward genomic regions and endpoints chosen for specific biological questions. For studies focused on overall mutation accumulation, representative neutral panels may remain appropriate. However, studies focused on functional consequences or disease relevance would benefit from panels that include coding and regulatory loci selected for relevance to age-related disease. Studies focused on mechanism could instead target regions that differ in transcriptional activity, chromatin state, replication timing, or other genomic features expected to influence regional mutation accumulation. As sequencing costs continue to decline, larger targeted panels and broader error-corrected sequencing approaches may make it possible to move beyond representative sampling toward more comprehensive, question-driven analysis.

Future studies should also consider which classes of genomic alteration are most relevant to the biological question. The DS analyses used here primarily assessed SNVs and small indels, but aging-associated genome instability may also involve larger copy-number changes, large deletions, rearrangements, and aneuploidy. These endpoints would require adapted DS strategies or complementary genomic methods, but including them would provide a broader view of how different forms of genomic instability accumulate with age.

4.4.2 Resolving tissue-, cell-type-, and sex-specific trajectories of age-related mutation accumulation

Future work should build on this thesis by resolving age-related mutation accumulation at finer biological resolution. The present study used DS on whole-tissue DNA extracts, which provided a controlled tissue-level view of nuclear mutagenesis across liver, lung, BM, and male germ cells. This approach was appropriate for comparing mutational trajectories across tissues with distinct biological properties. However, tissue-level analysis cannot determine whether the observed mutation patterns are distributed broadly across cells or driven by specific cell populations.

Stage-resolved analyses would be particularly valuable for clarifying the male germ cell findings. DNA extracted from seminiferous tubule preparations represents a mixture of germ cells at multiple stages of spermatogenesis, together with a smaller contribution from somatic support cells. A stable MF estimated from these preparations could therefore obscure differences among spermatogonial, meiotic, and post-meiotic germ cell populations. Earlier *lacI* reporter work using enriched spermatogenic cell populations found that mutant frequency varied across spermatogenesis and increased in spermatogenic cells from old mice, suggesting that mutation burden can differ by germ cell stage and age (Walter et al., 1998). Future DS studies that isolate defined spermatogenic stages would provide a higher-resolution view of male germline aging and help identify which mutation patterns are most relevant to the cells that ultimately contribute to sperm.

Similar considerations apply to somatic tissues. Whole-tissue DNA reflects a composite of resident functional cells, tissue-specific regenerative or progenitor compartments, immune populations, support cells, and age-related changes in cell abundance or clonal structure. In BM, cell-population-specific analyses would help distinguish mutation accumulation within haematopoietic stem and progenitor cells from shifts in differentiated lineages or clonal expansion. In liver and lung, compartment-enriched analyses could help determine whether age-related mutation burden is broadly distributed or concentrated in particular cell populations.

Future studies should also include additional tissues and female mice. High-turnover epithelial tissues such as skin or intestine, and protected tissues such as brain, could provide useful contrasts in tissue turnover, environmental exposure, barrier function, and cellular composition. Including female somatic tissues and oocytes would allow direct assessment of whether the tissue-specific trajectories observed in males are conserved and would inform sex-specific aging. Together, these extensions would help determine how broadly the patterns observed in this thesis apply across age, tissue type, cell population, and sex.

4.4.3 Identifying mechanisms that shape tissue-specific mutagenesis

The tissue-specific trajectories observed in this thesis raise the question of which biological processes drive differences in age-related nuclear mutagenesis among tissues. The present work used mutation burden, spectra, and genomic distribution to assess mutational patterns across liver, lung, BM, and male germ cells, but these endpoints do not directly identify their upstream biological sources. Future studies should therefore pair DS with tissue-matched measurements of candidate mechanisms, including DNA damage burden, repair pathway activity, transcriptional and chromatin state, inflammatory environment, tissue turnover, or clonal dynamics.

Mechanistic interpretation will also require designs that are powered for the mutation endpoints being used. Genic-context comparisons, mutation spectra, mutational signatures, and positional analyses all provide information beyond total mutation burden, but each

becomes increasingly difficult to interpret when rare mutation counts are subdivided across categories. Future mechanistic studies may therefore require larger targeted panels, more informative duplex bases, additional animals, or a combination of these approaches to distinguish true biological similarity from insufficient power.

Future work could test specific mechanisms suggested by the tissue-specific patterns observed here. The delayed increase in BM mutation burden could be examined alongside age-related remodeling of the haematopoietic stem and progenitor compartment, including inflammatory niche changes and clonal skew. In liver and lung, future work could test for alignment with tissue-specific metabolic stress, environmental exposure, transcriptional activity, or repair pathway activity. In germ cells, mechanistic studies could evaluate whether low mutation burden reflects germline genome-maintenance and quality-control processes. Linking DS mutation profiles to tissue-matched biological measurements would help future studies move beyond describing tissue-specific trajectories and toward testing the processes that shape them.

4.4.4 Defining appropriate age scales and study systems for cross-species mutagenesis studies

Future work should clarify how chronological and biological age should be used when comparing MutaMouse mutagenesis patterns with human aging. The controlled MutaMouse design used in this thesis was valuable because it allowed mutation accumulation to be compared across tissues within a single mammalian model. However, mouse age-related trajectories should not be mapped directly onto human aging using chronological age alone. Since mouse and human aging differ in lifespan, environmental context, reproductive biology, disease susceptibility, and the timing of age-related decline (Brunet, 2020; Demetrius, 2006), future studies should therefore define the most relevant age scale and study system for the tissue or cell type being examined.

For somatic tissues, chronological age relative to lifespan stage may provide a useful basis for cross-species comparison. Old mouse somatic tissues are not the same absolute age as old human somatic tissues, but they may represent a comparable point along the lifespan of

each species. The age-related increases in liver, lung, and BM MF observed in this thesis may therefore reflect mutational processes that are also relevant to late-life human somatic tissues. Future comparisons between MutaMouse and human somatic datasets should focus on whether broad late-life mutational patterns are conserved across species.

The germ cell findings require a different interpretation. In this thesis, germ cell MF did not show a detectable age-related increase over the ages examined, whereas human studies have reported paternal-age-associated increases in germline mutation burden, including evidence from sperm sequencing (Neville et al., 2025; Shoag et al., 2025). This apparent difference may mean that germ cells may be better compared using biological germline age rather than organismal chronological age. In Chapter 3, the oldest mouse examined in the ≥78-week cohort was 85 weeks old, which corresponded to only approximately 2.8 human post-pubertal germline years, or roughly a 17.8-year-old human in germ cell terms. Future germ cell studies should therefore identify which age scale makes the cross-species comparison biologically meaningful. Mice have a much shorter lifespan than humans, so MutaMouse males may not be able to model the biological germline age reached by older human males. Future studies of paternal-age-associated germline mutagenesis may therefore be better addressed directly in human sperm, which can be sampled post-puberty across much of the human lifespan. MutaMouse germ cell studies remain useful for controlled studies of murine germline biology, but the short mouse lifespan may limit their utility for modeling late-life human male germline aging.

4.4.5 Validating and refining targeted-panel signals

In this thesis, positional mapping revealed that most mutations were distributed across the Mouse Mutagenesis Panel, but a small number of positions showed recurrent signals across samples and tissues. Fixed-panel DS repeatedly interrogates the same genomic positions, so recurrent variants may reflect real mutation-prone sites or target-specific technical features. The chromosome 11 intergenic signal resolved in this thesis (§3.4.3) illustrates this concern, as its recurrence pattern and local sequence features were more consistent with a technical

artifact than with a biological hotspot. Other recurrent sites were flagged but not resolved and therefore remain important targets for future analyses.

Future work should characterize the remaining recurrent positions because these sites affect how the panel should be used. Technical recurrent signals could bias MF estimates, whereas biological recurrent signals may identify true local mutation enrichment. Distinguishing between these possibilities is therefore necessary before recurrent positional signals are used to support biological conclusions. Resolving these sites would also provide empirical validation of the Mouse Mutagenesis Panel by identifying positions that can be retained, flagged, redesigned, or excluded in future DS studies.

Characterizing these positions would also clarify whether recurrent positional signals provide useful information about age-related mutation accumulation across tissues. If some recurrent signals are biological, their distribution across tissues would become important for evaluating their biological relevance. A biological signal that appears across many tissues may reflect a sequence context that is mutation-prone in many cell types, whereas a signal that appears mainly in one tissue could suggest a local process that is more active in that tissue. A better-resolved panel would strengthen both the technical reliability of targeted DS and the ability to identify positional patterns that may reveal local features of tissue-specific mutation accumulation.

4.5 Concluding remarks

This thesis demonstrates the value of DS for studying rare nuclear mutations across aging tissues in a controlled mammalian model. The simulation-based framework developed in Chapter 2 showed that accurate DS mutation detection must be paired with sufficient statistical power to support meaningful biological conclusions. Applying this framework to MutaMouse male tissues revealed that age-related nuclear mutagenesis is tissue-dependent rather than uniform across the organism. MF increased with age in somatic tissues, whereas no detectable age-related increase was observed in male germ cells over the ages examined. Mutation spectra and genomic-context analyses further showed that tissue identity shapes broader

mutational patterns as well. Overall, this thesis provides two complementary contributions to the study of age-related nuclear mutagenesis: analytical guidance for designing DS-based mutation studies, and empirical evidence that mutation accumulation is shaped by tissue context in the MutaMouse male.

4.6 References

Abascal, F., Harvey, L. M. R., Mitchell, E., Lawson, A. R. J., Lensing, S. V., Ellis, P., Russell, A. J. C., Alcantara, R. E., Baez-Ortega, A., Wang, Y., Kwa, E. J., Lee-Six, H., Cagan, A., Coorens, T. H. H., Chapman, M. S., Olafsson, S., Leonard, S., Jones, D., Machado, H. E., … Martincorena, I. (2021). Somatic mutation landscapes at single-molecule resolution. *Nature*, *593*(7859), 405–410. https://doi.org/10.1038/s41586-021-03477-4

Arbeithuber, B., Hester, J., Cremona, M. A., Stoler, N., Zaidi, A., Higgins, B., Anthony, K., Chiaromonte, F., Diaz, F. J., & Makova, K. D. (2020). Age-related accumulation of de novo mitochondrial mutations in mammalian oocytes and somatic tissues. *PLOS Biology*, *18*(7), e3000745. https://doi.org/10.1371/journal.pbio.3000745

Brunet, A. (2020). Old and new models for the study of human ageing. *Nature Reviews Molecular Cell Biology*, *21*(9), 491–493. https://doi.org/10.1038/s41580-020-0266-4

Demetrius, L. (2006). Aging in Mouse and Human Systems. *Annals of the New York Academy of Sciences*, *1067*(1), 66–82. https://doi.org/10.1196/annals.1354.010

Dodge, A. E., LeBlanc, D. P. M., Zhou, G., Williams, A., Meier, M. J., Van, P., Lo, F. Y., Valentine Iii, C. C., Salk, J. J., Yauk, C. L., & Marchetti, F. (2023). Duplex sequencing provides detailed characterization of mutation frequencies and spectra in the bone marrow of MutaMouse males exposed to procarbazine hydrochloride. *Archives of Toxicology*, *97*(8), 2245–2259. https://doi.org/10.1007/s00204-023-03527-y

Kennedy, S. R., Salk, J. J., Schmitt, M. W., & Loeb, L. A. (2013). Ultra-Sensitive Sequencing Reveals an Age-Related Increase in Somatic Mitochondrial Mutations That Are Inconsistent with Oxidative Damage. *PLOS Genetics*, *9*(9), e1003794. https://doi.org/10.1371/journal.pgen.1003794

LeBlanc, D. P. M., Meier, M., Lo, F. Y., Schmidt, E., Valentine, C., Williams, A., Salk, J. J., Yauk, C. L., & Marchetti, F. (2022). Duplex sequencing identifies genomic features that determine susceptibility to benzo(a)pyrene-induced in vivo mutations. *BMC Genomics*, *23*(1), 542. https://doi.org/10.1186/s12864-022-08752-w

LeBlanc, D. P. M., Zhou, G., Williams, A., Meier, M. J., Valentine, C. C., Salk, J. J., Yauk, C. L., & Marchetti, F. (2025). Duplex sequencing identifies unique characteristics of ENU-induced mutations in male mouse germ cells. *Biology of Reproduction*, *112*(5), 1015–1027. https://doi.org/10.1093/biolre/ioaf029

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The Hallmarks of Aging. *Cell*, *153*(6), 1194–1217. https://doi.org/10.1016/j.cell.2013.05.039

Martincorena, I. (2019). Somatic mutation and clonal expansions in human tissues. *Genome Medicine*, *11*, 35. https://doi.org/10.1186/s13073-019-0648-4

Neville, M. D. C., Lawson, A. R. J., Sanghvi, R., Abascal, F., Pham, M. H., Cagan, A., Nicola, P. A., Bayzetinova, T., Baez-Ortega, A., Roberts, K., Lensing, S. V., Widaa, S., Alcantara, R. E., García, M. P., Wadge, S., Stratton, M. R., Campbell, P. J., Small, K., Martincorena, I., … Rahbari, R. (2025). Sperm sequencing reveals extensive positive selection in the male germline. *Nature*, *647*(8089), 421–428. https://doi.org/10.1038/s41586-025-09448-3

Partridge, L., Deelen, J., & Slagboom, P. E. (2018). Facing up to the global challenges of ageing. *Nature*, *561*(7721), Article 7721. https://doi.org/10.1038/s41586-018-0457-8

Ren, P., Dong, X., & Vijg, J. (2022). Age-related somatic mutation burden in human tissues. *Frontiers in Aging*, *3*. https://www.frontiersin.org/articles/10.3389/fragi.2022.1018119

Sanchez-Contreras, M., Sweetwyne, M. T., Tsantilas, K. A., Whitson, J. A., Campbell, M. D., Kohrn, B. F., Kim, H. J., Hipp, M. J., Fredrickson, J., Nguyen, M. M., Hurley, J. B., Marcinek, D. J., Rabinovitch, P. S., & Kennedy, S. R. (2023). The multi-tissue landscape of somatic mtDNA mutations indicates tissue-specific accumulation and removal in aging. *eLife*, *12*, e83395. https://doi.org/10.7554/eLife.83395

Shoag, J. E., Srinivasa, A., Loh, C. A., Liu, M. H., Lassen, E., Melanaphy, S., Costa, B. M., Grońska-Pęski, M., Jabara, N. T., Picciotto, S., Choi, U., Bohorquez, A. D., Barbieri, C. E., Callum, P., Skytte, A.-B., & Evrony, G. D. (2025). Direct measurement of the male germline mutation rate in individuals using sequential sperm samples. *Nature Communications*, *16*(1), 2546. https://doi.org/10.1038/s41467-025-57507-0

Turker, M. S., Lasarev, M., Connolly, L., Kasameyer, E., & Roessler, D. (2007). Age-related accumulation of autosomal mutations in solid tissues of the mouse is gender and cell type specific. *Aging Cell*, *6*(1), 73–86. https://doi.org/10.1111/j.1474-9726.2006.00264.x

Vijg, J. (2021). From DNA damage to mutations: All roads lead to aging. *Ageing Research Reviews*, *68*, 101316. https://doi.org/10.1016/j.arr.2021.101316

Walter, C. A., Intano, G. W., McCarrey, J. R., McMahan, C. A., & Walter, R. B. (1998). Mutation frequency declines during spermatogenesis in young mice but increases in old mice. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(17), 10015–10019. https://doi.org/10.1073/pnas.95.17.10015

Yauk, C. L., Lynch, A. M., Dobrovolsky, V. N., Schuler, M., Smith-Roe, S. L., Fitzgerald, D., Higgins, J., Honarvar, N., Le Curieux, F., Matsumura, S., Minocherhomji, S., Recio, L., Salk, J. J., Sugiyama, K.-I., Suzuki, T., Wills, J. W., & Marchetti, F. (2026). Application of error-corrected sequencing technologies for in vivo regulatory mutagenicity assessment. *Regulatory Toxicology and Pharmacology: RTP*, *164*, 105985. https://doi.org/10.1016/j.yrtph.2025.105985

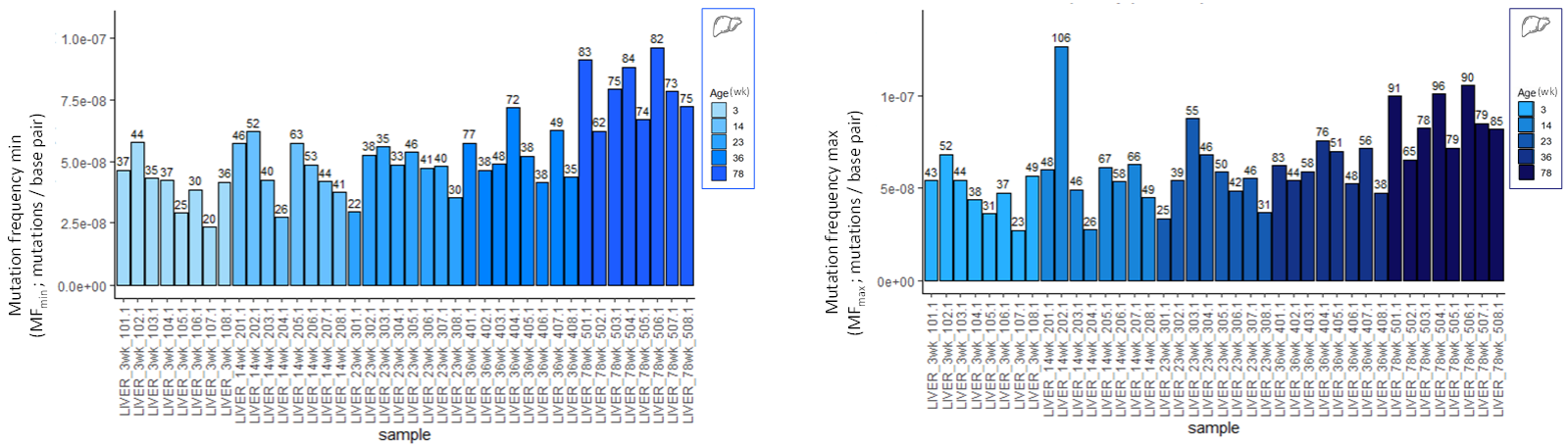

Figure S3.1: Duplex Sequencing mutation frequencies (MF_{Min} on the left and MF_{Max} on the right) in the liver of MutaMouse males at 3, 14, 23, 36, and ≥78 weeks by sample. Numbers above bars represent mutation counts in each sample.

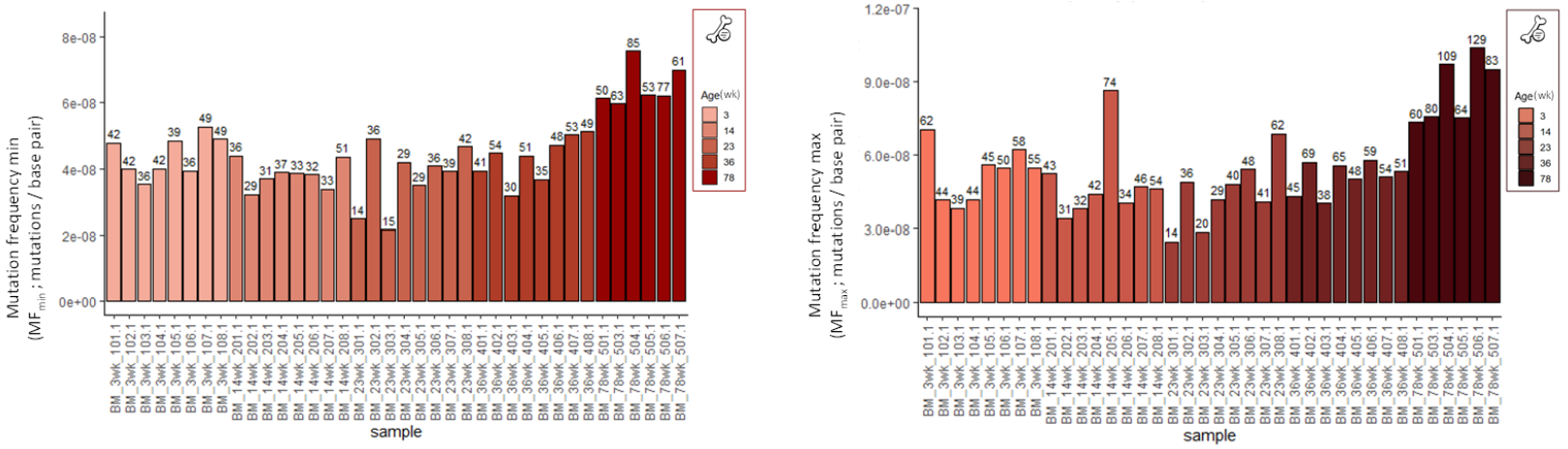

Figure S3.2: Duplex Sequencing mutation frequencies (MF_{Min} on the left and MF_{Max} on the right) in the bone marrow of MutaMouse males at 3, 14, 23, 36, and ≥78 weeks by sample. Numbers above bars represent mutation counts in each sample.

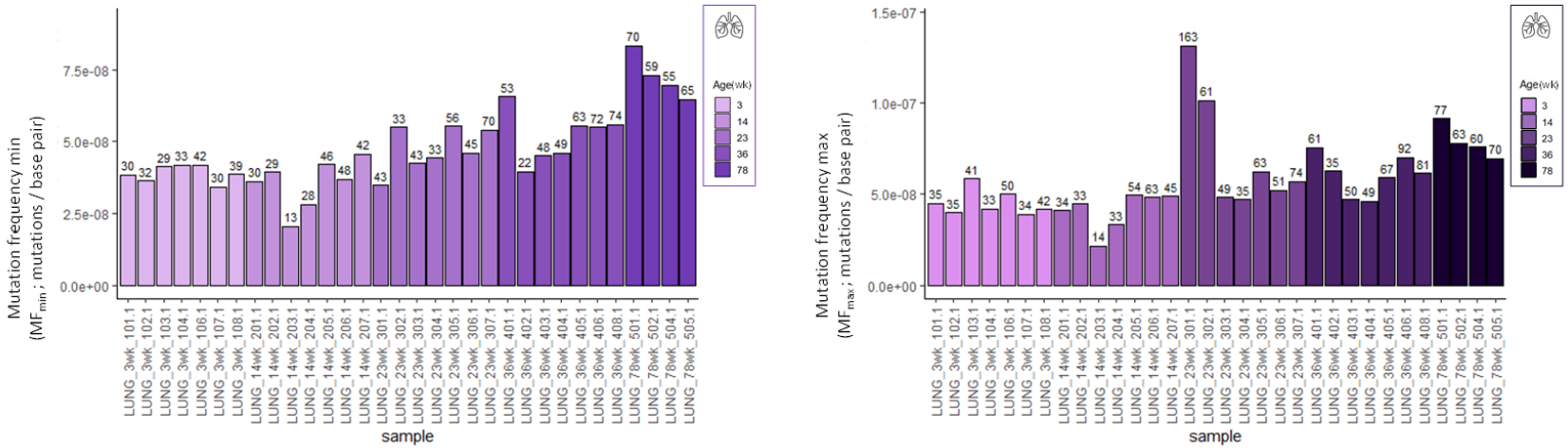

Figure S3.3: Duplex Sequencing mutation frequencies (MF_{Min} on the left and MF_{Max} on the right) in the lung of MutaMouse males at 3, 14, 23, 36, and ≥78 weeks by sample. Numbers above bars represent mutation counts in each sample.

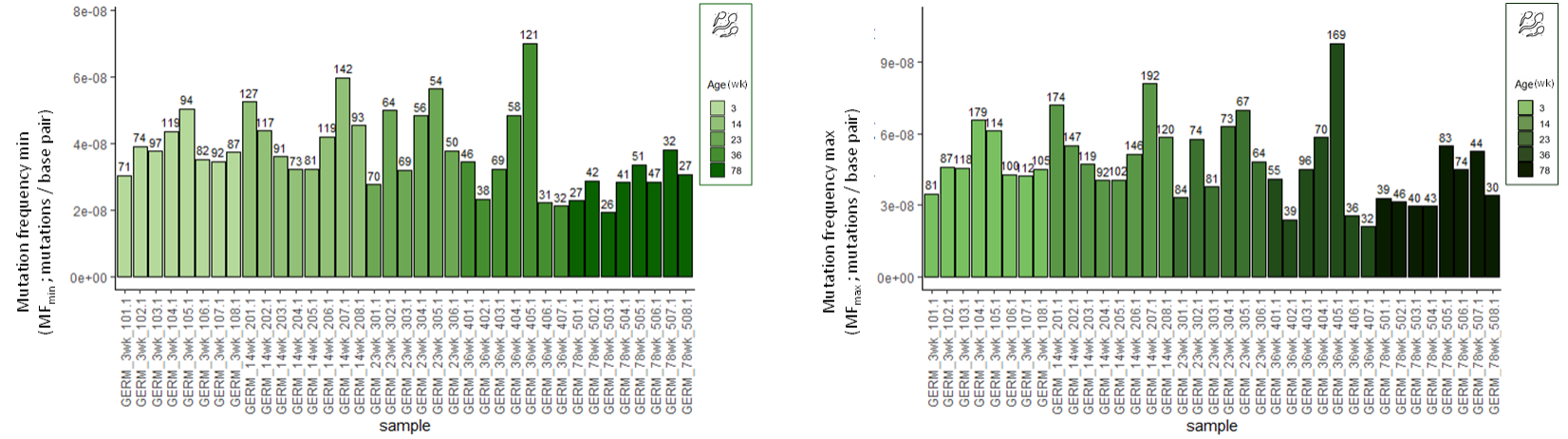

Figure S3.4: Duplex Sequencing mutation frequencies (MF_{Min} on the left and MF_{Max} on the right) in the germ of MutaMouse males at 3, 14, 23, 36, and ≥78 weeks by sample. Numbers above bars represent mutation counts in each sample.

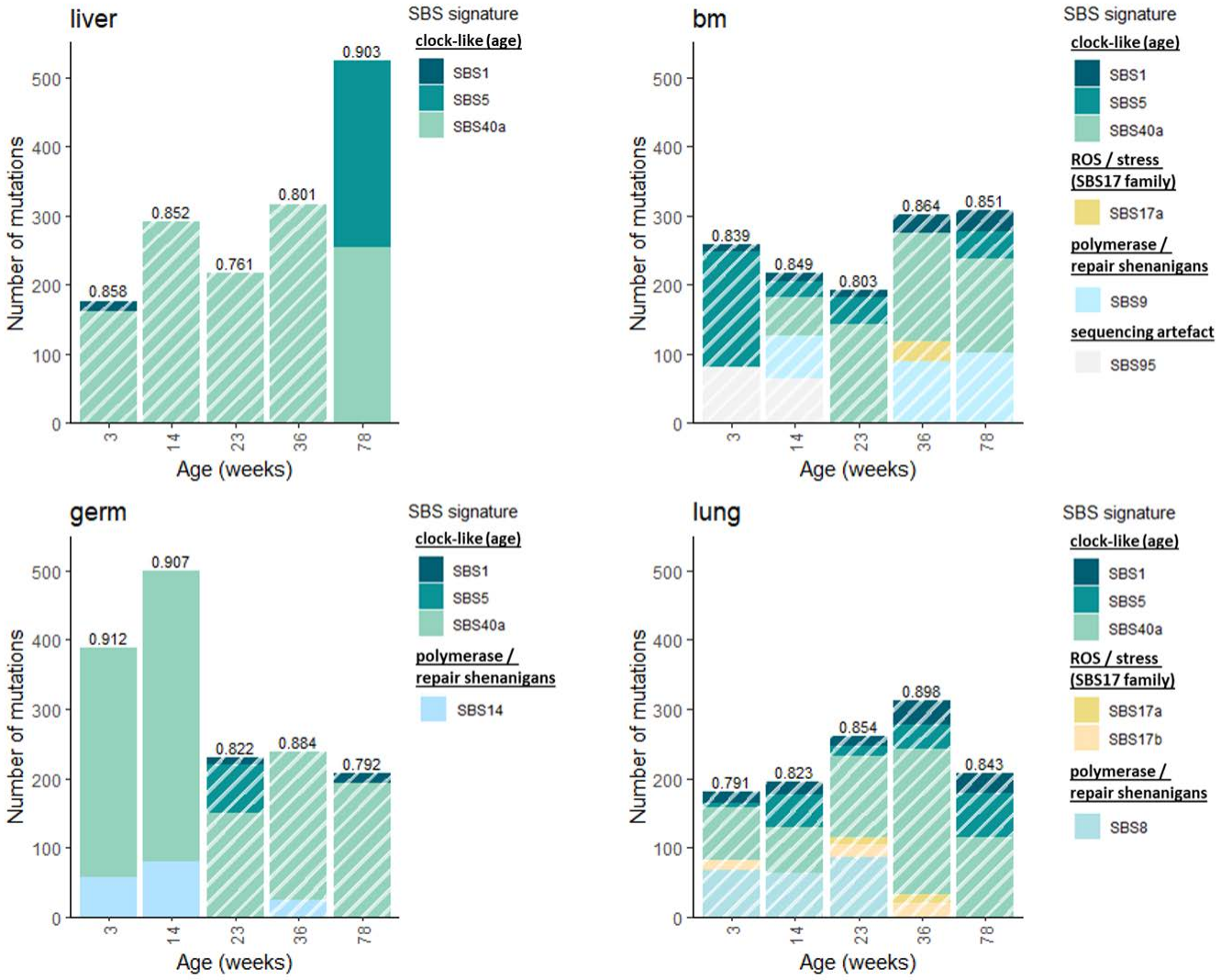

Figure S3.5: Exploratory SBS mutational signature profiles across tissue and age. Stacked bars show the number of mutations assigned to COSMIC v3.4 SBS signatures in liver, bone marrow, germ cells, and lung across age groups. Values above bars indicate cosine similarity between observed and reconstructed spectra; hatched bars indicate cosine similarity <0.90 and should be interpreted cautiously.

Table S3.1: Summary of recurrent positional mutation spikes identified across tissues during descriptive positional mapping.

Target	Relative position(s)	Event(s) observed	Local sequence-context note	Included in detailed follow-up
chr11	2351, 2356, 2363	CA→A; TG→T; marked drop in informative duplex bases	In or adjacent to multiple homopolymer runs	Yes (Figure 3.4)
chr13	356, 374	620-nt deletion-like event→T; 624-nt deletion-like event→T; TA→A	Position 377 marks the start of a homopolymer run; proximity noted	No
	990, 991	ACCAG→A; A→T; CCAGGGC→GGT; C→A	GC-rich local interval; position 1000 marks the start of a homopolymer run	
chr1.2	485, 501	727-nt deletion-like event→A; A→G; G→A; 727-nt deletion-like event→G	Position 501 lies at the end of a homopolymer run	No
	1168, 1191	G→C; G→A; G→–; GCAC→ACAT; GCACA→ACATG; GCACAGTA→ACATGGTG; GCACAGTACAA→ACATGGTGCAG	No obvious local sequence feature identified	
chr15	1791	T→G	Position 1792 marks the start of a homopolymer run	No
chr12	580	CT→C; C→A	Position 581 marks the start of a homopolymer run	No

Mutation spikes were defined descriptively as positions showing visibly elevated counts of mutated samples across tissues in the global positional maps (Figure 3.3). Event labels summarize the recurrent alternate calls observed at each position or local cluster. Only the chromosome 11 target was examined further in a composite positional plot in the main text (Figure 3.4).