

STOCHASTIC MODELLING AND INFERENCE FOR MITOCHONDRIA IN HEALTH AND DISEASE

A THESIS PRESENTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY OF IMPERIAL COLLEGE LONDON
BY
JUVID ARYAMAN

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE LONDON

NOVEMBER 2018

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

COPYRIGHT

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Stochastic Modelling and Inference for Mitochondria in Health and Disease

ABSTRACT

Mitochondria are crucial organelles for complex cellular life: their dysfunction causes multiple devastating heritable diseases, and is associated with healthy ageing. Cells possess a population of mitochondrial DNAs (mtDNA) which are continuously turned over. Consequently, mtDNA mutations may arise and expand throughout the cell. Here, we aim to use stochastic modelling to deepen our understanding of mitochondrial populations, using inference techniques to connect with data, to provide rational hypotheses for biomedical interventions.

Mitochondria are not static organelles: they form dynamic networks through organellar fusion and fission. We provide the first analytically-derived bridge between mitochondrial genetics and network dynamics. We show that the rate of increase in cell-to-cell variability in mitochondrial mutant load, as well as the rate of *de novo* mutation generation, is modulated by the fraction of mitochondria which are unfused from the mitochondrial network.

We suggest that cells respond to rising mutant load of a particular mtDNA mutant by initially reducing in volume to maintain the density of unmutated mtDNAs and reduce power demand. When a minimum cell volume is reached, we suggest that cells then switch to power supply production through the induction of bioenergetic pathways.

A non-linear relationship between cell size and mitochondrial functionality has recently been experimentally observed. We suggest that a metabolic scaling argument, combined with a simple model of cell death, plausibly accounts for the observed non-linearity.

Finally, we explore how replication errors in nuclear DNA during neurodevelopment may seed regions of pathologically mutated neurons in the adult brain. By modelling neurodevelopment as a deterministic branching process, we infer the mutation rate during neurodevelopment from a recent dataset. We extrapolate from our model to suggest that pathological islands consisting of approximately 10^5 neurons may be common.

To my family,

ACKNOWLEDGMENTS

I would like to thank my supervisors Nick Jones, Iain Johnston and Mauricio Barahona for giving me the opportunity to pursue this PhD. I would like to give special thanks to my main supervisors Nick and Iain for their countless hours of patient supervision, continual insights, and support. It has been a real privilege to work with, and learn from, you both.

Thank you to Patrick Chinnery, Michael Keogh and Wei Wei for involving me in a number of exciting collaborations. Thank you to Tristan Rodriguez for having me in his lab, and to Ana Lima for giving me cell culture lessons. Thank you to Gyorgy Szabadkai for also having me in his lab, and teaching me about microscopy.

Thank you to everyone in the Systems and Signals group I have had the pleasure to work with: Rory Brittain, Alex De Figueiredo, Abhishek Deshpande, Matt Garrod, Antonia Godoy-Lorite, Sam Greenbury, Luke Heaton, Till Hoffmann, Ben Ingledow, Ferdinando Insalata, Thomas Ouldrige, and Sarab Sethi. Thank you for making this such an interesting and enjoyable place to work.

Thank you to my friends Peter Bingham, Elwin Carlos, Jakov Marelic, and Stephen Stopyra for the hikes, their conversation, and support. Thank you to all of the inspirational teachers I have had over the years, especially to Dr. Holding for teaching me to look for what is meaningful, and to Dr. Galgut for helping me find my way.

I would like to give special thanks to Tom McGrath and Hanne Hoitzing. I am so grateful for having gotten to know you both over these years. Tom, thank you for all the lunchtime chats, and for always being thought-provoking. Hanne, thank you for your kindness, for making me laugh, and for making this period of my life richer.

Finally, I would like to thank my family, to whom I dedicate this thesis. Without your love and support, this work would not have been possible. Thank you for being there, through it all.

PUBLICATIONS & CONTRIBUTIONS

Mitochondrial heterogeneity, (submitted). Preprint: arXiv:1809.01878 (2018)

J. Aryaman, I. G. Johnston, and N. S. Jones

Contribution: performed literature survey and wrote the manuscript.

Appearance in thesis: Chapter 1.

Mitochondrial network fragmentation modulates mutant mtDNA accumulation independently of absolute fission-fusion rates, (submitted). Preprints: arXiv, 1809.01882 (2018); bioRxiv, 409128 (2018)

J. Aryaman, C. Bowles, N. S. Jones, and I. G. Johnston

Contribution: Refined the model structure, and also expanded and generalised the analysis to include copy-number control, identified the role of fragmentation on heteroplasmy variance, proved its relevance analytically, extended to wider classes of systems, made links to the Moran process and identified the role of intermediate fusion in the case of selection. Performed computational implementation of the analysis and wrote the manuscript.

Appearance in thesis: Chapter 2.

Mitochondrial DNA density homeostasis accounts for a threshold effect in a cybrid model of a human mitochondrial disease, Biochemical Journal, 474, 4019-4034 (2017)

J. Aryaman, I. G. Johnston, and N. S. Jones

Contribution: performed all simulations, analysis and wrote the manuscript.

Appearance in thesis: Chapter 3.

Mitochondrial heterogeneity, metabolic scaling and cell death, BioEssays, 39(7), 1700001 (2017)

J. Aryaman, H. Hoitzing, J. Burgstaller, I. G. Johnston, and N. S. Jones

Contribution: performed all simulations, analysis and wrote the manuscript.

Appearance in thesis: Chapter 4.

High prevalence of focal and multi-focal somatic genetic variants in the human brain, Nature Communications, Nature Communications, 9, 4257 (2018)

M. J. Keogh, W. Wei, J. Aryaman, L. Walker, J. van den Ameele, J. Coxhead, I. Wilson, M. Bashton, J. Beck, J. West, R. Chen, C. Haudenschild, G. Bartha, S. Luo, C. M. Morris, N. S. Jones, J. Attems and P. F. Chinnery

Contribution: designed mathematical analysis of sequencing data, performed mathematical and statistical analysis. Wrote the Supplementary Methods and Supplementary Notes.

Appearance in thesis: Chapter 5.

Oligogenic genetic variation of neurodegenerative disease genes in 980 postmortem human brains, Journal of Neurology, Neurosurgery & Psychiatry, 89(8), 813-816 (2018)

M. J. Keogh, W. Wei, J. Aryaman, I. Wilson, K. Talbot, M. R. Turner, C-A. McKenzie, C. Troakes, J. Attems, C. Smith, S. Al Sarraj, C. M. Morris, O. Ansorge, S. Pickering-Brown, N. S. Jones, J. W. Ironside, P. F. Chinnery

Contribution: Designed mathematical analysis of sequencing data, performed mathematical and statistical analysis. Wrote Supplementary File 2.

Appearance in thesis: Not included.

Frequency and signature of somatic variants in 1461 human brain exomes, Genetics in Medicine (2018)

W. Wei, M. J. Keogh, J. Aryaman, Z. Golder, P. J. Kullar, I. Wilson, K. Talbot, M. R. Turner, C-A. McKenzie, C. Troakes, J. Attems, C. Smith, S Al. Sarraj, CM. Morris, O. Ansorge, N. S. Jones, J. W. Ironside, P. F. Chinnery

Contribution: Designed statistical analysis of sequencing data, performed statistical analysis.

Appearance in thesis: Not included.

CONTENTS

1	INTRODUCTION	1
1.1	Basics of mitochondrial physiology	1
1.2	Mitochondrial heterogeneity	2
1.2.1	Genetic sources of mitochondrial heterogeneity	3
1.2.1.1	Inter-cellular mtDNA copy number appears to be under homeostatic control	5
1.2.1.2	Intra-cellular mutations in mitochondrial DNA are a source of genotypic heterogeneity	6
1.2.1.3	Inter-cellular mutations in mitochondrial DNA are a source of genotypic heterogeneity	7
1.2.2	Non-genetic sources of mitochondrial heterogeneity	9
1.2.2.1	Stoichiometric and structural heterogeneity in the inner mitochondrial membrane as a potential amplifier of the genotype-phenotype link	10
1.2.2.2	Mitochondrial membrane potential heterogeneity, mitochondrial networks and mitochondrial genotype	12
1.2.2.3	Mitochondrial complementation may dampen the genotype-phenotype link but its extent is incompletely understood	13
1.2.3	Other sources of mitochondrial heterogeneity	15
1.3	Outline	15
2	MITOCHONDRIAL NETWORK FRAGMENTATION MODULATES MUTANT mtDNA ACCUMULATION INDEPENDENTLY OF ABSOLUTE FISSION-FUSION RATES	18
2.1	Summary	18
2.2	Introduction	19
2.3	Results	22
2.3.1	Stochastic modelling of the coupling between genetic and network dynamics of mtDNA populations	22
2.3.2	Mitochondrial network state rescales the linear increase of heteroplasmy variance over time independently of fission-fusion rate magnitudes	25
2.3.3	Derivation of heteroplasmy variance behaviour for a model of mitochondrial network & genetic dynamics	29

2.3.3.1	Conversion of a chemical reaction network into ordinary differential equations	30
2.3.3.2	Deriving an ODE description of the mitochondrial network system	33
2.3.3.3	Proof of heteroplasmy relation for linear feedback control	35
2.3.4	The influence of mitochondrial dynamics upon heteroplasmy variance under different models of genetic mtDNA control	42
2.3.5	Control strategies against mutant expansions	43
2.3.6	Optimal mitochondrial network configurations for mitochondrial quality control	47
2.4	Discussion	50
3	MITOCHONDRIAL DNA DENSITY HOMEOSTASIS ACCOUNTS FOR A THRESHOLD EFFECT IN A CYBRID MODEL OF A HUMAN MITOCHONDRIAL DISEASE	55
3.1	Summary	55
3.2	Introduction	56
3.3	Materials and methods	58
3.3.1	Data normalisation	58
3.3.2	Data transformation to per-cell dimensions	60
3.3.3	Error propagation	60
3.3.4	Growth rate determination	61
3.3.5	Model inference	61
3.4	Results	62
3.4.1	Per-cell interpretation of omics data highlights multiphasic dynamics in response to heteroplasmic load	62
3.4.2	Integrated omics data motivates a model of the causal relationships between bioenergetic variables	62
3.4.3	Interactions between bioenergetic variables can be cast as a bottom-up quantitative model	64
3.4.4	Parametrisations of a simple biophysical model account for complex observations across a range of heteroplasmic load	67
3.4.5	Key claims and predictions of biophysical model of heteroplasmy	69
3.5	Discussion	78
4	MITOCHONDRIAL HETEROGENEITY, METABOLIC SCALING AND CELL DEATH	81
4.1	Summary	81
4.2	Introduction	82

4.3	Linking mitochondrial heterogeneity and cell size	82
4.4	Cell physiological models coupled to cell death	85
4.4.1	A bioenergetic proposal: a combination of scaling of cellular power demand with cell size, and cell death, may account for observed patterns of mitochondrial functionality	87
4.4.2	A parabolic function of mitochondrial functionality against cell size is compatible with a mathematical model of mitochondrial variability . . .	91
4.5	Discussion	93
5	SOMATIC MOSAICISM DURING NEURODEVELOPMENT IN GENES ASSOCIATED WITH NEURODEGENERATION	94
5.1	Summary	94
5.2	Introduction	95
5.3	Methods	97
5.3.1	Order-of-magnitude calculation for the mutation rate during neurodevelopment	97
5.3.2	Approximate Bayesian computation for a crude branching model of neurodevelopment	100
5.3.3	Forward-simulation to estimate the prevalence of somatic mosaicism in neurodevelopment	106
5.4	Results	107
5.4.1	A crude model of neurodevelopment predicts the prevalence of somatic mosaicism in the human brain	107
5.4.2	Modelling of neurodevelopment as a binary tree is robust to several model perturbations	107
5.4.2.1	Branching model and cell death	107
5.4.2.2	Initial conditions for neurodevelopment	114
5.4.2.3	Asymmetric division in neurodevelopment	116
5.4.2.4	Intuition for model robustness	119
5.4.2.5	Cell migration and spatial correlations in neurodevelopment . .	121
5.5	Discussion	122
6	SUMMARY	124
	APPENDIX A SUPPLEMENTAL INFORMATION COMPLEMENTING CHAPTER 2	127
A.1	Constant rates yield unstable copy numbers for a model describing mtDNA genetic and network dynamics	127

A.2	Heteroplasmy variance relations for alternative model structures and modes of genetic mtDNA control	128
A.3	Choice of nominal parametrisation	131
A.4	Rate renormalisation	134
A.5	A modified Moran process may account for the alternative forms of heteroplasmy variance dynamics under different models of genetic mtDNA control	135
A.6	Error quantification	138
APPENDIX B SUPPLEMENTAL INFORMATION COMPLEMENTING CHAPTER 3		140
B.1	Transformation to per-cell dimensions using cell volume	140
B.2	Generative model description	143
B.3	Justifications for model structure	148
B.3.1	Justification of ETC mRNA and protein	148
B.3.2	Justification of cell volume scaling for power demand	150
B.3.3	Expected cell volume and growth rate	151
B.4	Experimental tests for model of MELAS heteroplasmy	153
B.4.1	Wild-type mtDNA density homeostasis is maintained until a minimum volume is reached at the critical heteroplasmy	153
B.4.2	Mitochondrial tRNAs are enriched in the vicinity of their corresponding parental mtDNA	153
B.4.3	Mutant mtDNAs have a transcriptional defect	154
B.4.4	Cell volume is not explained by cell cycle variations	154
B.4.5	Cells proliferate inversely with their size	154
B.4.6	Maximum respiratory capacity linearly tracks ETC protein content . . .	155
B.4.7	Summary	155
B.5	Alternative Models	155
B.5.1	Mutant mtDNA and transcription	155
B.5.2	tRNA misincorporation model	156
B.6	Relative OXPHOS contribution to power supply	161
APPENDIX C SUPPLEMENTAL INFORMATION COMPLEMENTING CHAPTER 4		163
C.1	Correlation between mitochondrial mass and cell volume is not strongly constrained	163
C.2	Simulating metabolic scaling and cell death	164
C.3	Coefficient of variation of relative mitochondrial membrane potential and cell size	166

C.4	A parabolic function of cell death against cell size as an alternative account of maximal mitochondrial functionality	168
C.4.1	Simulating parabolic function of cell death against cell size	170
APPENDIX D	SUPPLEMENTAL INFORMATION COMPLEMENTING CHAPTER 5	172
D.1	Summary statistics for the proportion of individuals with a pathologically mutated region of a fixed size	172
D.2	Details for the visualisation of mutated regions	173
D.3	Multimodality in the number of pathological neuronal mutations is driven by the largest pathological region	174
D.4	Constant region size model	174
D.4.1	Summary	174
D.4.2	Approximate Bayesian computation for fixed region size model	177
REFERENCES		206

1

INTRODUCTION

1.1 BASICS OF MITOCHONDRIAL PHYSIOLOGY

Mitochondria are double-membrane bound organelles of the cell, which are best known for their role in converting molecules ultimately derived from food into a more fungible energy currency, adenosine triphosphate (ATP), through a bioenergetic pathway called oxidative phosphorylation (OXPHOS). The turnover of ATP is high, around 40 kg per day for a resting human [1], reflecting the profound importance of ATP in a wide range of cellular and physiological functions. It is therefore, perhaps, unsurprising that mitochondrial dysfunction causes a range of devastating heritable diseases [2], as well as being associated with neurodegeneration [3, 4], cancer [5, 6], and healthy ageing [7, 8].

OXPHOS involves the transfer of high-energy electrons between enzymes of the electron transport chain (ETC), from Complex I to Complex IV, which use this energy to pump protons across the inner mitochondrial membrane. The resulting electrochemical gradient (or protonmotive force) is used by the molecular motor ATP synthase (also situated in the inner membrane) to convert adenosine diphosphate (ADP) and inorganic phosphate into ATP. Importantly, mitochondria possess their own small, circular genomes (mtDNA) which are held in multiple copies per cell. MtDNA is much shorter than nuclear DNA (16.6 kb per molecule, whereas nuclear DNA consists of 3.2 Mb per haploid genome, in humans), consisting of 13 protein-coding genes, 2 ribosomal RNAs and 22 tRNAs required for the translation of the mitochondrially-encoded

proteins [9].

Whilst nuclear DNA is replicated once per cell cycle in proliferating cells (i.e. cells which are actively dividing), mtDNA is replicated and degraded continuously, even in quiescent (i.e. non-dividing) cells. Replication errors during copying can give rise to *de novo* mutations, and subsequent copying of mutated mtDNAs can cause them to expand amongst the intra-cellular population of mtDNAs. Defining m as the number of mutated mtDNAs of a particular allele in a cell, and w as the number of non-mutated mtDNAs, heteroplasmy is defined as the fraction of mutated alleles per cell

$$h := \frac{m}{m + w}. \quad (1.1)$$

Since both m and w are stochastic processes (i.e. random variables which vary with time), due to the continuous replication and degradation of their corresponding mtDNA molecules, mathematical understanding may provide insights into the nature of the time-evolution of heteroplasmy, and therefore allow us to better understand the diseases which mtDNA mutations can induce.

Another important aspect of mitochondrial physiology is that mitochondria are not generally discrete, static, organelles (as they are often depicted). Rather, mitochondria continuously fuse and fragment resulting in a dynamic network. The purpose of mitochondrial networks is not fully understood, although it has been suggested that they serve to remove damaged organelles; increase metabolic output through non-linearities involved in sharing of intra-mitochondrial components; and dampen biochemical fluctuations [10].

1.2 MITOCHONDRIAL HETEROGENEITY

Since mitochondria form a dynamic intra-cellular population, as discussed above, there exists potential for features of mitochondrial physiology to be variable within and between single cells. We define mitochondrial heterogeneity as (1) the variation of a mitochondrial feature within the mitochondrial population of a single cell, and (2) the variation of a mitochondrial feature, potentially aggregated at the per-cell level, between supposedly identical cells. Mitochondrial heterogeneity is a theme which is present throughout this thesis. In this section, we present a review of the literature in the field of mitochondrial heterogeneity, and indicate how the work in subsequent chapters is placed within the broader literature.

Inter-cellular heterogeneity, in a more general sense, plays central functional roles in a variety of biomedically important phenomena, such as development [11, 12], virus infection [13],

chemotherapeutic resistance [14, 15] and gene expression in ageing [16]. Inter-cellular heterogeneity may arise from the intrinsically stochastic nature of cellular processes including gene transcription [17, 18], but also from ‘extrinsic’ sources such as the cell cycle [19] and partitioning noise [20–22].

Mitochondrial heterogeneity has been found to be a strong contributing factor to extrinsic cellular noise. Das Neves *et al.* [23] and Johnston *et al.* [24] found that mitochondrial mass, scaled by mitochondrial membrane potential, correlates strongly with global transcription rate at the single-cell level (also accounting for $\sim 50\%$ of the heterogeneity observed in protein levels [25]). Since global transcription rate has diverse implications for cellular function [26], this provides compelling evidence for the importance of mitochondrial heterogeneity as a contributor to cellular noise. Indeed, mathematical modelling predicted that mitochondrial functionality influences stem cell differentiation [24], which has received experimental support [27–29]. Furthermore, mitochondrial heterogeneity has been directly associated with cell-to-cell heterogeneity in chemotherapeutic resistance in mammalian cells [15, 30, 31], as well as proliferation rate and drug resistance in yeast [32]. The importance of metabolic noise is also becoming recognised in prokaryotic systems [33, 34].

Since individual cells consist of a population of mitochondria, intra-cellular heterogeneity may exist and may give rise to inter-cellular heterogeneity. Mitochondria also display a rich physiology, apart from genetic aspects. We therefore partition our discussion into genetic and non-genetic sources of mitochondrial heterogeneity and indicate their inter-dependence through genotype-phenotype links (Figure 1.1). This thesis is biomedically focussed, so we will emphasise findings in humans and model organisms used to study human disease such as mouse, fly, worm and yeast. We provide a necessarily non-exhaustive discussion of recent developments in these topics, and point out potential areas for future development.

1.2.1 GENETIC SOURCES OF MITOCHONDRIAL HETEROGENEITY

Genetic and non-genetic sources of mitochondrial heterogeneity are not independent. Variation in the amount and sequence of mtDNA affects the number and sequence of corresponding transcripts, eventually affecting respiratory output (although these relationships may be non-linear [35–37]). Conversely, the non-genetic state of the cell may affect the mitochondrial genetic state, for instance AMP/ATP ratios may alter mitochondrial biogenesis and autophagy (mitophagy) rates [38]. Variation in mtDNA turnover rates have the potential to affect not only mtDNA copy number but also heteroplasmy (see Eq. (1.1)) through random genetic drift

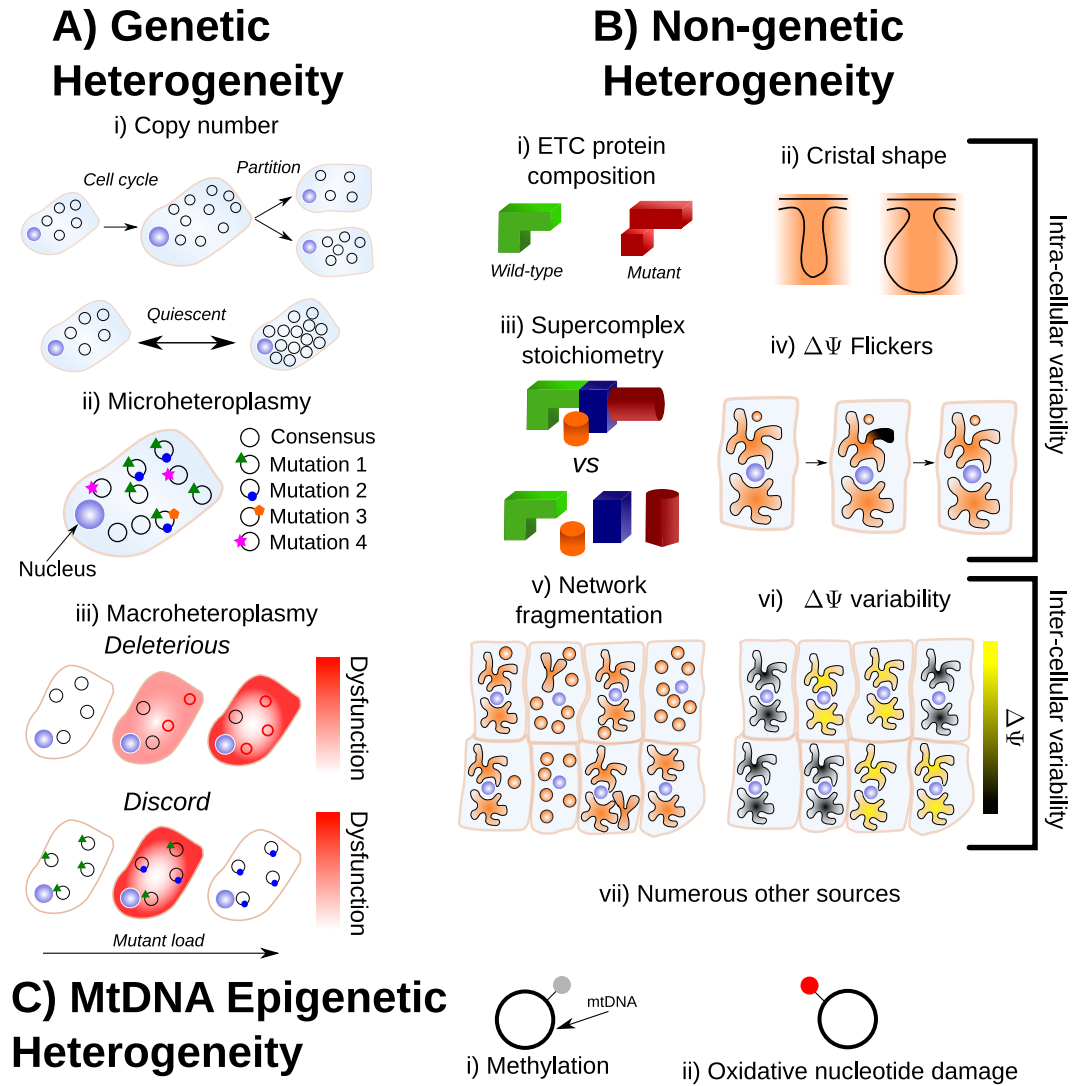


Figure 1.1: Sources of mitochondrial heterogeneity from genetic, non-genetic and epigenetic mitochondrial sources. (A) Potential sources of mitochondrial genetic heterogeneity. i) A proliferative cell which doubles in volume is expected to replicate its mtDNA complement by approximately $\times 2$ to avoid dilution, which may confound measurements of copy number heterogeneity; noise at partitioning is thought to be somewhat suppressed. The extent of copy number heterogeneity in quiescent cells, and its consequences, are not fully understood. ii) A cloud of mutations around the consensus sequence is expected given a finite mtDNA mutation rate; we term this ‘microheteroplasmy’. The actual mutant proportion for most mutations is expected to be very small in reality. iii) In the more canonical case, ‘macroheteroplasmy’, a deleterious sequence (such as a deleterious single nucleotide polymorphism or deletion, red circles) induces a pathological phenotype above a threshold heteroplasmy. By contrast, intermediate heteroplasmy has been observed to induce large fitness disadvantages but homoplasmy does not, indicating discord between the two alleles. (B) Potential sources of mitochondrial non-genetic heterogeneity; (i-iv) vary at the intra-cellular level, whereas (v-vi) vary between cells. i) MtDNA sequence variation can induce variation in the structure of corresponding proteins. ii) Cristal structure is variable and physiologically regulated to control respiratory output. iii) A landscape of supercomplex stoichiometries exists. iv) Spatially-restricted, transient, depolarisation/repolarisation cycles of the mitochondrial network are observed. v) The extent of mitochondrial fragmentation influences heteroplasmy dynamics (see Chapter 2). vi) Intercellular heterogeneity in mitochondrial membrane potential has been shown to be an important source of cellular noise. vii) There are numerous other sources of non-genetic heterogeneity: from distinct spatial niches to ER-mitochondrial interactions. (C) Potential sources of mitochondrial epigenetic heterogeneity. i) MtDNA is observed to undergo epigenetic modification by methylation, although the physiological significance of this is uncertain. ii) Oxidative damage to mtDNA nucleotides may cause transcriptional errors.

[39–43], as we discuss below.

1.2.1.1 INTER-CELLULAR MTDNA COPY NUMBER APPEARS TO BE UNDER HOMEOSTATIC CONTROL

MtDNA copy number appears to be under homeostatic control. A potential source of mitochondrial genotypic heterogeneity is simply the quantity of mtDNA per cell. A proliferating cell is expected to increase its complement of mtDNAs by a factor of 2 over the cell cycle, and noisily partition those molecules upon division [22] (Figure 1.1Ai). However, a number of studies suggest that cytoplasmic mtDNA density is tightly controlled in a variety of replicating cell types. Work in HeLa cells [44] and budding yeast [45] have indicated that mitochondrial volume increases approximately in proportion to cytoplasmic volume. Studies in proliferative human cell lines [46, 47], budding yeast [48] and fission yeast [49] have shown that the distribution of inter-nucleoid spacings is significantly perturbed from random, suggesting that mtDNA density is controlled in proliferating cells. Indeed, mathematical modelling using a constant mitochondrial density [24] was able to explain a range of single-cell data for replicating cells [23].

Interpretation of apparent mtDNA copy number homeostasis. The conservation of mitochondrial density is somewhat surprising, given that mitochondrial density is a potential axis for cells to control power production in response to differing demands, especially in the context of differing cell volume. Smaller cells have a larger surface area to volume ratio, therefore power demand is not expected to scale linearly with cell volume. In Chapter 4, in light of a recent experimental study [50], we discuss how cells may instead modulate their mitochondrial membrane potential, rather than their mtDNA density, to satisfy cellular demands in mammalian cells, perhaps affording the cell more control since membrane potential may change on a faster timescale than mtDNA biogenesis. The extent to which mtDNA density homeostasis holds in the absence of cell volume variation driven by the cell cycle, i.e. quiescent cells, has yet to be carefully explored (Figure 1.1Ai) despite its relevance for mosaic dysfunction in ageing post-mitotic tissues [7].

Pathological consequences of loss of mtDNA copy number homeostasis. A variety of nuclear mutations which induce defects in mtDNA maintenance cause mitochondrial depletion syndromes; these are severe disorders and clinically diverse in their physiological impact [51]. Conversely, it has been shown that increasing mtDNA copy number can rescue male infertility in mice engineered to accumulate mtDNA mutations, despite unaltered heteroplasmy [52, 53]. In Chapter 3, in light of a recent study [37], we discuss how failure to maintain homeostasis

in the density of functional mtDNAs may underlie the pathology of one of the most common mtDNA mutations associated with mitochondrial disease (3243A>G tRNA mutation). Indeed, assuming constant mitochondrial functionality, the study of Johnston *et al.* [24] predicts that a reduction in mtDNA density results in lowered ATP concentrations, which results in lowered transcription rate [23]. These studies highlight the potential pathophysiological relevance of maintaining mtDNA density homeostasis.

1.2.1.2 INTRA-CELLULAR MUTATIONS IN MITOCHONDRIAL DNA ARE A SOURCE OF GENOTYPIC HETEROGENEITY

MtDNA mutation as a source of heterogeneity. Mitochondrial DNA is replicated and degraded, even in non-proliferating tissues, which generates opportunities for mtDNA mutations to arise and proliferate. Studies of mtDNA mutation spectra in humans have suggested that point mutations predominantly arise from replication errors [54–56] as opposed to oxidative damage [57, 58], as is also the case for the “common” 4997 bp deletion [59].

Intra-cellular mtDNA mutation as a source of heterogeneity. Finite mutation rates during replication of mtDNA are expected to give rise to a set of closely-related sequences which do not all necessarily maximise fitness [60, 61]. Therefore, at the intra-cellular level, we expect to observe mtDNA sequence diversity (see e.g. [62]). Recent experimental work in primary cultures of mouse neurons and astrocytes has shown this to be the case [63]. The authors found 3.9 ± 5.7 single nucleotide variations (SNVs) ($\pm$ standard deviation) per mitochondrion [63], with a mitochondrion expected to contain around 5 molecules of mtDNA [64]. We use the term ‘microheteroplasmy’ to denote abundant ultra-low heteroplasmic mutations within single cells (Figure 1.1Aii).

Pathological consequences of microheteroplasmy. Whilst the pathological implications of pervasive microheteroplasmy remain unclear, comparisons could be drawn to experiments where a naturally occurring, but foreign, mtDNA haplotype is introduced into a cell with which it has not co-evolved. Increasing genetic distance amongst haplotypes has been shown to induce tissue-specific selective pressures in mice [65], and heteroplasmy between otherwise healthy haplotypes have been shown to induce fitness disadvantages [66, 67], as corroborated by bottom-up mathematical modelling [68]. Recent work has shown that such fitness disadvantages may be mediated through oxidative damage [69] (Figure 1.1Aiii).

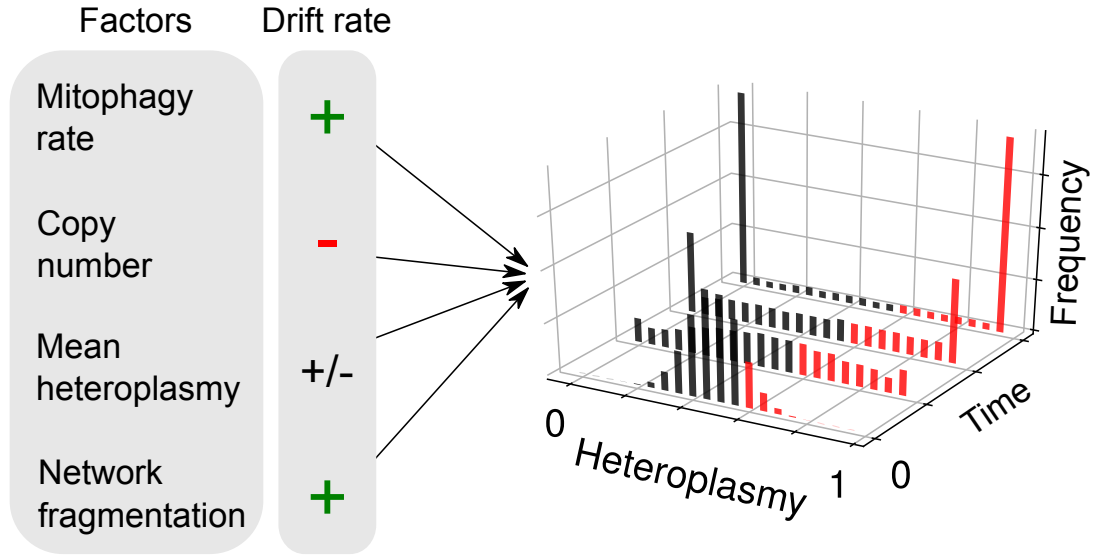


Figure 1.2: Factors influencing random genetic drift of mtDNA. Heteroplasmy (h , see Eq. (1.1)) is not generally constant between cells: it is a random variable and yields a distribution of cellular states. If the consensus and variant alleles experience the same instantaneous birth and death rates per cell, then the heteroplasmy distribution is subject to ‘neutral drift’. Neutral drift is characterised by the increase in variance of the heteroplasmy distribution with time. It is thought that when cells exceed a particular threshold heteroplasmy, a pathological phenotype may be expressed (red bars). Therefore, the number of pathological cells may increase with heteroplasmy variance. Mathematical modelling has shown that the rate of increase of heteroplasmy variance ($\nabla(h)$) increases with mitophagy rate, as higher turnover provides more opportunities for replication of either allele and cause a change in h . $\nabla(h)$ decreases with copy number, since large populations are more robust to fluctuations. $\nabla(h)$ is maximal when mean heteroplasmy is 0.5, and diminishes as one allele dominates over the other. In Chapter 2, we show that mitochondrial network fragmentation can rescale the genetic drift rate.

1.2.1.3 INTER-CELLULAR MUTATIONS IN MITOCHONDRIAL DNA ARE A SOURCE OF GENOTYPIC HETEROGENEITY

Cell-to-cell heterogeneity in heteroplasmy increases linearly with time. Another consequence of the finite mutation rate of mtDNA, and its stochastic turnover, is that mutations will occasionally drift to very large heteroplasmy simply by genetic drift [39, 70], i.e. ‘macroheteroplasmy’ (Figure 1.1Aiii). Cell-to-cell heterogeneity in heteroplasmy (heteroplasmy variance) therefore widens through time. Certain deleterious mutations may therefore drift to high levels of heteroplasmy and cause pathology [35] (see Figure 1.2). Theoretical work predicts that heteroplasmy variance increases approximately linearly, and increases with the rate of mitophagic turnover of mtDNAs [43, 71], but decreases with total mtDNA copy number with a non-linear dependence upon mean heteroplasmy [43]. A linear increase in heteroplasmy variance with time has recently been observed in mouse oocytes, and over pup lifetimes [72]. Mathematical modelling has suggested that increasing heteroplasmy variance might increase the energetic cost of maintaining a tissue [68].

Heteroplasmy variance and selective thresholds may counter the progressive increase in mu-

tant load during development. Increases in heteroplasmy variance can also purge mitochondrial mutations when combined with selection, as observed between generations in the human mitochondrial bottleneck [73–75]. During development, through a reduction of cellular mtDNA copy number, heteroplasmy variance is increased. When combined with a selective apoptotic threshold, whereby cells above a heteroplasmic threshold undergo cell death, mean heteroplasmy can be reduced. The mechanisms and timings of the mitochondrial bottleneck are debated [76–79]; however, stochastic modelling has shown that several proposals are compatible with the induction of heteroplasmy variance through a combination of random mtDNA partitioning at division and passive turnover of mtDNA [43, 73, 80].

Clonal expansions of mitochondrial mutations with age have pathological consequences. Mitochondrial mutations which survive the bottleneck can result in severe congenital diseases [2], and accelerate ageing in the case of severe mutant loads in mice [81]. Under physiological conditions, mtDNA mutations are also observed to accumulate with age across tissues in humans [7, 82–86]. The extent to which mtDNA mutations present at birth, versus somatic mutation post-development, contribute to healthy ageing remains open (although the vast amount of DNA replication during development itself is likely to seed and expand a number of potentially pathological mtDNA mutations, as explored in Chapter 5). Mitochondrial mutations which arise and focally expand post-development in skeletal muscle contribute to pathological age-related loss in muscle mass [87–89]. In the case of human muscle fibres of healthy individuals, the number of fibres exhibiting electron transport chain (ETC) abnormalities increases from approximately 6% at age 49 to 31% at age 92 [88]. Serial sections through laser-captured single skeletal muscle fibres showed that heteroplasmy in mtDNA deletions may exceed 90% [88]. Deletion mutations have been associated with local fibre atrophy and breakage in mice [90, 91] and rhesus monkeys [87] (although the causative role of mtDNA mutations in ageing for shorter-lived animals is contested [92, 93]). It is noteworthy that mitochondrial dysfunction in a particular tissue is able to induce stress responses in distal tissues through the mitochondrial unfolded protein response [94–96] and other hormonal signalling pathways [97, 98], suggesting potential organism-wide consequences of focal mitochondrial mutations. The development of novel, quantitative, dyes for electron transport chain (ETC) deficiency may allow more refined measurements of the causes of deletion proliferation in somatic tissues in the future [99, 100].

Inter-cellular mitochondrial exchange as a physiological means to slow heteroplasmy variance. Mitochondria have been observed to be transferred between cells [101, 102], as has the inter-cellular transfer of other organelles [103]. Several mechanisms by which mitochondrial transfer is achieved are known, including tunnelling nanotubes and microvesicles, amongst others [102].

Intercellular mitochondrial exchange has largely been understood through in vitro experiments [62, 101, 104]; however, recent developments have shown the effect to also be relevant in vivo e.g. [104–106]. For instance, Islam *et al.* [105] have shown that mouse mesenchymal stem cells (MSCs) are able to release microvesicles containing mitochondria, and protect against sepsis-induced acute lung injury in mice. It has also been suggested that astrocytes can donate mitochondria to neurons after stroke [107], although interpretations of these data exist which do not require mitochondrial exchange [108]. In terms of mitochondrial genetic dynamics, inter-cellular exchange of mtDNA appears to slow down heteroplasmic drift [62] and thereby reduce cell-to-cell mitochondrial genetic heterogeneity. It is possible that inter-cellular exchange of mitochondria, and mitochondrial DNA, is an evolved mechanism to ameliorate heteroplasmy variance and therefore the build-up of cells with pathological levels of heteroplasmy. Understanding the extent to which mitochondrial exchange buffers against a build-up in mitochondrial heterogeneity in vivo is a further avenue for research.

Gene therapy as a means to therapeutically control heteroplasmy distributions. Gene-editing technologies targeted at the mitochondrial genome are under development [109–112]. Such technologies sequence-specifically bind and cleave mitochondrial DNA, which is subsequently degraded rapidly [113]. Mathematical modelling of mitochondrially-targeted gene therapies predicts that tissues with high mean heteroplasmy and large heteroplasmy variance are generally more difficult to treat [68]; there is therefore a close link between these promising therapeutic technologies and inter-cellular heterogeneity in heteroplasmy.

1.2.2 NON-GENETIC SOURCES OF MITOCHONDRIAL HETEROGENEITY

Apart from the mitochondrial genome, there are many non-genetic properties of mitochondria which can vary. One such property, which is particularly clear, is ETC protein structure which is influenced by sequence heterogeneity in mtDNA (Figure 1.1Bi) – this being a potentially important source of both intra- and inter-cellular variability. Other non-genetic differences in mitochondria can include mitochondrial membrane composition/structure, ion content, membrane potential and network structure. Intra-cellular spatial heterogeneity through the existence of sub-cellular mitochondrial niches [114–116], e.g. perinuclear vs. peripheral locations, has been highlighted as an important axis of mitochondrial heterogeneity by previous authors, see [117]. Below, we discuss several other aspects in which non-genetic attributes of mitochondria may vary, and the potential pathological consequences of such variation. We also draw attention to how non-genetic heterogeneity may be driven by, or drive, genetic heterogeneity

through genotype-phenotype links.

1.2.2.1 STOICHIOMETRIC AND STRUCTURAL HETEROGENEITY IN THE INNER MITOCHONDRIAL MEMBRANE AS A POTENTIAL AMPLIFIER OF THE GENOTYPE-PHENOTYPE LINK

MtDNA genotype as a driver of IMM phenotype. The inner mitochondrial membrane (IMM) is heterogeneous in both its composition and its topology. The connection between the inner mitochondrial membrane and the mitochondrial genotype is particularly relevant since mitochondrial DNA is situated in the mitochondrial matrix and in close proximity to the IMM [118]. Consequently, one might expect mtDNAs to affect their local respiratory units more than distal mtDNAs [119], thus allowing greater control of respiration at a local level [120–122]. Below we discuss aspects of how IMM physiology may generate mitochondrial heterogeneity, and how cardiolipin may act as a sensor of the mitochondrial genotype. We discuss how cardiolipin reacts to differences in reactive oxygen species (ROS) production and pH, and may amplify the mitochondrial phenotype through supercomplex formation and cristal structure.

Mitochondrial supercomplex stoichiometry as an axis of IMM heterogeneity. The stoichiometry of respiratory units within the IMM can vary, since respiratory units organise into supramolecular structures termed supercomplexes [123, 124] (Figure 1.1Biii). A variety of supercomplex stoichiometries exist in mammalian cells [123, 125] as well as free respiratory subunits [123]. The ‘plasticity model’ has therefore been suggested [126], whereby a landscape of combinations between respiratory complexes coexist. Dysfunction in the assembly of mitochondrial supercomplexes results in decreased muscle activity and heat production in the cold due to reduced CIV activity [127], showing that supercomplex assembly is required for fully-functional respiration. Consequently, dysfunction in supercomplex assembly could affect mitochondrial quality control pathways and influence the mitochondrial genotype.

Cristal structure as an axis of IMM heterogeneity. The topology of mitochondrial membranes are heterogeneous [124, 128]. Notably, ATP-synthase forms dimers in the inner mitochondrial membrane (IMM) which often arrange into micron-scale rows, which are associated with high local curvature to form mitochondrial cristae [129, 130]. Recent work has suggested that the mitochondrial fusion protein OPA1 stabilises ATP synthase oligomers by modulating cristal shape [131]. Crista membranes show an enrichment of respiratory complexes relative to the inner boundary membrane [132, 133], and it has been suggested that cristae exist to increase the packing density of respiratory units [134]. Individual cristae are morphologically heterogeneous

in mammalian cells [135], and can be modulated in response to altered metabolic demands [136]: cristae become narrower in mammalian cells in response to starvation [137], suggesting that cristal shape influences respiratory efficiency, perhaps by modulating local substrate concentrations [138] (Figure 1.1Bii). Remodelling of mitochondrial morphology occurs during cell death [139, 140], whereby supercomplexes and dimers of ATP synthase disassemble and cristal structure becomes disorganised [141], allowing the release of cytochrome c (an electron carrier of the ETC) to trigger the intrinsic cell death pathway [142] (although the importance of crista remodelling for cytochrome c release has been questioned [143]). The intimate connection between mitochondrial physiology and cell death provides insight into the recent observation that mitochondrial heterogeneity can partially explain variability in chemotherapeutic resistance [15].

Cardiolipin is necessary for supercomplex and crista stabilisation. Cardiolipin is a phospholipid found in the inner mitochondrial membrane, and stabilises both supercomplexes and cristal structure. When exposed to a pH gradient, cardiolipin-containing lipid vesicles spontaneously form crista-like membrane invaginations [144, 145], thus providing a connection between respiratory activity and cristal shape. Furthermore, flies with deficient cardiolipin levels show reduced ATP synthase abundance in high-curvature regions of cristae, resulting in disorganised cristae, cardiac insufficiency, motor weakness and early death [146]. In yeast, cardiolipin has been shown to be necessary for supercomplex stabilisation [147, 148]. Indeed, patients with Barth syndrome who are deficient in cardiolipin due to a mutation in the tafazzin gene show both aberrant crista formation [146] and reduced supercomplex formation [149].

Cardiolipin as a potential amplifier of mtDNA heterogeneity through alterations in pH and ROS generation. Cardiolipin is particularly susceptible to damage by ROS. In isolated bovine mitochondria, ROS exposure resulted in loss of CI activity, but exogenously added cardiolipin could restore CI activity [150]. It has recently been shown that a mitochondrially-targeted antioxidant (MitoQ) is able to increase cardiolipin expression and content in liver mitochondria of rats fed on a high-fat diet, resulting in increased mitochondrial functionality and ATP synthase activity [151]. Together, this suggests that mitochondrial dysfunction, which alters pH gradients across the IMM and ROS production, can cause cristae to become disorganised and affect supercomplex assembly via cardiolipin, potentially resulting in further loss of mitochondrial efficiency and ROS production. It is possible that heterogeneity in pH and ROS production, for instance through mtDNA mutation heterogeneity [122], could be amplified through such mechanisms and thus strengthen the genotype-phenotype link between mtDNA and their local respiratory complexes.

1.2.2.2 MITOCHONDRIAL MEMBRANE POTENTIAL HETEROGENEITY, MITOCHONDRIAL NETWORKS AND MITOCHONDRIAL GENOTYPE

Inter-cellular mitochondrial membrane potential heterogeneity as a predictor of cell-physiological heterogeneity. The inner membrane potential ($\Delta\Psi$) is an indicator of mitochondrial functionality, generated by the ETC, which drives the synthesis of ATP by ATP-synthase. Mitochondrial output is highly sensitive to $\Delta\Psi$: a 14 mV change in $\Delta\Psi$ corresponds to a 10-fold change in the maximum ATP/ADP ratio [152], where $\Delta\Psi$ typically ranges between 150-180 mV [153]. As discussed above, mitochondrial mass scaled by $\Delta\Psi$ has been shown to explain much of the variation in transcript elongation rate [23, 24] and protein noise [25], as well as predicting phenomena such as cell cycle duration [24] and chemotherapeutic resistance [15] (Figure 1.1Bvi).

Calcium and pH transients mediated by the mPTP as determinants of intra-cellular $\Delta\Psi$ fluctuations. In addition to inter-cellular $\Delta\Psi$ heterogeneity, individual mitochondria have been shown to undergo transient depolarisation/repolarisation cycles, termed ‘flickers’ in mammals [154, 155] or ‘pulses’ in plants [156] (Figure 1.1Biv). In mammalian cells, flickers range from <10 mV to >100 mV [155] and appear to be regulated by the mitochondrial permeability transition pore (mPTP), although other mechanisms have been identified [117]. The mPTP is a non-selective mitochondrial channel that induces cell death when open for prolonged periods [157]. Short openings of the mPTP induces efflux of Ca^{2+} ions; Ca^{2+} being a stimulant of the tricarboxylic acid cycle and modulates ATP production [158]. Transient Ca^{2+} efflux induces transient mitochondrial membrane depolarisation [159]. Transient opening of the mPTP has also been correlated with circularly permuted yellow fluorescent protein (cpYFP) fluorescence [160], which is associated with transient increases in pH (i.e. loss of protons) in the mitochondrial matrix [161, 162]. Such events were correlated with transient mitochondrial membrane depolarisation [160]. Since mPTP activity is, by definition, unselective (‘*permeability* transition’), the mPTP ought to dissipate any gradient; yet, alkalinisation of the mitochondrial matrix is observed to correlate with mPTP opening [160–162], indicating that the inner membrane must be sealed for protons during the event. Hence the correlation between the opening of the mitochondrial pore and transient increases in matrix pH may potentially be mediated by a selective mode of the same pore. Flickers have been shown to exist in mouse astrocytes and are enhanced by both neuronal activity and ROS production [163]. Hence, transient depolarisation cycles mediated by mPTP appear to be a means of regulating metabolic rate, and perhaps ROS production, at an intra-cellular scale in vivo.

Mitochondrial membrane potential influences mitochondrial genotype through quality control.

The canonical means by which mitochondrial membrane potential feeds back into the genetic state is through mitophagy and mitochondrial network dynamics [164]. Mitochondria are not static organelles but undergo dynamic fusion and fission, the purpose of which is incompletely understood [10]. In rat pancreas cells, it has been shown that fission often results in a daughter mitochondrion which has a lower $\Delta\Psi$ than its sister [164]. These depolarised mitochondria have a lowered propensity for fusion, and are more likely to be degraded [164]. Selective fusion, when combined with non-selective mitophagy, is sufficient to preferentially degrade depolarised/damaged mitochondria [10] (see also Chapter 2), although the selective strength of mitophagy itself is not yet fully understood. The extent to which mitochondria which are degraded via selective/non-selective forms of mitophagy possess mitochondrial genomes which are perturbed from the consensus sequence also remains incompletely understood [165]. However, if depolarised mitochondria are more likely to be perturbed from the consensus mitochondrial sequence, mitochondrial quality control mechanisms would exert a negative selective pressure against variant alleles. Indeed, negative selection pressures have been observed in human tissues [166] and at the intracellular level in mice [63], which may be due to a combination of mitochondrial networks and mitophagy. In Chapter 2 we explore the relationship between mitochondrial networks and genetics in greater depth.

1.2.2.3 MITOCHONDRIAL COMPLEMENTATION MAY DAMPEN THE GENOTYPE-PHENOTYPE LINK BUT ITS EXTENT IS INCOMPLETELY UNDERSTOOD

Mitochondrial complementation is thought to partially buffer genotype-phenotype links via the threshold effect. Mitochondrial ‘complementation’ consists of mitochondria sharing their contents through fusion-fission events, and potentially complementing each others’ genetic defects [167–170]. This is supported by experiments involving photoactivatable fluorescent proteins which show that intra-mitochondrial contents mix over time [164, 171]. Inhibition of fusion and fission have also been shown to induce heterogeneity in the distribution of aged mitochondrial proteins [172]. Beyond depending on fission and fusion rates, and the degree of mitochondrial mobility, the strength of complementation is closely linked to the diffusivity of the mitochondrial matrix: if the matrix is a high-diffusivity environment then gene products are expected to be promiscuous within the matrix and not remain local to their parental mtDNA. Therefore, in cells heteroplasmic between wild-type and a pathological mutant, healthy versions of any particular transcript would be found in the matrix. The idea of complementation is also closely linked to the mitochondrial threshold effect [35, 173], whereby complementation effects allow cells to withstand large levels of mutant load (60-90%, [35, 174, 175]) before displaying a

respiratory defect.

Mitochondrial complementation remains incompletely understood. Much of the progress in understanding mitochondrial complementation derives from cell fusion studies where cells harbouring different mtDNA mutations are fused, and mitochondrial functionality recovers through sharing of transcription products [170, 176, 177]. Such experiments highlighted physiological subtleties which remain incompletely understood. For instance, the experiment of Ono *et al.* [170] between cells harbouring two different mt-tRNA mutations required an adaptive period of 10-14 days before respiratory activity was restored – the reason for this was not fully understood. Furthermore, cell fusion studies between cells harbouring partially-functional (3271 T>C) and completely dysfunctional (3243 A>G) mitochondrial tRNA-Leu mutations resulted in a sigmoidal relationship between heteroplasmy and COX-activity, whereas a linear relationship was observed for simple co-culture between the two cell types [178], which was not fully understood. In contrast to the experiment of Ono *et al.* [170], an experiment by Gilkerson *et al.* [176] between two non-overlapping mitochondrial deletions showed recovery of MTCO2 protein production on a relatively fast timescale (4 days). Furthermore, Gilkerson *et al.* [176] found that fused cells cycled between heteroplasmic and homoplasmic states during long-term culture, despite being on a medium which mildly selected for mitochondrial function (but still allowed ATP production through glycolysis). Yet, a straightforward interpretation of complementation would suggest that heteroplasmic states would show maximal fitness. This effect was originally explained as a tension between a neutral genetic drift effect, deriving from the assumption that nucleoids possess several mtDNAs [179], and selective pressure for heteroplasmy [176]. However super-resolution microscopy studies have revealed that nucleoids harbour only 1-2 mtDNAs per nucleoid [180, 181]. It is possible that large drift rates could be explained by the passive partitioning of very large clusters of genetically homogeneous mtDNA at cell division. Furthermore, active asymmetric apportioning of damaged or aged mitochondria has been observed in yeast [182] and stem cells [183] (which has been shown to influence cellular fate decisions in immune cells [184]); therefore, active mechanisms of asymmetric apportioning of mitochondrial genomes may also exist. Cycles between heteroplasmy and homoplasmy highlight an area for future experimental and theoretical investigation.

Restricted diffusion in the IMM may inhibit complementation. Experiments by Wilkens *et al.* [185] fused HeLa cells with mitochondrially-targeted fluorescent proteins of differing colours to investigate the extent of mitochondrial diffusivity. The authors found that, through cycles of mitochondrial fusion and fission, IMM proteins appear to experience slow diffusion relative to the outer mitochondrial membrane, and retain cristalline structure [171]. Consequently,

complementation may be diminished since gene products may remain local to their parental mtDNA [119], allowing local control of respiration [120–122]. This is supported by the observation that mitochondrial transcripts are particularly concentrated around mtDNA [186]. The idea of a local enrichment of tRNAs around the parental mtDNA is used in Chapter 3 and is compatible with a diverse set of omics data [37]. Hence, restricted complementation and local phenotype-genotype links appear to be explanatory.

1.2.3 OTHER SOURCES OF MITOCHONDRIAL HETEROGENEITY

We note that, apart from genetic and non-genetic heterogeneity, mitochondrial epigenetic modification is also a potential layer of heterogeneity (Figure 1.1C). Mitochondrial DNA can undergo methylation [187]; although its physiological impact is still being unravelled, it has been suggested that methylation may regulate mtDNA gene expression [188]. Furthermore, whilst oxidative damage to mtDNA nucleotides have classically been known for their potential role in mutagenesis, it has been proposed that the formation of 8-OHdG, which is a ROS-modified version of guanine, induces mitochondrial mutations at the transcriptional level [189] and may be responsible for premature ageing phenotypes in mice bred to rapidly accumulate mtDNA mutations [52, 190, 191]. Further downstream sources of mitochondrial heterogeneity, such as translation errors, have also been shown to have physiological consequences in budding yeast [192]. Therefore, there exists a variety of potential mechanisms downstream from mtDNA mutation, which are nevertheless constrained by mtDNA, which may be able to explain the age-related functional decline of mitochondria in shorter-lived animals [193, 194] where direct mtDNA mutations might not [92, 93].

1.3 OUTLINE

As discussed in Section 1.2, multiple aspects of mitochondrial physiology display heterogeneity. To understand the significance of mitochondrial heterogeneity, we may use mathematical tools such as stochastic modelling, as well as inference to connect with data, allowing us to constrain our models and improve our understanding. This thesis focusses on the role of mitochondria in human health and disease through the application of stochastic modelling and inference. Our discussion of mitochondrial heterogeneity in Section 1.2 has appeared as a pre-print on ArXiv and has been submitted for publication [195].

In Chapter 2, we provide the first analytic treatment to bridge mitochondrial genetics with

network dynamics. We show for a wide range of neutral models of mtDNA genetics that the rate of increase of heteroplasmy variance, and the rate of accumulation of *de novo* mutations, is proportional to the fraction of unfused mitochondria, independently of absolute fission-fusion rates. We then relax the assumption of neutral genetic dynamics to explore the role of mitochondrial networks in the context of genetic selection. We find that intermediate fusion/fission ratios are optimal for the clearance of mtDNA mutants. We use our findings to suggest that promoting fusion when mean heteroplasmy is low may serve to slow random genetic drift, which could be potentially therapeutic for mitochondrial disease and healthy ageing. This chapter has appeared as a pre-print on BioRxiv and has been submitted for publication [196].

In Chapter 3, we analyse a recent dataset which consisted of transcriptomic and metabolic data for cells containing different levels of heteroplasmy in the 3243A>G tRNA-Leu mtDNA mutation, which can cause mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes (MELAS) syndrome [37]. We construct a mathematical model linking ETC and glycolysis transcript levels, ETC protein, cell volume, maximum respiratory capacity and cellular proliferation rate. Our model suggests that for progressively increasing heteroplasmy, cells attempt to maintain homeostasis in the cytoplasmic density of wild-type mtDNAs through cell volume reduction, and therefore power demand reduction. When heteroplasmy passes a critical threshold, a minimum cell volume is reached whereby cells toggle from power demand reduction to power supply increase. Our model provides further support for the importance of mtDNA density in cell physiology. We propose novel experiments, as well as therapeutic strategies, based on our results. This chapter has appeared as a publication in Biochemical Journal [197].

In Chapter 4, we analyse another recent dataset which investigated the relationship between cellular volume and mitochondrial functionality [50], finding a concave relationship between cellular volume and mitochondrial functionality. We find that a simple mathematical model of cell size scaling of cellular power demand may explain the observed reduction in mitochondrial functionality with increasing cell size, in large cells. When combined with a simple model of cell death, whereby smaller cells are at an increased risk of cell death, our model may account for many of the observations in Ref. [50]. Our work emphasises the need to account for confounding variables such as cell size, cell death, cell cycle stage and mitochondrial density in future single-cell studies of mitochondrial physiology. This chapter has appeared as a publication in BioEssays [198].

Finally, in Chapter 5, in contrast to the previous chapters which focus on the mitochondrial genome, we provide a set of initial estimates for the propensity of mosaic somatic mutations

(or simply ‘somatic mosaicism’) in the nuclear genome induced during neurodevelopment. Our study focusses on estimating the propensity of somatic mosaicism for a set of genes associated with neurodegenerative disease. We model neurodevelopment, crudely, as a deterministic branching process, and connect with a dataset which sequenced samples of brain tissue from post-mortem control subjects, and subjects with neurodegenerative disease. We perturb our model to include simple models of cell death, founder effects and asymmetric division, finding that our extrapolations are surprisingly robust given the simplicity of our modelling approach. This robustness may be explained by a simple order-of-magnitude calculation. We suggest that if neurodevelopment consists of unbiased, spatially proximal, replication of cells once the brain consists of $\gtrsim 10^6$ cells, then adult individuals are expected to possess 1 spatially contiguous, pathologically mutated, region of 10^4 – 10^5 neurons. This opens the possibility that mutated regions of neurons seeded during development could contribute to neurodegenerative disease risk. Our approach could be extended in the future to various genes and tissues of interest, including mitochondrial genes. Somatic mosaicism in nuclear genes linked with mtDNA maintenance (e.g. *POLG*, *PINK1* and *PRKN*) could seed regions of mutated mtDNA, possibly inducing localised ageing phenotypes within tissues (see Section 1.2.1.3). This chapter has been accepted for publication in Nature Communications [199].

Approximately two months of this thesis was dedicated to laboratory work. In collaboration with Dr. Tristan Rodriguez, we performed preliminary investigations of genotype-phenotype links for heteroplasmic embryonic stem cell lines. This work involved cell culture, FACS analysis, and a number of other wet-lab techniques. In connection with this work, I co-supervised three Masters students (Daria Dicu, Madeleine Hall and Heather Jackson), working on inference techniques for single-cell mtDNA quantification using qPCR and confocal microscopy data. Furthermore, in collaboration with Prof. Gyorgy Szabadkai, we investigated single-cell mtDNA quantification using high-throughput confocal microscopy. I contributed to statistical modelling and analysis in a further two publications during the course of this PhD [200, 201], which will not be discussed here.

2

MITOCHONDRIAL NETWORK FRAGMENTATION MODULATES MUTANT MTDNA ACCUMULATION INDEPENDENTLY OF ABSOLUTE FISSION-FUSION RATES

2.1 SUMMARY

Mitochondrial DNA (mtDNA) mutations cause severe congenital diseases but may also be associated with healthy ageing. MtDNA is stochastically replicated and degraded, and exists within organelles which undergo dynamic fusion and fission. The role of the resulting mitochondrial networks in determining the time evolution of the cellular proportion of mutated mtDNA molecules (heteroplasmy), and cell-to-cell variability in heteroplasmy (heteroplasmy variance), remains incompletely understood. Heteroplasmy variance is particularly important since it modulates the number of pathological cells in a tissue. Here, we provide the first wide-reaching mathematical treatment which bridges mitochondrial network and genetic states. We show that, for a range of models, the rate of increase in heteroplasmy variance, and the rate of *de novo* mutation, is proportionately modulated by the fraction of unfused mitochondria, independently of the absolute fission-fusion rate. In the context of selective fusion, we show

that intermediate fusion/fission ratios are optimal for the clearance of mtDNA mutants. Our findings imply that modulating network state, mitophagy rate and copy number to slow down heteroplasmy dynamics when mean heteroplasmy is low, could have therapeutic advantages for mitochondrial disease and healthy ageing.

Code for simulations and analysis in this chapter can be accessed at <https://github.com/ImperialCollegeLondon/MitoNetworksGenetics>

2.2 INTRODUCTION

As discussed in Chapter 1, an important observation in mitochondrial physiology is the threshold effect, which we will discuss in greater detail for a particular pathological mtDNA mutation in Chapter 3. It is thought that cells can often tolerate relatively high levels of mtDNA mutation, until heteroplasmy exceeds a certain critical value where a pathological phenotype occurs [35, 37, 173]. Fluctuations within individual cells mean that heteroplasmy is not constant within a tissue (Figure 2.1A), but follows a probability distribution which changes with time (Figure 2.1B). In the following, motivated by a general picture of ageing, we will largely focus on the setting of non-dividing cells, which possess two mtDNA variants (we will consider *de novo* mutation in Section 2.3.5). The variance of the distribution of heteroplasmy informs upon the fraction of cells above a given pathological threshold (Figure 2.1B). Therefore heteroplasmy variance is related to the number of dysfunctional cells within a tissue, and is a quantity of importance in addition to mean heteroplasmy.

Mitochondria exist within a network which dynamically fuses and fragments. Although the function of mitochondrial networks remains an open question [10], it is often thought that a combination of network dynamics and mitochondrial autophagy (termed mitophagy) perform quality control on the mitochondrial population [164]. However, the observation of pervasive intra-mitochondrial mtDNA mutation [63] and universal heteroplasmy in the human population [202] suggest that mitochondrial quality control mechanisms at the genetic level are weak, at least for some mitochondrial mutations. Neutral genetic models are therefore a rational means to understand mtDNA dynamics, whereby each species has the same probability per unit time of degradation, and neither species has a replicative advantage [40, 43, 71].

A number of studies have attempted to understand the impact of the mitochondrial network on mitochondrial dysfunction through computer simulation (reviewed in [203]). Mouli *et al.* [204] use a stochastic model to investigate the role of the mitochondrial network in quality control,

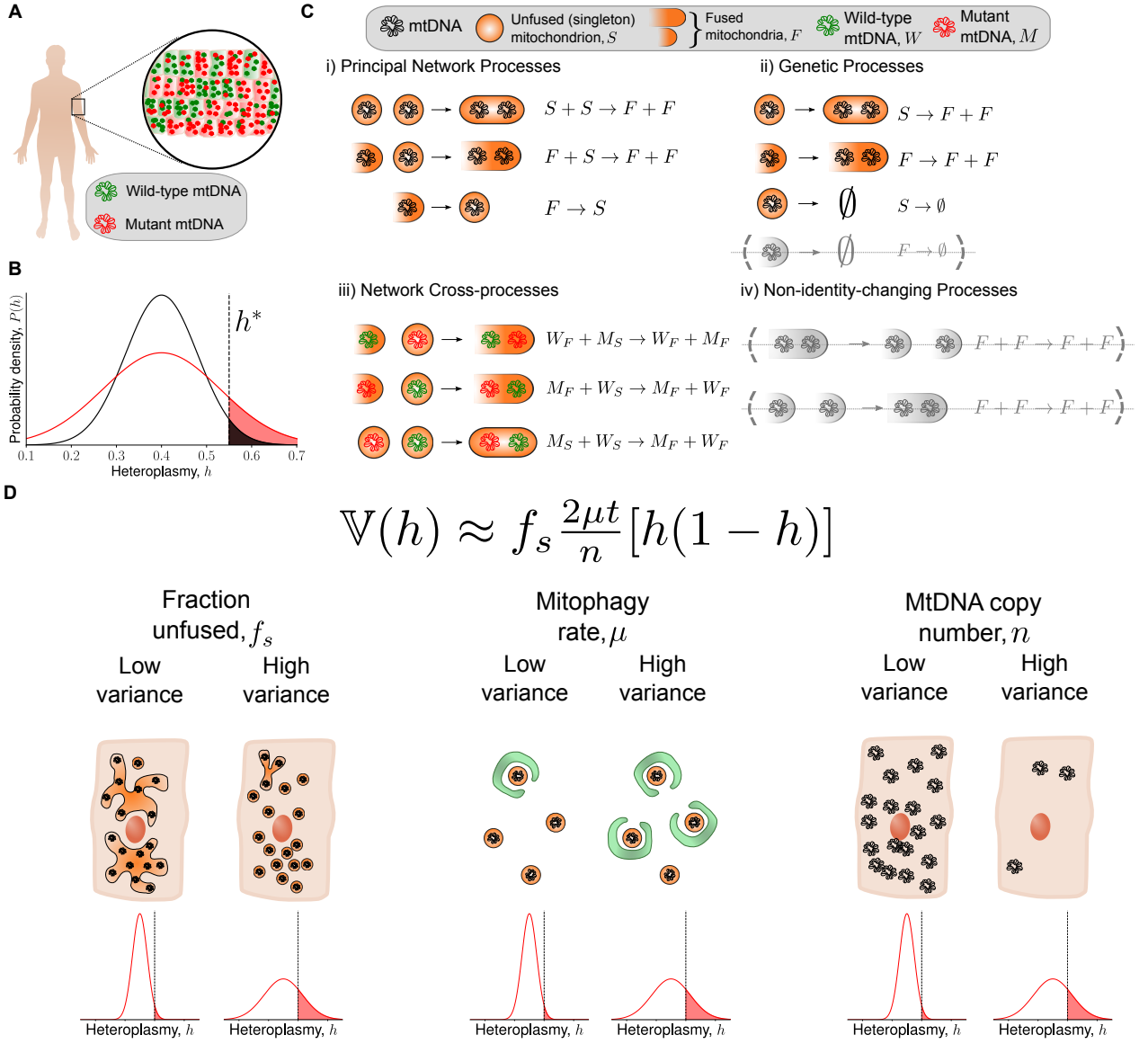


Figure 2.1: A simple model bridging mitochondrial networks and genetics yields a wide-reaching, analytically obtained, description of heteroplasmy variance dynamics. (A) A population of cells from a tissue exhibit inter-cellular heterogeneity in mitochondrial content: both mutant load (heteroplasmy) and copy number. (B) Inter-cellular heterogeneity implies that heteroplasmy is described by a probability distribution. Cells above a threshold heteroplasmy (h^* , black dashed line) are thought to exhibit a pathological phenotype. The low-variance distribution (black line) has fewer cells above a pathological threshold heteroplasmy than the high-variance distribution (red line). (C) The chemical reaction network we use to model the dynamics of mitochondrial DNA (see Main Text for a detailed description). MtDNAs are assigned a genetic state: mutant (M) or wild-type (W), and a network state: singleton (i.e. unfused, S) or fused (F). (D) The central result of our work is, assuming that a cell at time $t = 0$ is at its (deterministic) steady-state, heteroplasmy variance ($\mathbb{V}(h)$) approximately increases with time (t), mitophagy rate (μ) and the fraction of mitochondria that are unfused (f_s), and decreases with mtDNA copy number (n). Importantly, $\mathbb{V}(h)$ does not depend on the absolute magnitude of the fission-fusion rates.

where mitochondria incur “damage” at a constant rate, which was linked sigmoidally to a qualitative mitochondrial “activity”. The authors investigated selective fusion (where fusion occurs between sufficiently active mitochondria), and assumed selective mitophagy (where only mitochondria with sufficiently low activity are degraded). The authors concluded that i) a higher fusion rate means more effective clearance of damaged mitochondria when fusion is selective; ii) intermediate fusion rates are optimal for selective mitophagy. Patel *et al.* [205] take a similar approach, simulating autophagy, replication, fusion, fission and transport of mitochondria, finding that when network dynamics are coupled to transport, that transport can indirectly modulate mitochondrial health through mitochondrial dynamics. In another pair of studies by Tam *et al.* [206, 207] the authors applied a chemical reaction model where mutant and wild-type mtDNAs undergoing selective fusion, fission, mitophagy and replication, where the cell is partitioned into spatial compartments. The authors argue that while more frequent fission-fusion events increase the occurrence of mutant-rich mitochondria, which may be removed by selective mitophagy, these mitochondria exist for a shorter period of time, hence the existence of a trade-off [207]. The authors also argue that slower fission-fusion rates result in slow spatial mixing of mtDNA, implying greater stochasticity, and a greater propensity for the expansion of mutant mtDNAs [206]. Finally, a study by Figge *et al.* [208] which models transitions within an abstract “quality” space, argues that if mitochondrial fission induces damage, decelerating fission-fusion cycles may improve mitochondrial quality; note that mitochondrial fission is usually interpreted as exposing damaged mitochondria rather than itself being harmful [164, 203].

Previous attempts to link mitochondrial genetics and network dynamics, while important for breaking ground, have centred around complex computer simulations, making it difficult to deduce general laws and principles. In contrast, we take a simpler approach in terms of our model structure (Figure 2.1C), allowing us to derive explicit, interpretable, mathematical formulae which provide intuitive understanding, and give a direct account for the phenomena which are observed in our model (Figure 2.1D). Our results hold for a range of variant model structures. Simplified approaches using stochastic modelling have shown success in understanding mitochondrial physiology from a purely genetic perspective [40, 41, 43]. Furthermore, there currently exists limited evidence for pronounced, universal, selective differences of mitochondrial variants *in vivo* [56, 209]. Our basic approach therefore also differs from previous modelling attempts, since our model is neutral with respect to genetics (no replicative advantage or selective mitophagy) and the mitochondrial network (no selective fusion). Evidence for negative selection of particular mtDNA mutations has been observed *in vivo* [63, 166]; we therefore

extend our analysis to explore selectivity in the context of mitochondrial quality control using our simplified framework.

Here, we reveal the first general mathematical principle linking network state and heteroplasmy statistics (Figure 2.1D). This link shows analytically, for a broad range of situations, that the expansion of mtDNA mutants is strongly modulated by network state, such that the rate of increase of heteroplasmy variance, and the rate of accumulation of *de novo* mutation, is proportional to the fraction of unfused mitochondria. This result stems from the notion that fusion shields mtDNAs from turnover, since autophagy of large fragments of the mitochondrial network are unlikely, which effectively rescales time. Importantly, we show that heteroplasmy variance is independent of the absolute magnitude of the fusion and fission rates due to a separation of timescales between genetic and network processes (in contrast to [207]). Furthermore, we find that when fusion is selective, intermediate fusion/fission ratios are optimal for the clearance of mutated mtDNAs (in contrast to [204]). When mitophagy is selective, complete fragmentation of the network results in the most effective elimination of mitochondrial mutants (in contrast to [204]). We also confirm that mitophagy and mitochondrial DNA copy number also affect the rate of accumulation of *de novo* mutations [43]. We suggest that pharmacological interventions which promote fusion, slow mitophagy and increase copy number during youth may slow the rate of accumulation of pathologically mutated cells, with implications for mitochondrial disease and ageing.

2.3 RESULTS

2.3.1 STOCHASTIC MODELLING OF THE COUPLING BETWEEN GENETIC AND NETWORK DYNAMICS OF MTDNA POPULATIONS

Our modelling approach takes a chemical master equation perspective by combining a general model of neutral genetic drift (for instance, see [40, 43]) with a model of mitochondrial network dynamics. We consider the existence of two mitochondrial alleles, wild-type (W) and mutant (M). MtDNAs exist within mitochondria, which undergo fusion and fission. We therefore assign mtDNAs a network state: fused (F) or unfused (we term “singleton”, S). This representation of the mitochondrial network allows us to include the effects of the mitochondrial network in a simple way, without the need to resort to a spatial model or consider the precise network structure, allowing us to make analytic progress and derive interpretable formulae in a more general range of situations.

Our model can be decomposed into three notional blocks (Figure 2.1C). Firstly, the principal network processes denote fusion and fission of mitochondria containing mtDNAs of the same allele



where X denotes either a wild-type (W) or a mutant (M) mtDNA (therefore a set of chemical reactions analogous to Eq.(2.1)-Eq.(2.3) exist for both DNA species). γ and β are the stochastic rate constants for fusion and fission respectively. Note that we constrain all fusion reactions to have the same rate constant, since this is the simplest model of fusion.

Secondly, mtDNAs are replicated and degraded through a set of reactions termed genetic processes. A central assumption is that all degradation of mtDNAs occur through mitophagy, and that only small pieces of the mitochondrial network are susceptible to mitophagy; for parsimony we take the limit of only the singletons being susceptible to mitophagy



where λ and μ are the replication and mitophagy rates respectively, which are shared by both W and M resulting in a so-called ‘neutral’ genetic model. $\emptyset$ denotes removal of the species from the system. The effect of allowing non-zero degradation of fused species is discussed in Section A.2 (see Eq.(A.11) and Figure A.2). Note that replication of a singleton changes the network state of the mtDNA into a fused species, since replication occurs within the same membrane-bound organelle. An alternative model of singletons which replicate into singletons leaves our central result (Figure 2.1D) unchanged (see Eq.(A.10)). The system may be considered neutral since both W and M possess the same replication and degradation rates per molecule of mtDNA at any instance in time.

Finally, mtDNAs of different genotypes may interact through fusion via a set of reactions we

term network cross-processes:



We note that any fusion or fission event which does not involve the generation or removal of a singleton leaves our system unchanged; we term such events as non-identity-changing processes, which can be ignored in our system (see Section A.4 for a discussion of rate renormalisation). Note that we have neglected *de novo* mutation in the model description above, see Section 2.3.5 for a treatment of *de novo* mutation using a modified infinite sites Moran model.

We found that treating $\lambda = \text{const}$ led to instability in total copy number (see Section A.1), which is not credible. We therefore favoured a state-dependent replication rate such that copy number is controlled to a particular value, as has been done by previous authors [40, 41, 43]. Allowing lower-case variables to denote the copy number of their respective molecular species, we will focus on a linear replication rate of the form [68, 209]:

$$\lambda = \lambda(w_T, m_T) = \mu + b(\kappa - (w_T + \delta m_T)) \quad (2.10)$$

where $w_T = w_s + w_f$ is the total wild-type copy number, and similarly for m_T . b is a parameter which determines the strength with which total copy number is controlled to a target copy number, and κ is a parameter which is indicative of (but not equivalent to) the steady state copy number. δ indicates the relative contribution of mutant mtDNAs to the control strength and is linked to the “maintenance of wild-type” hypothesis [173, 210]. When $0 \leq \delta < 1$, and both mutant and wild-type species are present, mutants have a lower contribution to the birth rate than wild-types. When wild-types are absent, the population size will be larger than when there are no mutants: hence mutants have a higher carrying capacity in this regime. We have modelled the mitophagy rate as constant per mtDNA. We do, however, explore relaxing this constraint below by allowing mitophagy to be a function of state, and also affect mutants differentially under quality control. Analogues of this model (without a network) have been applied to mitochondrial systems [40, 41]. Overall, our simple model consists of 4 species (W_S, W_F, M_S, M_F), 6 parameters and 15 reactions, and captures the central property that mitochondria fragment before degradation [164].

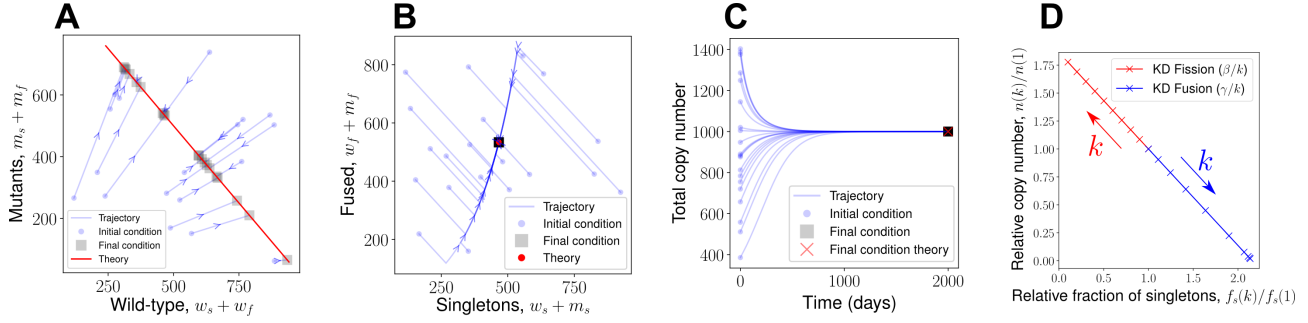


Figure 2.2: Deterministic analysis of a mathematical model of mitochondrial network and genetic dynamics. (A) Dynamics of total wild-type and mutant copy numbers. Trajectories move towards a line of steady states, as indicated by arrows. (B) Dynamics of total number of fused and unfused species. All trajectories move towards a particular value in this subspace. (C) Deterministic dynamics of total copy number, which is controlled to a particular steady state value. (D) Defining a knock-down (KD) factor ($k^{-1} = 0.1, 0.2 \dots, 1.0$), the fission rate was rescaled to $\beta \rightarrow \beta/k$ (red) and the fusion rate to $\gamma \rightarrow \gamma/k$ (blue), causing a linear increase and decrease in total copy number respectively under a deterministic treatment.

2.3.2 MITOCHONDRIAL NETWORK STATE RESCALES THE LINEAR INCREASE OF HETEROPLASMY VARIANCE OVER TIME INDEPENDENTLY OF FISSION-FUSION RATE MAGNITUDES

We first performed a deterministic analysis of the system presented in Eq.(2.1)–Eq.(2.10), by converting the reactions into an analogous set of four coupled ordinary differential equations (for a derivation see Section 2.3.3.2), and choosing a biologically-motivated approximate parametrisation (which we will term as the ‘nominal’ parametrisation, see Section A.3). Figures 2.2A&B show that copy numbers of each individual species change in time such that the state approaches a line of steady states (Eq.(2.36)–Eq.(2.38)), as seen in other neutral genetic models [41, 209]. Upon reaching this line, total copy number remains constant (Figure 2.2C) and the state of the system ceases to change with time. This is a consequence of performing a deterministic analysis, which neglects stochastic effects, and our choice of replication rate in Eq.(2.10) which decreases with total copy number when $w_T + \delta m_T > \kappa$ and vice versa, guiding the total population to a fixed total copy number. Varying the fission (β) and fusion (γ) rates revealed a negative linear relationship between the steady-state fraction of singletons and copy number (Figure 2.2D).

We may also simulate the system in Eq.(2.1)–Eq.(2.9) stochastically, using the stochastic simulation algorithm [211], which showed that mean copy number is slightly perturbed from the deterministic prediction due to the influence of variance upon the mean [209, 212] (Figure 2.3A). The stationarity of total copy number is a consequence of using $\delta = 1$ for our nominal parametrisation (i.e. the line of steady states is also a line of constant copy number). Choosing

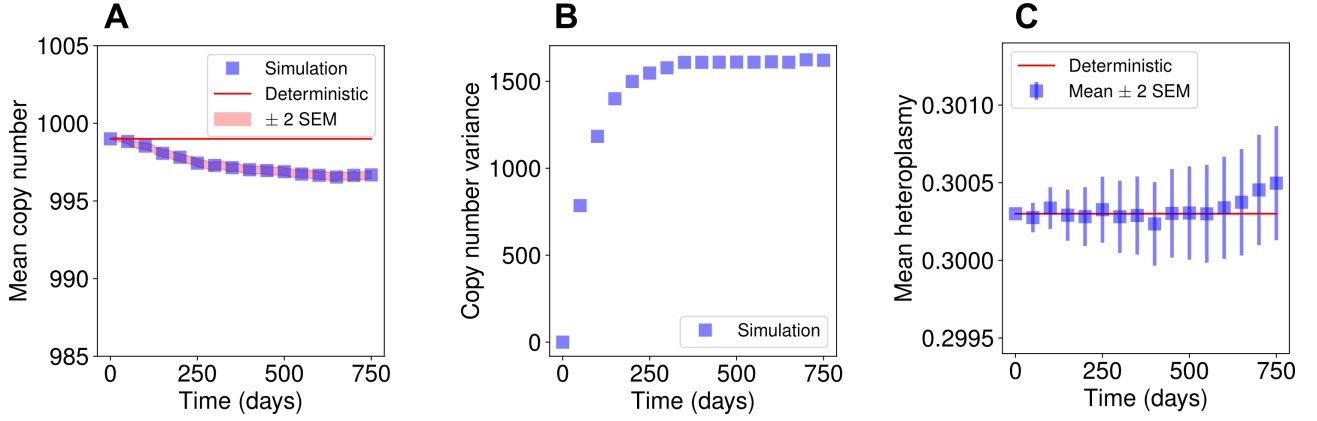


Figure 2.3: Stochastic simulations of a mathematical model of mitochondrial network and genetic dynamics. Stochastic simulations of the network system for the nominal parametrisation in Table 2.1 (10^5 realisations). Trajectories are initialised on the deterministic steady state line at $h = 0.3$. **(A)** Mean copy number is slightly perturbed from the deterministic treatment of the system, and then remains constant. **(B)** Copy number variance for stochastic simulations initially increases, since all stochastic simulations begin with the same initial condition, but then plateaus since the steady state line is globally attracting. **(C)** Stochastic simulations are concordant with mean heteroplasmy dynamics being invariant with time (see Eq.(2.63)).

$\delta \neq 1$ results in a difference in carrying capacities between the two species, and non-stationarity of mean total copy number, as trajectories spread along the line of steady states to different total copy numbers. Copy number variance initially increases since trajectories are all initialised at the same state, but plateaus because trajectories are constrained in their copy number to remain near the attracting line of steady states (Figure 2.3B). Mean heteroplasmy remains constant through time under this model (Figure 2.3C, see [39]). This is unsurprising since each species possesses the same replication and degradation rate, so neither species is preferred.

From stochastic simulations we observed that, for sufficiently short times, heteroplasmy variance increases approximately linearly through time for a range of parametrisations (Figure 2.4), which is in agreement with recent single-cell oocyte measurements in mice [72]. Previous work has also shown a linear increase in heteroplasmy variance through time for purely genetic models of mtDNA dynamics (see [43]). We sought to understand the influence of mitochondrial network dynamics upon the rate of increase of heteroplasmy variance.

To this end, we analytically explored the influence of mitochondrial dynamics on mtDNA variability. Assuming that the state of system above ($\mathbf{x} = (w_s, w_f, m_s, m_f)$) is initialised at its deterministic steady state ($\mathbf{x}(t = 0) = \mathbf{x}_{ss}$), we took the limit of limit of large mtDNA copy numbers, fast fission-fusion dynamics, and applied a second-order truncation of the Kramers-Moyal expansion [213] to the chemical master equation describing the dynamics of the system (see Section 2.3.3). This yielded a stochastic differential equation for heteroplasmy, via Itô's formula [214]. Upon forcing the state variables onto the steady-state line [215], we derived

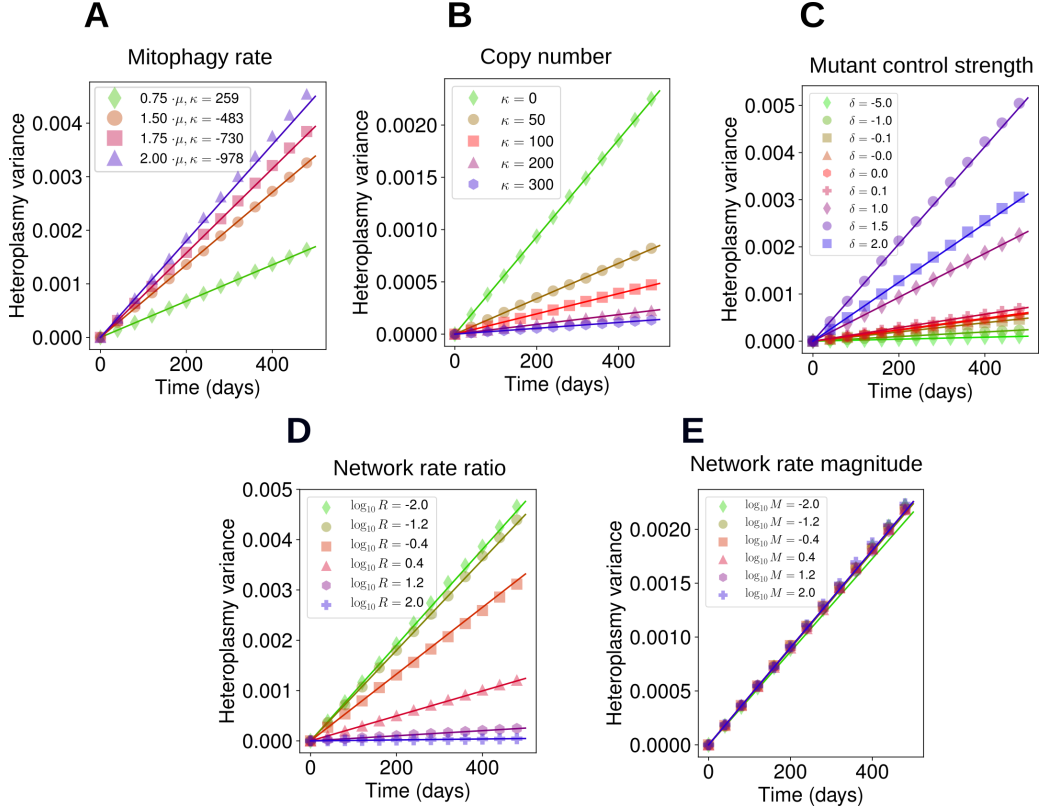


Figure 2.4: Heteroplasmy variance is modulated by network fragmentation but not absolute fission-fusion rates over a wide range of biologically plausible parametrisations. Robustness of Eq.(2.11). Solid lines are from Eq.(2.11) and markers are from stochastic simulation (10^4 realisations per parametrisation). **(A)** Increasing the mitophagy rate, and correspondingly varying the value of κ to maintain a constant copy number of $n = 1000$, results in an increase in heteroplasmy variance. All simulations are initialised at $h = 0.3$ and μ is varied relative to the nominal parametrisation. **(B)** Increasing the value of the deterministic steady state copy number, by increasing κ , reduces heteroplasmy variance. **(C)** Eq.(2.11) predicts $\mathbb{V}(h)$ for a variety of values of δ , which is hypothesised to model different types of mtDNA mutation. **(D)** Fusion and fission rates are redefined as $\gamma \rightarrow \gamma_0 MR$ and $\beta \rightarrow \beta_0 M$ where M and R denote the relative magnitude and ratio of the network rates, and γ_0, β_0 denote the nominal parametrisations of the fusion and fission rates respectively (see Table 2.1). Increasing R (namely the fusion rate at a fixed fission rate) at $M = 1$ reduces heteroplasmy variance. **(E)** Sweeps of the network rate magnitude M . Heteroplasmy variance is approximately independent of absolute network rates over a broad range of network magnitudes. See Figure A.3 and Table A.1 for parameter sweeps numerically demonstrating the generality of the result for different mtDNA control modes.

Eq.(2.65), which may be approximated for sufficiently short times as

$$\mathbb{V}(h) \approx f_s(\mathbf{x}) \frac{2\mu t}{n(\mathbf{x})} h(\mathbf{x})(1 - h(\mathbf{x})) \Big|_{\mathbf{x}=\mathbf{x}_{ss}}. \quad (2.11)$$

$h(\mathbf{x}) = (m_s + m_f)/(w_s + w_f + m_s + m_f)$ is heteroplasmy (a random variable since it is a function of random state variables $\mathbf{x}$), $\mathbb{V}(h)$ is the variance of heteroplasmy, μ is the mitophagy rate, $n(\mathbf{x})$ is the total copy number and $f_s(\mathbf{x}) = (m_s + w_s)/n(\mathbf{x})$ is the fraction of unfused (singleton) mtDNAs, and is thus a measure of the fragmentation of the mitochondrial network. Note that the three functions of state in Eq.(2.11) have been forced to be their deterministic steady state value ($\mathbf{x}_{ss}$), which is equivalent to their initial values by model construction. We find that Eq.(2.11) closely matches heteroplasmy variance dynamics from stochastic simulation, for sufficiently short times after initialisation, for a variety of parametrisations of the system (Figure 2.4 and A.3A). In Section 2.3.3 we present our derivation of Eq.(2.11).

To our knowledge, Eq.(2.11) reflects the first analytical principle linking mitochondrial dynamics and the cellular population genetics of mtDNA variance. Its simple form allows several intuitive interpretations. As time progresses, replication and degradation of both species occurs, allowing the ratio of species to fluctuate; hence we expect $\mathbb{V}(h)$ to increase with time according to random genetic drift (Figure 2.4). The rate of occurrence of replication/degradation events is set by the mitophagy rate μ , since degradation events are balanced by replication to maintain population size; hence, random genetic drift occurs more quickly if there is a larger turnover in the population (Figure 2.4A). We expect $\mathbb{V}(h)$ to increase more slowly in large population sizes, since the birth of e.g. 1 mutant in a large population induces a small change in heteroplasmy (Figure 2.4B). The factor of $h(1 - h)$ encodes the state-dependence of heteroplasmy variance, exemplified by the observation that if a cell is initialised at $h = 0$ or $h = 1$, heteroplasmy must remain at its initial value (since the model above does not consider *de novo* mutation, see Section 2.3.5) and so heteroplasmy variance is zero. Furthermore, the rate of increase of heteroplasmy variance is maximal when a cell's initial value of heteroplasmy is $1/2$. In Figure 2.4C, we show that Eq.(2.11) is able to recapitulate the rate of heteroplasmy variance increase across different values of δ , which are hypothesised to correspond to different replicative sensing strengths of different mitochondrial mutations [209]. We also show in Figure 2.5 that Eq.(2.11) is robust to the choice of feedback control strength b in Eq.(2.10).

In Eq.(2.6), we have made the important assumption that only unfused mitochondria can be degraded via mitophagy, as seen by Twig *et al.* [164], hence the total propensity of mtDNA turnover is limited by the number of mtDNAs which are actually susceptible to mitophagy.

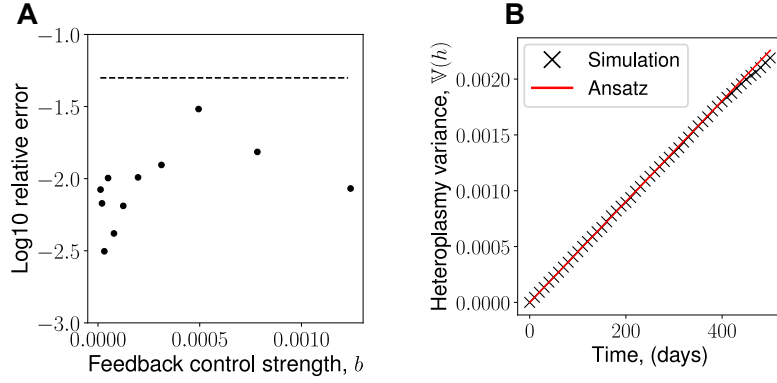


Figure 2.5: Heteroplasmy variance ansatz is robust to different replication control strengths. (A) Error in Eq.(2.11) in a sweep over the feedback control strength, b . Dotted line denotes a 5% error according to Eq.(A.28). (B) $\mathbb{V}(h)$ profile for the parametrisation with the largest error in (A).

Strikingly, we find that *the dynamics of heteroplasmy variance are independent of the absolute rate of fusion and fission, only depending on the fraction of unfused mtDNAs at any particular point in time* (see Figure 2.4D&E). This observation, which contrasts with the model of [206, 207] (see Discussion), arises from the observation that mitochondrial network dynamics are much faster than replication and degradation of mtDNA, by around a factor of $\beta/\mu \approx 10^3$ (see Table 2.1), resulting in the existence of a separation of timescales between network and genetic processes. In the derivation of Eq.(2.11) (see Section 2.3.3), we assume that fission-fusion rates are infinite, which simplifies $\mathbb{V}(h)$ into a form which is independent of the magnitude of the fission-fusion rate. A parameter sweep of the magnitude and ratio of the fission-fusion rates reveals that, if the fusion and fission rates are sufficiently small, Eq.(2.11) breaks down and $\mathbb{V}(h)$ gains dependence upon the magnitude of these rates (see Figure A.3A). This regime is, however, for network rates which are approximately 100 times smaller than the biologically-motivated nominal parametrisation shown in Figure 2.4A-D where the fission-fusion rate becomes comparable to the mitophagy rate. Since fission-fusion takes place on a faster timescale than mtDNA turnover, we may neglect this region of parameter space as being implausible.

In the following section, we explore in mathematical detail how Eq.(2.11) may be derived from the chemical master equation using the Kramers-Moyal expansion and a large copy number assumption, in the limit of large network fission-fusion rates.

2.3.3 DERIVATION OF HETEROPLASMY VARIANCE BEHAVIOUR FOR A MODEL OF MITOCHONDRIAL NETWORK & GENETIC DYNAMICS

In this section we seek to prove Eq.(2.11) from first principles. This section is more mathematical in its content; less mathematically-inclined readers may wish to skip to Section 2.3.4.

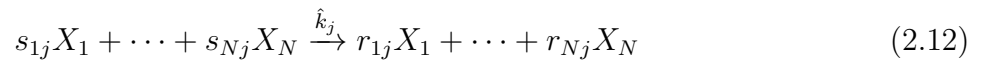
Parameter	Description	Value	Dimensions	Remarks
β	Fission rate	33.12	day ⁻¹	For cortical neurons, see [216]
γ	Fusion rate	3.79×10^{-2}	day ⁻¹	Approximately balances fission
μ	Mitophagy rate	0.023	day ⁻¹	From [207]
δ	Mutant feedback sensitivity	1	dimensionless	Potentially appropriate for e.g. non-coding mtDNA mutations
b	Feedback control strength	1.24×10^{-5}	day ⁻¹	Chosen as the weakest control strength which has a non-trivial steady state and total copy number of 1000
κ	Steady state copy number parameter	11.7	dimensionless	See remark for b

Table 2.1: Nominal parametrisation for network system. Dimensions of parameters are derived using individual terms in Eq.(2.31)–Eq.(2.34); copy numbers of particular species are pure numbers and therefore dimensionless. Since we have chosen a system size of $\Omega = 1$ throughout, we set the dimension of volume to 1. See Section A.3 for further details of parameter justification.

2.3.3.1 CONVERSION OF A CHEMICAL REACTION NETWORK INTO ORDINARY DIFFERENTIAL EQUATIONS

In order to proceed with deriving Eq.(2.11), we must first be able to cast our chemical reaction network in Eq.(2.1)–Eq.(2.10) into a set of ordinary differential equations: this is necessary for both preliminary deterministic analysis (see Figure 2.2) and is required in the subsequent derivation. We will therefore devote this section to understanding how to convert a general chemical reaction network into a set of ODEs. In particular, we pay special attention to the fact that the rate of a chemical reaction with a stochastic treatment is not always equivalent to the rate in a deterministic treatment [217], as we will explain below. This subtlety is sometimes overlooked in the literature. This section draws on a number of standard texts [211, 217–219] as well as Ref. [220]. We hope this harmonised treatment will be of help as a future reference.

Consider a general chemical system consisting of N distinct chemical species (X_i) interacting via R chemical reactions, where the j^{th} reaction is of the form



where s_{ij} and r_{ij} are stoichiometric coefficients. We define $\hat{k}_j$ as the *microscopic* rate for this reaction. The dimensionality of this parameter will vary depending upon the stoichiometric coefficients s_{ij} . $\hat{k}_j$ may be loosely interpreted as setting the characteristic timescale (i.e. the cross section [217]) of reaction j .

The chemical master equation (CME) describes the dynamics of the joint probability distribu-

tion of the state of the system and time, moving forwards through time. Defining the state of the system as $\mathbf{x} = (x_1, \dots, x_N)^T$, where x_i is the copy number of the i^{th} species, allows us to write the CME as [220]

$$\frac{\partial P(\mathbf{x}, t | \mathbf{x}_0, t_0)}{\partial t} = \Omega \sum_{j=1}^R \left(\prod_{i=1}^N E_i^{-S_{ij}} - 1 \right) \hat{f}_j(\mathbf{x}, \Omega) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \quad (2.13)$$

where Ω is the volume of the compartment in which the reactions occur (also known as the system size), $S_{ij} = r_{ij} - s_{ij}$ is the stoichiometry matrix, and $E_i^{-S_{ij}}$ is referred to as the step operator and is defined through the relation $E_i^{-S_{ij}}(g(\mathbf{x})) = g(x_1, \dots, x_i - S_{ij}, \dots, x_N)$, for any function of state $g(\mathbf{x})$. $\hat{f}_j(\mathbf{x}, \Omega)$ is the microscopic rate function of reaction j , which in general depends on both the state and the system size. A factor of Ω is explicitly included in this definition of the chemical master equation so that our treatment is compatible with Van Kampen's system size expansion [218]. As a consequence of this, the probability that, given the current state $\mathbf{x}$, the j^{th} reaction occurs in the time interval $[t, t + dt)$ somewhere in Ω [219] is

$$\hat{a}_j(\mathbf{x}, \Omega) dt := \Omega \hat{f}_j(\mathbf{x}, \Omega) dt. \quad (2.14)$$

$\hat{a}_j(\mathbf{x}, \Omega)$ is termed the propensity function (or ‘‘hazard’’) and is of particular relevance in the stochastic simulation algorithm [211], since $\hat{a}_j(\mathbf{x}, \Omega) / \sum_j \hat{a}_j(\mathbf{x}, \Omega)$ determines the probability that the j^{th} reaction occurs next.

For the microscopic rate function, we may write

$$\hat{f}_j(\mathbf{x}, \Omega) = \hat{k}_j \prod_{i=1}^N \Omega^{-s_{ij}} \binom{x_i}{s_{ij}}. \quad (2.15)$$

This equation counts the number of available combinations of reacting molecules [211, 217], whilst taking into account scaling with system size [220].

We also introduce the deterministic rate equation (generally considered to be the macroscopic analogue of the CME) which is defined as [218, 220]

$$\frac{d\phi_i}{dt} = \sum_{j=1}^R S_{ij} \tilde{f}_j(\phi) \quad (2.16)$$

where $\phi = (\phi_1, \dots, \phi_N)^T$ is the vector of macroscopic concentrations (of dimensions molecules

per unit volume) and $\tilde{f}_j(\phi)$ is the macroscopic rate function satisfying

$$\tilde{f}_j(\phi) = \tilde{k}_j \prod_{i=1}^N \phi_i^{s_{ij}} \quad (2.17)$$

where $\tilde{k}_j$ is the *macroscopic* rate for the j^{th} reaction. Note that we distinguish between $\hat{k}_j$ and $\tilde{k}_j$, respectively the rate constants for the discrete and continuous pictures, although this distinction is sometimes not emphasised in the literature [212, 218, 220]. The physical meaning of $\tilde{k}_j$ is not immediately obvious: we argue that this parameter only gains physical meaning through the following procedure.

As stated by Wilkinson [217], if we intend for the microscopic description in Eq.(2.13) to correspond to the macroscopic description in Eq.(2.16), the rate of consumption/production of particles for every reaction must be the same in the deterministic limit of the stochastic system (the conditions for which we define below). Therefore, we apply the following constraint in the limit of large copy numbers

$$\lim_{x_i \rightarrow \infty} \hat{f}_j(\mathbf{x}, \Omega) = \tilde{f}_j(\phi) \quad \forall i, j. \quad (2.18)$$

In applying this constraint on all species i and all reactions j , we may derive a general relationship between $\hat{k}_j$ and $\tilde{k}_j$

$$\lim_{x_i \rightarrow \infty} \hat{k}_j \prod_{i=1}^N \Omega^{-s_{ij}} \frac{x_i!}{s_{ij}! (x_i - s_{ij})!} = \tilde{k}_j \prod_{i=1}^N \phi_i^{s_{ij}}. \quad (2.19)$$

We can make two approximations to generate a more convenient relationship between the microscopic and macroscopic rates. Firstly, we assume that

$$x_i \approx \Omega \phi_i. \quad (2.20)$$

This is a small noise approximation, since it is often assumed that $x_i = \Omega \phi_i + \Omega^{1/2} \xi_i$, where ξ_i is a noise term [218]. If ξ_i is small then $x_i \approx \Omega \phi_i$ is a valid approximation. Secondly, we assume that

$$x_i(x_i - 1) \dots (x_i - s_{ij} + 1) \approx x_i^{s_{ij}}. \quad (2.21)$$

This is a large copy number approximation: in the case of e.g. a bimolecular reaction ($2X_i \rightarrow *$) with $s_{ij} = 2$, the approximation is of the form $x_i(x_i - 1) \approx x_i^2$ or $x_i \approx x_i - 1$. By applying Eq.(2.21) to the factor of $x_i!$ in Eq.(2.19), the factor of $(x_i - s_{ij})!$ cancels from the left-hand side. Simplifying using Eq.(2.20), $\phi_i^{s_{ij}}$ cancels from both sides and we arrive at the important

relationship

$$\tilde{k}_j \approx \hat{k}_j \prod_{i=1}^N \frac{1}{s_{ij}!}. \quad (2.22)$$

With Eq.(2.16), Eq.(2.17) and Eq.(2.22) one may therefore write down a set of ODEs for an arbitrary chemical reaction network, with constant reaction rates, in terms of the microscopic rates $\hat{k}_j$. This equation highlights that for reactions with $s_{ij} \geq 2$, $\tilde{k}_j \neq \hat{k}_j$, as is the case for bimolecular reactions of the form $2X_i \rightarrow *$ (see Eq.(2.1) and Eq.(A.1)).

It is important to note that if the microscopic rate function is a function of state then $\hat{k} = \hat{k}(\mathbf{x})$ and $\tilde{k} = \tilde{k}(\phi) \approx \tilde{k}(\mathbf{x}/\Omega)$. In this case, Eq.(2.22) still applies since the above argument assumed nothing about the particular forms of $\hat{k}$ and $\tilde{k}$. However, additional factors of Ω^{-1} are induced by applying Eq.(2.20), which may carry through to the individual parameters of $\tilde{k}(\phi)$. A demonstration of this is given in the following section for the linear feedback control in Eq.(2.10).

2.3.3.2 DERIVING AN ODE DESCRIPTION OF THE MITOCHONDRIAL NETWORK SYSTEM

In this section we show how to derive an ODE description of the network system described in Eq.(2.1)-Eq.(2.9) in Section 2.3.1. In accordance with the notation in the previous section, we will redefine all of the rates in Eq.(2.1)-Eq.(2.10) with a hat notation ($\hat{a}$, for a general rate parameter a), to reflect that these are stochastic rates. Deterministic rates will be denoted with a tilde ($\tilde{a}$). Our aim will be to write a set of ODEs in terms of the stochastic rates, $\hat{a}$, for which we are able to estimate values.

We will begin by considering the fusion network equations Eq.(2.1) and Eq.(2.2). For clarity, we rewrite Eq.(2.1) to allow the reaction to proceed with some arbitrary rate $\hat{\rho}$:



where X denotes either a wild-type (W) or mutant (M). We will subsequently fix $\hat{\rho}$ to the rate of all other fusion reactions $\hat{\gamma}$. We do this because Eq.(2.1) is a bimolecular reaction involving one species: a fundamentally different reaction to bimolecular reactions involving two species, as we will now see.

Since $\hat{\rho}, \hat{\gamma} = \text{const}$, we may use Eq.(2.22), resulting in the deterministic rates

$$\tilde{\rho} = \frac{\hat{\rho}}{2} \quad (2.24)$$

$$\tilde{\gamma} = \hat{\gamma} \quad (2.25)$$

for Eq.(2.23) and Eq.(2.2) respectively. If we then enforce the microscopic rates to be equal for both of these fusion reactions, i.e. $\hat{\rho} = \hat{\gamma}$, then $\tilde{\rho} = \hat{\gamma}/2$. All other fusion reactions have $\tilde{\gamma} = \hat{\gamma}$ by application of Eq.(2.22). Application of Eq.(2.22) to the fission reaction in Eq.(2.3) shows that $\hat{\beta} = \tilde{\beta}$.

For Eq.(2.4), we have chosen a $\hat{\lambda}$ which is not a constant, but a function of the copy numbers of the chemical species ($\hat{\lambda} = \hat{\lambda}(\mathbf{x})$ where $\mathbf{x} = (w_s, w_f, m_s, m_f)$, see Eq.(2.10)). As pointed out in the previous section, care must be taken in writing down the deterministic analogue of $\hat{\lambda}$. Applying Eq.(2.22), we have

$$\hat{\mu} + \hat{b} \left(\hat{\kappa} - (w_s + w_f + \hat{\delta}m_s + \hat{\delta}m_f) \right) = \tilde{\mu} + \tilde{b} \left(\tilde{\kappa} - (\phi_{w_s} + \phi_{w_f} + \tilde{\delta}\phi_{m_s} + \tilde{\delta}\phi_{m_f}) \right). \quad (2.26)$$

Applying Eq.(2.20) and equating individual terms, we arrive at

$$\hat{\mu} = \tilde{\mu} \quad (2.27)$$

$$\hat{b}w_s = \tilde{b}\frac{w_s}{\Omega} \implies \hat{b} = \tilde{b}/\Omega \quad (2.28)$$

$$\hat{b}\hat{\kappa} = \tilde{b}\tilde{\kappa} \implies \hat{\kappa} = \tilde{\kappa}\Omega \quad (2.29)$$

$$\hat{b}\hat{\delta}m_s = \tilde{b}\tilde{\delta}\phi_{m_s} \implies \hat{\delta} = \tilde{\delta}. \quad (2.30)$$

In this study, we let $\Omega = 1$ so the above 4 parameters are identical to their deterministic counterparts. Hence, by application of Eq.(2.16), we arrive at the following set of ODEs

$$\begin{aligned} \frac{d\phi_{w_s}}{dt} = & -2 \cdot \frac{\hat{\gamma}}{2} \phi_{w_s}^2 - \hat{\gamma} \phi_{w_s} \phi_{w_f} + \hat{\beta} \phi_{w_f} - \left(\hat{\mu} + \hat{b} \left(\hat{\kappa} - (\phi_{w_s} + \phi_{w_f} + \hat{\delta}\phi_{m_s} + \hat{\delta}\phi_{m_f}) \right) \right) \phi_{w_s} \\ & - \hat{\mu} \phi_{w_s} - \hat{\gamma} \phi_{m_f} \phi_{w_s} - \hat{\gamma} \phi_{w_s} \phi_{m_s} \end{aligned} \quad (2.31)$$

$$\begin{aligned} \frac{d\phi_{m_s}}{dt} = & -2 \cdot \frac{\hat{\gamma}}{2} \phi_{m_s}^2 - \hat{\gamma} \phi_{m_s} \phi_{m_f} + \hat{\beta} \phi_{m_f} - \left(\hat{\mu} + \hat{b} \left(\hat{\kappa} - (\phi_{w_s} + \phi_{w_f} + \hat{\delta}\phi_{m_s} + \hat{\delta}\phi_{m_f}) \right) \right) \phi_{m_s} \\ & - \hat{\mu} \phi_{m_s} - \hat{\gamma} \phi_{w_f} \phi_{m_s} - \hat{\gamma} \phi_{w_s} \phi_{m_s} \end{aligned} \quad (2.32)$$

$$\begin{aligned} \frac{d\phi_{w_f}}{dt} = & 2 \cdot \frac{\hat{\gamma}}{2} \phi_{w_s}^2 + \hat{\gamma} \phi_{w_s} \phi_{w_f} - \hat{\beta} \phi_{w_f} \\ & + \left(\hat{\mu} + \hat{b} \left(\hat{\kappa} - (\phi_{w_s} + \phi_{w_f} + \hat{\delta}\phi_{m_s} + \hat{\delta}\phi_{m_f}) \right) \right) (2\phi_{w_s} + \phi_{w_f}) + \hat{\gamma} \phi_{m_f} \phi_{w_s} + \hat{\gamma} \phi_{w_s} \phi_{m_s} \end{aligned} \quad (2.33)$$

$$\begin{aligned} \frac{d\phi_{m_f}}{dt} = & 2 \cdot \frac{\hat{\gamma}}{2} \phi_{m_s}^2 + \hat{\gamma} \phi_{m_s} \phi_{m_f} - \hat{\beta} \phi_{m_f} \\ & + \left(\hat{\mu} + \hat{b} \left(\hat{\kappa} - (\phi_{w_s} + \phi_{w_f} + \hat{\delta}\phi_{m_s} + \hat{\delta}\phi_{m_f}) \right) \right) (2\phi_{m_s} + \phi_{m_f}) + \hat{\gamma} \phi_{w_f} \phi_{m_s} + \hat{\gamma} \phi_{w_s} \phi_{m_s}. \end{aligned} \quad (2.34)$$

The steady state solution of this system of ODEs may be calculated, but its form is complex. For notational simplicity, we will drop the hat notation. Defining

$$x_1 = (b^2 (\beta^2 + 2\beta(\gamma\kappa + 3\mu) + \gamma^2\kappa^2 + \mu^2 + 2\gamma\mu(\kappa + 2(\delta - 1)\phi_{m_s})) + 2b\gamma\mu(-\beta + \mu + \gamma(\kappa - 2\delta\phi_{m_s} + 2\phi_{m_s})) + \gamma^2\mu^2)^{1/2} \quad (2.35)$$

the non-trivial, physically-realisable, component of the steady state may be parameterised in terms of ϕ_{m_s} and written as

$$\phi_{w_s} = -(\beta b^2\kappa + b^2\gamma\kappa^2 + b^2\kappa\mu + 2b^2\delta\mu\phi_{m_s} + \beta^2b + \beta b\gamma\kappa + 5\beta b\mu + 3b\gamma\kappa\mu - \beta\gamma\mu + 2b\mu^2 - 2b\gamma\delta\mu\phi_{m_s} - 2b\gamma\mu\phi_{m_s} - x_1(b\kappa + \beta + 2\mu) + 2\gamma\mu^2 + 2\gamma^2\mu\phi_{m_s})/(2\mu(b - \gamma)^2) \quad (2.36)$$

$$\begin{aligned} \phi_{w_f} = & (b^3 (\beta^2\kappa + \beta(\gamma\kappa(\kappa + (\delta - 1)\phi_{m_s}) + \mu(3\kappa + \delta\phi_{m_s})) \\ & + \phi_{m_s} (\gamma^2(\delta - 1)\kappa^2 - \delta\mu^2 + \gamma\mu((2\delta - 3)\kappa + 2(\delta - 1)\delta\phi_{m_s}))) + b^2 (\beta^3 \\ & + \beta^2(6\mu + \gamma(\kappa + (\delta - 1)\phi_{m_s})) + \beta (5\mu^2 + \gamma^2\delta\kappa\phi_{m_s} + \gamma^2\kappa(-\phi_{m_s}) + 6\gamma\delta\mu\phi_{m_s} - 8\gamma\mu\phi_{m_s} - \kappa x_1) \\ & + \phi_{m_s} (2\gamma^2\mu (\delta\kappa + \delta^2(-\phi_{m_s}) + \phi_{m_s}) + \gamma ((6\delta - 5)\mu^2 - (\delta - 1)\kappa x_1) - \delta\mu x_1)) \\ & - b (\beta (\gamma^2\mu(-(\kappa + (4 - 3\delta)\phi_{m_s})) + \gamma (2\mu^2 + (\delta - 1)\phi_{m_s}x_1) + 3\mu x_1) \\ & + \gamma\mu\phi_{m_s} (\gamma(\delta - 2)\mu + \gamma^2(\kappa - 2(\delta - 1)\phi_{m_s}) + (\delta - 3)x_1) + \beta^2(2\gamma\mu + x_1)) \\ & + \gamma\mu(\beta - \gamma\phi_{m_s})(\gamma\mu + x_1))/(2b\mu(b - \gamma)^2(\beta + \gamma(\delta - 1)\phi_{m_s})) \end{aligned} \quad (2.37)$$

$$\phi_{m_f} = (\phi_{m_s}(-b\beta + b\gamma\kappa + b\mu - 2b\gamma\delta\phi_{m_s} + 2b\gamma\phi_{m_s} + \gamma\mu + x_1))/(2b(\beta + \gamma(\delta - 1)\phi_{m_s})). \quad (2.38)$$

Note that, since the steady state is parametrised by ϕ_{m_s} , the steady state is therefore a line.

2.3.3.3 PROOF OF HETEROPLASMY RELATION FOR LINEAR FEEDBACK CONTROL

In this section we show that Eq.(2.11) holds for the system described by Eq.(2.1)-Eq.(2.9) given the replication rate in Eq.(2.10) using the Kramers-Moyal expansion under conditions of large copy number and fast network churn (to be defined below); the approach used here is similar to Ref. [215].

KRAMERS-MOYAL EXPANSION OF THE CHEMICAL MASTER EQUATION FOR LARGE COPY NUMBERS Customarily, the Kramers-Moyal expansion is formed using a continuous-space notation [213], so we will initially proceed in this way. Following the treatment by Gardiner [213],

we begin by re-writing the chemical master equation Eq.(2.13) (CME) as

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = \int_{-\infty}^{\infty} d\mathbf{x}' [T(\mathbf{x}|\mathbf{x}')P(\mathbf{x}', t) - T(\mathbf{x}'|\mathbf{x})P(\mathbf{x}, t)] \quad (2.39)$$

where we have set $\Omega = 1$. $T(\mathbf{x}|\mathbf{x}')$ is the transition rate from state $\mathbf{x}' \rightarrow \mathbf{x}$, and the dependence upon the initial condition has been suppressed for notational convenience. We now proceed by expanding the CME. The multivariate Kramers-Moyal expansion may be written as

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} \approx \int_{-\infty}^{\infty} \left(-\nabla (T(\mathbf{x}'|\mathbf{x})P(\mathbf{x}))^T \cdot (\mathbf{x}' - \mathbf{x}) + \frac{1}{2}(\mathbf{x}' - \mathbf{x})^T \cdot \mathbf{H} \cdot (\mathbf{x}' - \mathbf{x}) \right) d\mathbf{x}' \quad (2.40)$$

where $\mathbf{H}(\mathbf{x})$ is the Hessian matrix of $T(\mathbf{x}'|\mathbf{x})P(\mathbf{x})$

$$\mathbf{H} := \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} & \cdots & \frac{\partial^2}{\partial x_1 \partial x_N} \\ \vdots & & \vdots \\ \frac{\partial^2}{\partial x_N \partial x_1} & \cdots & \frac{\partial^2}{\partial x_N^2} \end{pmatrix} T(\mathbf{x}'|\mathbf{x})P(\mathbf{x}) \quad (2.41)$$

(see [213] for a proof of this in the univariate case).

Note that a transition to each possible neighbouring state $\mathbf{x}'$ corresponds to some reaction j which moves the state from $\mathbf{x} \rightarrow \mathbf{x}'$. Since we know the influence of each reaction on state $\mathbf{x}$ through the constant stoichiometry matrix S_{ij} , and that the propensity of a reaction does not depend upon $\mathbf{x}'$ itself (see Eq.(2.15)), we may transition from a notation involving $\mathbf{x}$ and $\mathbf{x}'$ into a notation involving $\mathbf{x}$ and j . We may therefore define $T_j(\mathbf{x}) := T(\mathbf{x}'|\mathbf{x}) \equiv \hat{f}_j(\mathbf{x})$ (see Eq.(2.15)), and let $\mathbf{H}(\mathbf{x}) \rightarrow \mathbf{H}_j(\mathbf{x})$.

We now make a large copy number assumption in order to simplify $T_j(\mathbf{x})$. To take a large copy number limit, we assume that $x_i! \approx x_i^{s_{ij}}(x_i - s_{ij})!$ resulting in

$$T_j(\mathbf{x}) \approx \hat{k}_j \prod_{i=1}^N \frac{x_i^{s_{ij}}}{s_{ij}!}. \quad (2.42)$$

Note that this approximation is exact when $s_{ij} = 0, 1$, but inexact when $s_{ij} \geq 2$. For example, if we consider the second-order bimolecular reaction in Eq.(2.1), Eq.(2.42) is equivalent to assuming $w_s^2 \approx w_s(w_s - 1)$; consequently, a factor of $1/(s_{ij}!) = 1/2$ arises in $T_j(\mathbf{x})$ as a combinatorial factor from stochastic considerations.

FOKKER-PLANCK EQUATION FOR CHEMICAL REACTION NETWORKS We now wish to re-write Eq.(2.40) as a Fokker-Planck equation. Since the integral in Eq.(2.40) is over $\mathbf{x}'$, and every $\mathbf{x}'$ corresponds to a reaction j , we may interpret the integral in Eq.(2.40) as a sum over all reactions, i.e. $\int d\mathbf{x}' \rightarrow \sum_{j=1}^R$. Hence, for the j^{th} reaction, $[(\mathbf{x}' - \mathbf{x})]_i = S_{ij}$. With these observations, we may write the first integral of Eq.(2.40) as

$$\begin{aligned} \int_{-\infty}^{\infty} -\nabla(T(\mathbf{x}'|\mathbf{x})P(\mathbf{x}))^T \cdot (\mathbf{x}' - \mathbf{x}) d\mathbf{x}' &= \int_{-\infty}^{\infty} -\nabla(T_j(\mathbf{x})P(\mathbf{x}, t))^T \cdot (\mathbf{x}' - \mathbf{x}) d\mathbf{x}' \\ &= -\sum_{j=1}^R \sum_{i=1}^N \frac{\partial}{\partial x_i} (T_j(\mathbf{x})P(\mathbf{x}, t)) S_{ij} \\ &= -\sum_{i=1}^N \frac{\partial}{\partial x_i} A_i P(\mathbf{x}, t) \end{aligned} \quad (2.43)$$

where

$$\mathbf{A} := \mathbf{S} \cdot \mathbf{T}. \quad (2.44)$$

$\mathbf{A}$ is a vector of length N , $[\mathbf{S}]_{ij} := r_{ij} - s_{ij}$ is the $N \times R$ stoichiometry matrix Eq.(2.12), and $\mathbf{T}$ is the vector of transition rates, of length R (for which we have taken a large copy number approximation in Eq.(2.42)). To re-write the second integral of Eq.(2.40), we write an element of the Hessian $\mathbf{H}_j$ in Eq.(2.41) as

$$H_{jlm} = \frac{\partial^2}{\partial x_l \partial x_m} T_j(\mathbf{x}) P(\mathbf{x}, t) \quad (2.45)$$

where $j = 1, \dots, R$ and $l, m = 1, \dots, N$. Thus, we may write

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{2} (\mathbf{x}' - \mathbf{x})^T \cdot \mathbf{H}_j \cdot (\mathbf{x}' - \mathbf{x}) d\mathbf{x}' &= \frac{1}{2} \sum_{j=1}^R \sum_{l=1}^N \sum_{m=1}^N S_{lj} H_{jlm} S_{mj} \\ &= \frac{1}{2} \sum_{j=1}^R \sum_{l=1}^N \sum_{m=1}^N S_{lj} \frac{\partial^2}{\partial x_l \partial x_m} T_j P(\mathbf{x}, t) S_{mj} \\ &= \frac{1}{2} \sum_{l=1}^N \sum_{m=1}^N \frac{\partial^2}{\partial x_l \partial x_m} \left(\sum_{j=1}^R S_{lj} T_j S_{mj} \right) P(\mathbf{x}, t) \\ &= \frac{1}{2} \sum_{i,m=1}^N \frac{\partial^2}{\partial x_i \partial x_m} B_{im} P(\mathbf{x}, t) \end{aligned} \quad (2.46)$$

where

$$\mathbf{B} := \mathbf{S} \cdot \text{Diag}(\mathbf{T}) \cdot \mathbf{S}^T. \quad (2.47)$$

$\mathbf{B}$ is an $N \times N$ matrix, and $\text{Diag}(\mathbf{Y})$ is a diagonal matrix whose main diagonal is the vector $\mathbf{Y}$. We may therefore re-write Eq.(2.40) as a Fokker-Planck equation for the state vector $\mathbf{x}$ of the form

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} \approx - \sum_{i=1}^N \frac{\partial}{\partial x_i} [A_i(\mathbf{x}) P(\mathbf{x}, t)] + \frac{1}{2} \sum_{i,m=1}^N \frac{\partial^2}{\partial x_i \partial x_m} [B_{im}(\mathbf{x}) P(\mathbf{x})]. \quad (2.48)$$

FOKKER-PLANCK EQUATION FOR AN ARBITRARY FUNCTION OF STATE We now wish to make a change of variables in Eq.(2.48) to write down a Fokker-Planck equation for an arbitrary scalar function of state $\mathbf{x}$ (which we will later set to be heteroplasmy). To do this, we wish to make use of Itô's formula, which allows a change of variables for an SDE. In general, the Fokker-Planck equation in Eq.(2.48) is equivalent [214] to the following Itô stochastic differential equation (SDE)

$$d\mathbf{x} = \mathbf{A} dt + \mathbf{G} d\mathbf{W} \quad (2.49)$$

where $\mathbf{G}\mathbf{G}^T \equiv \mathbf{B}$ (where $\mathbf{G}$ is an $N \times R$ matrix) and $d\mathbf{W}$ is a vector of independent Wiener increments of length R , and a Wiener increment dW satisfies

$$\int_0^t dW := W(t), \quad P(W, t) \equiv \frac{1}{\sqrt{2\pi t}} e^{-W^2/(2t)}. \quad (2.50)$$

Itô's formula states that, for an arbitrary function $h(\mathbf{x}, t)$ where $\mathbf{x}$ satisfies Eq.(2.49), we may write the following SDE

$$dh(\mathbf{x}, t) = \left\{ \frac{\partial h}{\partial t} + (\nabla h)^T \mathbf{A} + \frac{1}{2} \text{Tr} [\mathbf{G}^T \mathbf{H}_h(\mathbf{x}) \mathbf{G}] \right\} dt + (\nabla h)^T \mathbf{G} d\mathbf{W}, \quad (2.51)$$

where $\mathbf{H}_h(\mathbf{x})$ is the Hessian matrix of $h(\mathbf{x}, t)$ (see Eq.(2.41), where $T(\mathbf{x}'|\mathbf{x})P(\mathbf{x})$ should be replaced with $h(\mathbf{x}, t)$) and Tr is the trace operator. Given the form of $\mathbf{B}$ in Eq.(2.47) we let

$$\mathbf{G} = \mathbf{S} \cdot \text{Diag}(\sqrt{\mathbf{T}}), \quad (2.52)$$

which satisfies $\mathbf{G}\mathbf{G}^T \equiv \mathbf{B}$.

For convenience, we may also perform the transformation purely at the level of Fokker-Planck equations. Let $h(\mathbf{x}, t)$ satisfy the general Fokker-Planck equation

$$\frac{\partial P(h, t)}{\partial t} = - \frac{\partial}{\partial h} [\tilde{A}(h, t) P(h, t)] + \frac{1}{2} \frac{\partial^2}{\partial h^2} [\tilde{B}(h, t) P(h, t)] \quad (2.53)$$

for scalar functions $\tilde{A}(h, t)$ and $\tilde{B}(h, t)$. Using the cyclic property of the trace in Eq.(2.51), we

may identify

$$\tilde{A} = \frac{\partial h}{\partial t} + (\nabla h)^T \mathbf{A} + \frac{1}{2} \text{Tr} [\mathbf{B} H_h(\mathbf{x})]. \quad (2.54)$$

Also, from Eq.(2.51),

$$\tilde{B} = [(\nabla h)^T \mathbf{G}] [(\nabla h)^T \mathbf{G}]^T = (\nabla h)^T \mathbf{B} (\nabla h). \quad (2.55)$$

Hence, using Eq.(2.53), Eq.(2.54) and Eq.(2.55), we may write down a Fokker-Planck equation for an arbitrary function of state in terms of $\mathbf{A}$ and $\mathbf{B}$.

AN SDE FOR HETEROPLASMY FORCED ONTO THE STEADY STATE LINE IN THE HIGH-CHURN LIMIT It has been demonstrated that SDE descriptions of stochastic systems which possess a globally-attracting line of steady states may be formed in the long-time limit by forcing the state variables onto the steady state line [215, 221]. Such descriptions may be formed in terms of a parameter which traces out the position on the steady state line, hence reducing a high-dimensional problem into a single dimension [215, 221]. In our case, heteroplasmy is a suitable parameter to trace out the position on the steady state line. We seek to use similar reasoning to verify Eq.(2.11). In what follows, we will assume that $\mathbf{x}(t = 0) = \mathbf{x}_{ss}$, where $\mathbf{x}_{ss}$ is the state which is the solution of $\mathbf{A} = \mathbf{0}$ (which is equivalent to finding the steady state solution of the deterministic rate equation in Eq.(2.16) due to our assumption of large copy numbers and $\Omega = 1$), so that we may neglect any deterministic transient dynamics.

Inspection of the steady state of the ODE description of our system reveals that the set of steady state solutions forms a line (see Eq.(2.36)–Eq.(2.38)). Inspection of the steady state solution reveals that the steady state depends on the fusion (γ) and fission (β) rates. Mitochondrial network dynamics occur on a much faster timescale than the replication and degradation of mtDNA: the former occurring on the timescale of minutes [164] whereas the latter is hours or days [43]. We seek to use this separation of timescales to arrive at a simple form for $\mathbb{V}(h)$. We redefine the fusion and fission rates such that

$$\begin{aligned} \gamma &\rightarrow M\gamma \\ \beta &\rightarrow M\beta \end{aligned} \quad (2.56)$$

where M is a constant which determines the magnitude of the fusion and fission rates, which we call the “network churn”.

We now wish to use heteroplasmy

$$h(\mathbf{x}, t) = h(\mathbf{x}) := (m_s + m_f)/(w_s + w_f + m_s + m_f), \quad (2.57)$$

as our choice for the function of state in the Fokker-Planck equation in Eq.(2.53). We will first compute the diffusion term $\tilde{B}$ for heteroplasmy using Eq.(2.55). If we constrain the state $\mathbf{x}$ to be forced onto the steady state line $\mathbf{x}_{ss}$ (as per [215, 221]) in the high-churn limit, then upon defining

$$\theta := \sqrt{b^2(\beta + \gamma\kappa)^2 - 2b\gamma\mu(\beta - \gamma\kappa + 2\gamma(\delta - 1)m_s) + \gamma^2\mu^2} \quad (2.58)$$

we have

$$\begin{aligned} \lim_{M \rightarrow \infty} \left(\tilde{B} \Big|_{\mathbf{x}=\mathbf{x}_{ss}} \right) = & (16b^2\gamma^2\mu m_s (\beta \\ & + \gamma(\delta - 1)m_s)^2 (b^2 (\beta^3 + 2\beta^2\gamma(\delta - 1)m_s + \beta\gamma^2 (\kappa^2 + (1 - 2\delta)\kappa m_s + \\ & (\delta - 1)(2\delta - 1)m_s^2) + \gamma^3\kappa m_s ((\delta - 1)m_s - \kappa)) \\ & + b (-\beta^2(2\gamma\mu + \theta) + \beta\gamma (\kappa(2\gamma\mu + \theta) + m_s(\gamma(5 - 4\delta)\mu - 2(\delta - 1)\theta)) + \\ & \gamma^2 m_s ((\delta - 1)m_s(3\gamma\mu + \theta) - \kappa(2\gamma\mu + \theta))) \\ & + \gamma\mu(\gamma\mu + \theta) (\beta - \gamma m_s))) / (\beta^3 (b (\gamma (\kappa - 2\delta m_s + 2m_s) - \beta) + \gamma\mu + \theta)^4). \end{aligned} \quad (2.59)$$

Eq.(2.59) is difficult to understand. In order to perform further simplification, we make an ansatz for the form of $\tilde{B}$ ($\tilde{B}_{An}$) and seek to determine whether our ansatz is equivalent to the derived form of $\tilde{B}$ under the constraints defined on the left-hand side of Eq.(2.59). Our ansatz takes the form

$$\tilde{B}_{An} := \lim_{M \rightarrow \infty} \left(\frac{2\mu h(1 - h)}{n(\mathbf{x})} \cdot f_s(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}_{ss}} \right) \quad (2.60)$$

where $f_s(\mathbf{x}) := (w_s + m_s)/(w_s + w_f + m_s + m_f)$ and $n(\mathbf{x}) := w_s + w_f + m_s + m_f$. Notice that this ansatz is more general than Eq.(2.59), since it has no explicit dependence upon the parameters of the control law assumed in Eq.(2.10), and only explicitly depends upon functions of state $\mathbf{x}$.

Upon substituting the steady state $\mathbf{x}_{ss}$ into the ansatz in Eq.(2.60) and taking the high-churn limit, we find that

$$\begin{aligned} \tilde{B}_{An} = & -(m_s (\beta + \gamma(\delta - 1)m_s) (b(\beta + \gamma\kappa) - \gamma\mu + \theta)(b(\beta + \gamma\kappa) + \gamma\mu - \theta) (b(m_s(2\beta\delta - \beta + \gamma\kappa) - 2\beta\kappa) \\ & + m_s(\theta - \gamma\mu))) / (2b\beta^3 (\kappa - \delta m_s + m_s)^2 (b(\gamma(\kappa - 2\delta m_s + 2m_s) - \beta) + \gamma\mu + \theta)). \end{aligned} \quad (2.61)$$

After some algebra (see GitHub repository for Mathematica notebook), it can be shown that

Eq.(2.59) and Eq.(2.61) are equivalent, i.e.

$$\tilde{B}_{\text{An}} \equiv \lim_{M \rightarrow \infty} \left(\tilde{B} \Big|_{\mathbf{x}=\mathbf{x}_{\text{ss}}} \right). \quad (2.62)$$

As such, we may use $\tilde{B}$ and $\tilde{B}_{\text{An}}$ interchangeably in the limit of high network churn. Furthermore, it can be shown after some algebra that the drift of heteroplasmy when forced onto the steady state line is 0, i.e.

$$\tilde{A} \Big|_{\mathbf{x}=\mathbf{x}_{\text{ss}}} \equiv 0. \quad (2.63)$$

A similar result is shown in [215] ((Equation S59) therein). Substituting $h(\mathbf{x}, t) = h$, $\tilde{A}$ and $\tilde{B}$ into the Fokker-Planck equation for an arbitrary function of state Eq.(2.53), we have

$$\frac{\partial P(h, t)}{\partial t} = \frac{1}{2} \frac{\partial^2}{\partial h^2} \left[\left(\frac{2\mu h(1-h)}{n(\mathbf{x})} \cdot f_s(\mathbf{x}) \right) \Big|_{\mathbf{x}=\mathbf{x}_{\text{ss}}(h)} P(h, t) \right] \quad (2.64)$$

which is equivalent to the following SDE for heteroplasmy

$$dh = \sqrt{\frac{2\mu h(1-h)f_s(\mathbf{x})}{n(\mathbf{x})}} \Big|_{\mathbf{x}=\mathbf{x}_{\text{ss}}(h)} dW \quad (2.65)$$

in the limit of large network churn, large copy numbers, and a second-order truncation of the Kramers-Moyal expansion. Note that, although the state has been forced onto the steady state, stochastic fluctuations mean that trajectories may move along the line of steady states, so the diffusion coefficient is not constant in general. We may calculate the new value of $\mathbf{x}_{\text{ss}}(h)$ for every displacement due to Wiener noise in h , and substitute into $f_s(\mathbf{x})$ and $n(\mathbf{x})$ to determine the diffusion coefficient at the next time step.

However, for sufficiently short times, and large copy numbers (i.e. low diffusivity of h), we may assume that the diffusion coefficient in Eq.(2.65) may be approximated as constant. Since the general solution of the SDE

$$dy = \sqrt{B} dW \quad (2.66)$$

for $B = \text{const}$ is

$$y \sim \mathcal{N}(y|y_0, Bt) \quad (2.67)$$

where $\mathcal{N}(y|y_0, \sigma^2)$ is a Gaussian distribution on y with mean y_0 and variance σ^2 , and $y_0 = y(t=0)$. Since we have assumed that the state is initialised at $\mathbf{x}(t=0) = \mathbf{x}_{\text{ss}}$, there are no

deterministic transient dynamics, so we may write

$$\mathbb{V}(h) \approx \frac{2\mu t}{n(\mathbf{x})} h(\mathbf{x})(1 - h(\mathbf{x})) f_s(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{x}_{ss}}, \quad (2.68)$$

where $\mathbb{V}$ returns variance of a random variable. Note that in this equation, we take $\mathbf{x} = \mathbf{x}_{ss} = \text{const}$, since we have assumed a low-diffusion limit. Also note that this equation is of precisely the same form as (Equation 12) of Ref. [43], except with an additional proportionality factor of f_s induced by the inclusion of a mitochondrial network.

2.3.4 THE INFLUENCE OF MITOCHONDRIAL DYNAMICS UPON HETEROPLASMY VARIANCE UNDER DIFFERENT MODELS OF GENETIC MTDNA CONTROL

To demonstrate the generality of Eq.(2.11), we explored several alternative forms of cellular mtDNA control (several of which may be found in [43], see Section A.2). We found that when copy number is controlled through the replication rate function (i.e. $\lambda = \lambda(\mathbf{x})$, $\mu = \text{const}$), when the fusion and fission rates were high and the fixation probability ($P(h = 0)$ or $P(h = 1)$) was negligible, Eq.(2.11) accurately described $\mathbb{V}(h)$ across all of the replication rates investigated, see Figure A.3A-F. The same mathematical argument to show Eq.(2.11) for the replication rate in Eq.(2.10) may be applied to these alternative replication rates where a closed-form solution for the deterministic steady state may be written down (see Section 2.3.3). Interestingly, when copy number is controlled through the degradation rate (i.e. $\lambda = \text{const}$, $\mu = \mu(\mathbf{x})$), heteroplasmy variance loses its dependence upon network state entirely and the f_s term is lost from Eq.(2.11) (see Eq.(A.15) and Figure A.3G-I). A similar mathematical argument was applied to reveal how this dependence is lost (see Section 2.3.3 and Section A.2).

In order to provide an intuitive account for why control in the replication rate, versus control in the degradation rate, determines whether or not heteroplasmy variance has network dependence, we investigated a time-rescaled form of the Moran process (see Section A.5). The Moran process is structurally much simpler than the model presented above, to the point of being unrealistic, in that the mitochondrial population size is constrained to be constant between consecutive time steps. Despite this, the modified Moran process proved to be insightful. We find that, when copy number is controlled through the replication rate, the absence of death in the fused subpopulation means the timescale of the system (being the time to the next death event) is proportional to f_s . In contrast, when copy number is controlled through the degradation rate, the presence of a constant birth rate in the entire population means the timescale of the system (being the time to the next birth event) is independent of f_s (see Sec-

tion A.5, Eq.(A.27)). Further discussion of Moran models in the context of *de novo* mutations is presented in Section 2.3.5.

2.3.5 CONTROL STRATEGIES AGAINST MUTANT EXPANSIONS

Thus far, we have argued that the rate of increase of heteroplasmy variance, and therefore the rate of accumulation of pathologically mutated cells within a tissue, increases with mitophagy rate (μ), decreases with total mtDNA copy number per cell (n) and increases with the fraction of unfused mitochondria (termed “singletons”, f_s), see Eq.(2.11). Below, we explore how biological modulation of these variables influences the accumulation of mutations. We use this new insight to propose three classes of strategy to control mutation accumulation and hence address associated issues in ageing and disease, and discuss these strategies through the lens of existing biological literature.

TARGETING NETWORK STATE

In Eq.(A.26) we argue that when population size is controlled in the birth rate, the inter-event rate (Γ) is effectively rescaled by the fraction of unfused mitochondria, i.e. $\Gamma = \mu n f_s$. We may use this idea to explore the role of the mitochondrial network in the accumulation of *de novo* mutations using an infinite sites Moran model [222] (see Figure 2.6A). Single cells were modelled over time as having a fixed mitochondrial copy number (n), and at each time step one mtDNA is randomly chosen for duplication and one (which can be the same) for removal. The individual replicated incurs Q *de novo* mutations, where Q is binomially distributed according to

$$Q \sim \text{Binomial}(L_{\text{mtDNA}}, \eta) \quad (2.69)$$

where $\text{Binomial}(N, p)$ is a binomial random variable with N trials and probability p of success. $L_{\text{mtDNA}} = 16569$ is the length of mtDNA in base pairs and $\eta = 5.6 \times 10^{-7}$ is the mutation rate per base pair per doubling [223]; hence each base pair is idealised to have an equal probability of mutation upon replication.

Figure 2.6B shows that in the infinite sites model, the consequence of Eq.(A.26) is that the rate of accumulation of mutations per cell reduces as the mitochondrial network becomes more fused, as does the mean number of mutations per mtDNA (Figure 2.6C). These observations are intuitive: since fusion serves to shield the population from mitophagy, mtDNA turnover slows down, and therefore there are fewer opportunities for replication errors to occur per unit

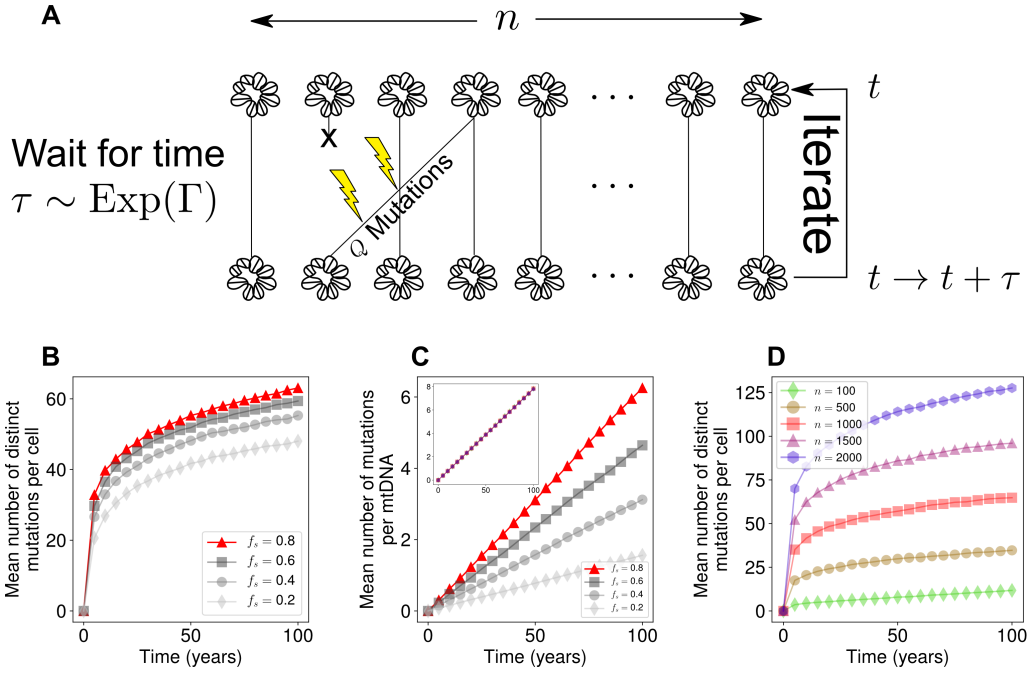


Figure 2.6: Rate of de novo mutation accumulation is sensitive to the network state/mitophagy rate and copy number for a time-rescaled infinite sites Moran model. (A) An infinite sites Moran model where Q mutations occur per Moran step (see Eq.(2.69)). (B-D) Influence of our proposed intervention strategies. (B) Mean number of distinct mutations increases with the fraction of unfused mitochondria. This corresponds to a simple rescaling of time, so all but one of the parametrisations are shown in grey. (C) The mean number of mutations per mtDNA also increases with the fraction of unfused mitochondria. Inset shows that the mean number of mutations per mtDNA is independent of the number of mtDNAs per cell; values of n are the same as in (D). (D) Mean number of mutations per cell increases according to the population size of mtDNAs. Standard error in the mean is too small to visualise, so error bars are neglected, given 10^3 realisations.

time. Different values of f_s in Figures 2.6B&C therefore correspond to a rescaling of time i.e. stretching of the time-axis.

Concurrently, a study by Chen *et al.* [224] observed the effect of deletion of two genes which are involved in mitochondrial fusion (*MFN1* and *MFN2*) in mouse skeletal muscle. Fragmentation of the mitochondrial network induced severe depletion of mtDNA copy number (as found in Figure 2.2D). The authors also observed that the number of mutations per base pair increased upon fragmentation, which is in agreement with the infinite sites model where fragmentation effectively results in a faster turnover of mtDNA (Figure 2.6F).

Our models predict that promoting mitochondrial fusion has a two-fold effect: firstly, it slows the increase of heteroplasmy variance (see Eq.(2.11) and Figure 2.4D); secondly, it reduces the rate of accumulation of distinct mutations (see Figure 2.6B&C). These two effects are both a consequence of mitochondrial fusion rescaling the time to the next turnover event, and therefore the rate of random genetic drift. As a consequence, we predict that promoting fusion in youth (assuming mean heteroplasmy is low) could slow down the accumulation and spread of mitochondrial mutations, and perhaps slow ageing. Indeed, the premature ageing phenotypes of the mitochondrial mutator mouse, which rapidly accumulates mtDNA mutations [52], can be rescued through endurance exercise with a resultant normalisation of mitochondrial morphology [225]. Since endurance exercise regulates mitochondrial network processes [226], it is possible that promotion of mitochondrial fusion contributes to the rejuvenation of the mutator mouse.

If we assume that fusion is selective in favour of wild-type mtDNAs, which appears to be the case at least for some mutations under therapeutic conditions [227, 228], we predict that a balance between fusion and fission is the most effective means of removing mutant mtDNAs (see below, Figure 2.7A), perhaps explaining why mitochondrial networks are often observed to exist as balanced between mitochondrial fusion and fission [229, 230]. In contrast, if selective mitophagy pathways are induced then promoting fragmentation is predicted to accelerate the clearance of mutants (see below, Figure 2.7B).

TARGETING MITOPHAGY RATE

Alterations in the mitophagy rate μ have a comparable effect to changes in f_s in terms of reducing the rate of heteroplasmy variance (see Eq.(2.11)) and the rate of *de novo* mutation (Figure 2.6B&C) since they both serve to rescale time. Our theory therefore suggests that inhibition of basal mitophagy in youth may be able to slow down the rate of random genetic drift, and perhaps healthy ageing, by locking-in low levels of heteroplasmy. Indeed, it has

been shown that mouse oocytes [231] as well as mouse hematopoietic stem cells [232] have comparatively low levels of mitophagy, which is consistent with the idea that these pluripotent cells attempt to minimise genetic drift by slowing down mtDNA turnover. A previous modelling study has also shown that mutation frequency increases with mitochondrial turnover [71].

Alternatively, it has also been shown that the presence of heteroplasmy, in genotypes which are healthy when present at 100%, can induce fitness disadvantages [66, 67, 69]. In cases where heteroplasmy itself is disadvantageous, especially in later life where such mutations may have already accumulated, accelerating heteroplasmy variance increase to achieve fixation of a species could be advantageous. However, this will not avoid cell-to-cell variability, and the physiological consequences for tissues of such mosaicism is unclear.

TARGETING COPY NUMBER

To investigate the role of mtDNA copy number (mtCN) on the accumulation of *de novo* mutations, we set $f_s = 1$ such that $\Gamma = \mu n$ (i.e. a standard Moran process). We found that varying mtCN did not affect the mean number of mutations per molecule of mtDNA (Figure 2.6C, inset). However, as the population size becomes larger, the total number of distinct mutations increases accordingly (Figure 2.6D). In contrast to our predictions, a recent study by Wachsmuth *et al.* [233] found a negative correlation between mtCN and the number of distinct mutations in skeletal muscle. However, Ref. [233] also found a correlation between the number of distinct mutations and age, in agreement with our model. Furthermore, the authors used partial regression to find that age was more explanatory than mtCN in explaining the number of distinct mutations, suggesting age as a confounding variable to the influence of copy number. Our work shows that, in addition to age and mtCN, turnover rate and network state also influence the proliferation of mtDNA mutations. Therefore, one would ideally account for these four variables for jointly, in order to fully constrain our model.

A study of single neurons in the substantia nigra of healthy human individuals found that mtCN increased with age [234]. Furthermore, mice engineered to accumulate mtDNA deletions through faulty mtDNA replication [52] display compensatory increases in mtCN [235], which potentially explains the ability of these animals to resist neurodegeneration. It is possible that the observed increase in mtCN in these two studies is an adaptive response to slow down random genetic drift (see Eq.(2.11)). In contrast, mtCN reduces with age in skeletal muscle [233], as well as in a number of other tissues such as pancreatic islets [236] and peripheral blood cells [237]. Given the beneficial effects of increased mtCN in neurons, long-term increases in mtCN

in youth could delay other age-related pathological phenotypes.

2.3.6 OPTIMAL MITOCHONDRIAL NETWORK CONFIGURATIONS FOR MITOCHONDRIAL QUALITY CONTROL

Whilst the above models of mtDNA dynamics are neutral (i.e. m and w share the same replication and degradation rates), it is often proposed that damaged mitochondria may experience a higher rate of degradation [238, 239]. There are two principal ways in which selection may occur on mutant species. Firstly, mutant mitochondria may be excluded preferentially from the mitochondrial network in a background of unbiased mitophagy. If this is the case, mutants would be unprotected from mitophagy for longer periods of time than wild-types, and therefore be at greater hazard of degradation. We can alter the fusion rate (γ) in the mutant analogues of Eq.(2.1), Eq.(2.2) and Eq.(2.7)–Eq.(2.9) by writing

$$\gamma \rightarrow \gamma/(1 + \epsilon_f) \quad (2.70)$$

for all fusion reactions involving 1 or more mutant mitochondria where $\epsilon_f > 0$. The second potential selective mechanism we consider is selective mitophagy. In this case, the degradation rate of mutant mitochondria is larger than wild-types, i.e. we modify the mutant degradation reaction to



for $\epsilon_m > 0$.

In these two settings, we explore how varying the fusion rate for a given selectivity (ϵ_f and ϵ_m) affects the extent of reduction in mean heteroplasmy. Figure 2.7A shows that, *in the context of selective fusion* ($\epsilon_f > 0$) *and non-selective mitophagy* ($\epsilon_m = 0$) *the optimal strategy for clearance of mutants is to have an intermediate fusion/fission ratio*. This was observed for all fusion selectivities investigated (see Figure 2.8). Intuitively, if the mitochondrial network is completely fused then, due to mitophagy only acting upon smaller mitochondrial units, mitophagy cannot occur – so mtDNA turnover ceases and heteroplasmy remains at its initial value. In contrast, if the mitochondrial network completely fissions, there is no mitochondrial network to allow the existence of a quality control mechanism: both mutants and wild-types possess the same probability per unit time of degradation, so mean heteroplasmy does not change. Since both extremes result in no clearance of mutants, the optimal strategy must be to have an intermediate fusion/fission ratio.

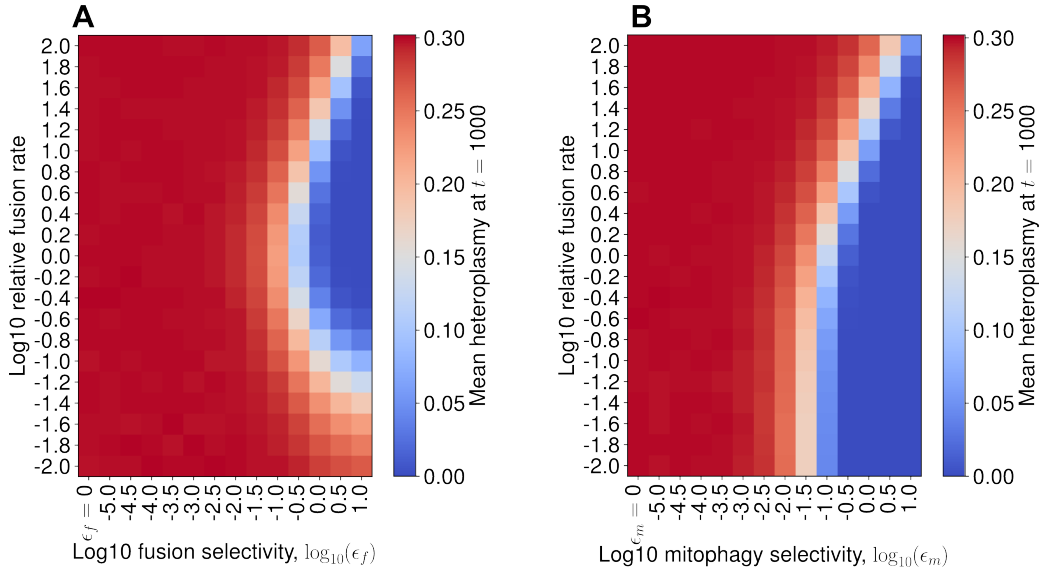


Figure 2.7: Selective fusion implies intermediate fusion rates are optimal for mutant clearance whereas selective mitophagy implies complete fission is optimal. Numerical exploration of the shift in mean heteroplasmy for varying fusion/fission ratio, across different selectivity strengths. Stochastic simulations for mean heteroplasmy, evaluated at 1000 days, with an initial condition of $h = 0.3$ and $n = 1000$; the state was initialised on the steady state line for the case of $\epsilon_f = \epsilon_m = 0$, for 10^4 iterations. **(A)** For selective fusion (see Eq.(2.70)), for each value of fusion selectivity (ϵ_f), the fusion rate (γ) was varied relative to the nominal parametrisation (see Table 2.1). When $\epsilon_f > 0$, the largest reduction in mean heteroplasmy occurs at intermediate values of the fusion rate; a deterministic treatment reveals this to be true for all fusion selectivities investigated (see Figure 2.8). **(B)** For selective mitophagy (see Eq.(2.71)), when mitophagy selectivity $\epsilon_m > 0$, a lower mean heteroplasmy is achieved, the lower the fusion rate (until mean heteroplasmy = 0 is achieved). Hence, complete fission is the optimal strategy for selective mitophagy.

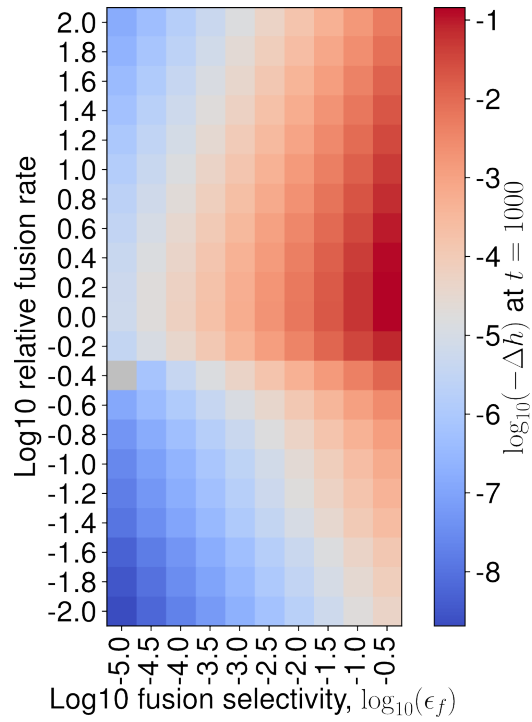


Figure 2.8: A deterministic parameter sweep of fusion selectivity and the relative fusion rate for Eq.(2.70), according to an ODE treatment. An ODE treatment allows smaller heteroplasmy changes to be probed without the need to resort to an infeasible number of stochastic simulations. Displaying the relative change in heteroplasmy (Δh) after $t = 1000$ days. We observe that a reduction in heteroplasmy is achieved at intermediate fusion rates at all non-zero fusion selectivities investigated. Grey denotes a change which is smaller than floating point precision.

In contrast, in Figure 2.7B, *in the context of non-selective fusion ($\epsilon_f = 0$) and selective mitophagy ($\epsilon_m > 0$), the optimal strategy for clearance of mutants is to completely fission the mitochondrial network*. Intuitively, if mitophagy is selective, then the more mtDNAs which exist in fragmented organelles, the greater the number of mtDNAs which are susceptible to selective mitophagy, the greater the total rate of selective mitophagy, the faster the clearance of mutants.

2.4 DISCUSSION

In this work, we have sought to unify our understanding of three aspects of mitochondrial physiology – the mitochondrial network state, mitophagy, and copy number – with genetic dynamics. The principal virtue of our modelling approach is its simplified nature, which makes general, analytic, quantitative insights available for the first time. Our models capture the central observation that fused mitochondria are at a lower susceptibility to degradation than unfused mitochondria, since mitophagy acts upon small mitochondrial fragments [164]. In using parsimonious models, we are able to make the first analytic link between the mitochondrial network state and heteroplasmy dynamics. This is in contrast to other computational studies in the field, whose structural complexity make analytic progress difficult, and accounting for their predicted phenomena correspondingly more challenging.

Using our approach we find that, for a wide class of models, the rate of linear increase of heteroplasmy variance is modulated in proportion to the fraction of unfused mitochondria (see Eq.(2.11)), which serves to rescale time. This central observation provides a substantial change in our understanding of mitochondrial genetics, as it suggests that the mitochondrial network state, in addition to mitochondrial turnover and copy number, must be accounted for in order to predict the rate of spread of mitochondrial mutations in a cellular population. Crucially, we find that the dynamics of heteroplasmy variance is independent of the absolute rate of fission-fusion events, since network dynamics occur approximately 10^3 times faster than mitochondrial turnover, inducing a separation of timescales. Since heteroplasmy variance depends purely on the absolute number of unfused mitochondria, this makes way for the possibility of quantifying the contribution of the mitochondrial network to heteroplasmy variance dynamics without the need to quantify the absolute rate of fission and fusion (for instance, the fraction of unfused mitochondria may be quantified from static confocal micrographs, whereas quantification of absolute fission-fusion rates would require more complex, time-resolved, data). By linking with classical statistical genetics, we find that the mitochondrial network also modulates the rate of

accumulation of *de novo* mutations, also due to the fraction of unfused mitochondria serving to rescale time. We find that, in the context of mitochondrial quality control through selective fusion, an intermediate ratio the fusion/fission ratio is optimal due to the finite selectivity of fusion. This latter observation perhaps provides an indication for the reason why we observe mitochondrial networks in an intermediate fusion state under physiological conditions [229, 230].

Our approach has, broadly speaking, been to consider neutral models of mtDNA genetic dynamics. It is, however, typically suggested that increasing the rate of mitophagy promotes mtDNA quality control, and therefore shrinks the distribution of heteroplasmies towards 0% mutant (see Eq.(2.70) and Eq.(2.71)). If this is the case, then a neutral genetic model appears to be inappropriate, as mutants experience a higher rate of degradation. Stimulation of the PINK1/Parkin pathway has been shown to select against deleterious mtDNA mutations *in vitro* [227] and *in vivo* [228], as has repression of the mTOR pathway via treatment with rapamycin [228, 240]. However, the necessity of performing a genetic/pharmacological intervention to clear mutations via this pathway suggests that the ability of tissues to selectively remove mitochondrial mutants under physiological conditions is weak. Consequently, neutral models such as our own are useful in understanding how the distribution of heteroplasmy evolves through time under physiological conditions. Indeed, it has been recently shown that mitophagy is basal [241] and can proceed independently of PINK1 *in vivo* [242], perhaps suggesting that mitophagy has non-selective aspects – although this is yet to be verified conclusively.

We have presented the case of copy number control in the replication rate as being a more intuitive model than control in the degradation rate. The former has the interpretation of biogenesis being varied to maintain a constant population size, with all mtDNAs possessing a characteristic lifetime. The latter has the interpretation of all mtDNA molecules being replicated with a constant probability per unit time, regardless of how large or small the population size is, and changes in mitophagy acting to regulate population size. Such a control strategy seems wasteful in the case of stochastic fluctuations resulting in a population size which is too large, and potentially slow if fluctuations result in a population size which is too small. Furthermore, control in the replication rate means that the mitochondrial network state may act as an additional axis for the cell to control heteroplasmy variance (Figure 2.4D) and the rate of accumulation of *de novo* mutations (Figure 2.6B&C). Single-mtDNA tracking through confocal microscopy in conjunction with mild mtDNA depletion could shed light on whether the probability of degradation per unit time per mtDNA varies when mtDNA copy number is perturbed, and therefore provide evidence for or against these two possible control strategies.

Our general approach reveals some apparent differences with previous studies which link mitochondrial genetics with network dynamics. Firstly, Tam *et al.* [206, 207] found that slower fission-fusion dynamics resulted in larger increases in heteroplasmy variance with time, in contrast to Eq.(2.11) which only depends on fragmentation state and not absolute network rates. The simulation approach of Tam *et al.* [206, 207] allowed for mitophagy to act on whole mitochondria, where mitochondria consist of multiple mtDNAs. Faster fission-fusion dynamics tended to form heteroplasmic mitochondria whereas slower dynamics formed homoplasmic mitochondria. It is intuitive that mitophagy of a homoplasmic mitochondrion induces a larger shift in heteroplasmy than mitophagy of a single mtDNA, hence slower network dynamics form more homoplasmic mitochondria. However, this apparent difference with our findings can naturally be resolved if we consider the regions in parameter space where the fission-fusion rate is much larger than the mitophagy rate, as is empirically observed to be the case [65, 216]. If the fission-fusion rates are sufficiently large to ensure heteroplasmic mitochondria, then further increasing the fission-fusion rate is unlikely to have an impact on heteroplasmy dynamics. Hence, this finding is potentially compatible with our study, although future experimental studies investigating intra-mitochondrial heteroplasmy would help constrain these models. Tam *et al.* [207] also found that fast fission-fusion rates could induce an increase in mean heteroplasmy [207], in contrast to Figure 2.3A which shows that mean heteroplasmy is constant with time after a small initial transient due to stochastic effects. We may speculate that the key difference between our treatment and that of Tam *et al.* [206, 207] is the inclusion of cellular subcompartments whereby fusion/fission induces migration between such compartments, which may induce dynamics in mean heteroplasmy. We note the uncertainty in accounting for the phenomena observed in such complex models highlights the virtues of a simplified approach which may yield interpretable laws and principles through analytic treatment.

The study of Mouli *et al.* [204] suggested that, in the context of selective fusion, higher fusion rates are optimal. This initially seems to contrast with our finding which states that intermediate fusion rates are optimal for the clearance of mutants (Figure 2.7A). However, the high fusion rates in that study do not correspond directly to the highly fused state in our study. Fission automatically follows fusion in [204], ensuring at least partial fragmentation, and the high fusion rates for which they identify optimal clearing are an order of magnitude lower than the top fusion rate they consider. In the case of complete fusion, mitophagy cannot occur in the model of [204], so there is no mechanism to remove dysfunctional mitochondria. It is perhaps more accurate to interpret the observations of Mouli *et al.* [204] as implying that selective fusion shifts the optimal fusion rate higher, when compared to the case of selective mitophagy alone.

Therefore the study of Mouli *et al.* [204] is compatible with Figure 2.7A. Furthermore, Mouli *et al.* [204] also found that when fusion is non-selective and mitophagy is selective, intermediate fusion rates are optimal whereas Figure 2.7B shows that complete fragmentation is optimal for clearance of mutants. Optimality of intermediate fusion in the context of selective mitophagy in the model of Mouli *et al.* [204] likely stems from two aspects of their model: i) mitochondria consist of several units which may or may not be functional; ii) the sigmoidal relationship between number of functional units per mitochondrion and mitochondrial ‘activity’ (the metric by which optimality is measured). Points (i) and (ii) imply that small numbers of dysfunctional mitochondrial units have very little impact on mitochondrial activity, so fusion may boost total mitochondrial activity in the context of small amounts of mutation. So whilst Figure 2.7B remains plausible in light of the study of Mouli *et al.* [204] if reduction of mean heteroplasmy is the objective of the cell, it is also plausible that non-linearities in mitochondrial output under cellular fusion [10] result in intermediate fusion being optimal in terms of energy output in the context of non-selective fusion and selective mitophagy. Future experimental studies quantifying the importance of selective mitophagy under physiological conditions would be beneficial for understanding heteroplasmy variance dynamics. We note again that the ubiquity of heteroplasmy [63, 166, 202] suggests that a neutral drift approach to mitochondrial genetics may be justified, which contrasts with the studies of Tam *et al.* [206, 207] and Mouli *et al.* [204] which focus purely on the selective effects of mitochondrial networks.

In order to fully test our model, further single-cell longitudinal studies are required. For instance, the study by Burgstaller *et al.* [72] found a linear increase in heteroplasmy variance through time in single oocytes. Our work here has shown that measurement of the network state, as well as turnover and copy number, are required to account for the rate of increase in heteroplasmy variance. Joint longitudinal measurements of f_s , μ and n , with heteroplasmy quantification, would allow verification of Eq.(2.11) and aid in determining the extent to which neutral genetic models are explanatory. This could be achieved, for instance, using the mito-QC mouse [241] which allows visualisation of mitophagy and mitochondrial architecture *in vivo*. Measurement of f_s , μ and n , followed by e.g. destructive single-cell whole-genome sequencing of mtDNA would allow validation of how μ , n and f_s influence $\mathbb{V}(h)$ and the rate of *de novo* mutation (see Figure 2.6). One difficulty is sequencing errors induced through e.g. PCR, which hampers our ability accurately measure mtDNA mutation within highly heterogeneous samples [243]. Morris *et al.* [63] have suggested that single cells are highly heterogeneous in mtDNA mutation, with each mitochondrion possessing 3.9 single-nucleotide variants on average. Error correction strategies during sequencing may pave the way towards high-accuracy mtDNA

sequencing [243, 244], and allow us to better constrain models of heteroplasmy dynamics.

3

MITOCHONDRIAL DNA DENSITY HOMEOSTASIS ACCOUNTS FOR A THRESHOLD EFFECT IN A CYBRID MODEL OF A HUMAN MITOCHONDRIAL DISEASE

3.1 SUMMARY

Mitochondrial dysfunction is involved in a wide array of devastating diseases but the heterogeneity and complexity of these diseases' symptoms challenges theoretical understanding of their causation. With the explosion of -omics data, we have the unprecedented opportunity to gain deep understanding of the biochemical mechanisms of mitochondrial dysfunction. This goal raises the outstanding need to make these complex datasets interpretable. Quantitative modelling allows us to translate such datasets into intuition and suggest rational biomedical treatments. Taking an interdisciplinary approach, we use a recently published large-scale dataset, and develop a descriptive and predictive mathematical model of progressive increase in mutant load of the MELAS 3243A>G mtDNA mutation. The experimentally observed behaviour is surprisingly rich, but we find that our simple, biophysically-motivated model intuitively accounts for this heterogeneity and yields a wealth of biological predictions. Our

findings suggest that cells attempt to maintain wild-type mtDNA density through cell volume reduction, and thus power demand reduction, until a minimum cell volume is reached. Thereafter, cells toggle from demand reduction to supply increase, upregulating energy production pathways. Our analysis provides further evidence for the physiological significance of mtDNA density, and emphasises the need for performing single-cell volume measurements jointly with mtDNA quantification. We propose novel experiments to verify the hypotheses made here, to further develop our understanding of the threshold effect, and connect with rational choices for mtDNA disease therapies.

3.2 INTRODUCTION

An important class of diseases which affect mitochondria are those which are directly caused by a mutation in mtDNA. The most common, and most studied, of these is MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes) syndrome, which is often associated with a mitochondrial tRNA mutation at position 3243A>G of the mitochondrial genome [245, 246]. Its incidence rate shows large regional variability, with prevalence of 1:6000 in Finland [247], to 1:424 in Australia [248]. tRNAs affected by the mutation cause amino-acid misincorporations during translation and generate defective mitochondrial proteins, inducing defective respiration when heteroplasmy is high [249].

A common feature in many diseases associated with mutations of mitochondrial DNA, including MELAS, is the non-linear physiological response of cells and tissues to increasing levels of mtDNA heteroplasmy. In particular, cells appear to be able to withstand high levels of heteroplasmy without showing any significant metabolic or physiological defect. For instance, fibroblasts possessing the MELAS mutation were shown to have unaffected respiratory enzyme activity until mutant load exceeded around 60% [175]. Also, Chomyn *et al.* [174] showed that oxygen consumption of cells does not significantly reduce until MELAS heteroplasmy exceeds ~90%. This observation has been named the threshold effect (reviewed in [35]).

It has been argued that the threshold effect occurs because mitochondria possess spare capacity at the translational, enzymatic and biochemical levels, which are each able to absorb some degree of stress, and thus delay the phenotypic response of increasing heteroplasmy, until a particular threshold heteroplasmy is exceeded, which is typically large [35]. Within this picture, each physiological feature (such as enzymatic activity or oxygen consumption), may be expected to display step-like behaviour with respect to increasing heteroplasmy. Asynchrony of thresholds between different features, such as 60% for enzyme activity [175] and 90% for

oxygen consumption [174], may be explained by spare capacity at intermediate levels. For example, a biochemical threshold effect could account for asynchrony in oxygen consumption and enzyme activity thresholds, where metabolic fluxes are altered to compensate for fewer functional enzymes and hence prolong an oxygen consumption deficit to higher levels of enzymatic dysfunction [35].

A recent study published by Picard *et al.* [37] established 143B TK⁻ cybrid osteosarcoma cell lines containing the MELAS 3243A>G mutation across the full dose response of mutant load. These cybrid cell lines were generated via the transfer of mtDNAs from a lymphoblastoid cell line, derived from a heteroplasmic 3243A>G patient, into 143B TK⁻ mtDNA-deficient cells. Although cybrid models such as this have had a number of non-physiological perturbations made to them, the advantage in using this system is that all cells possess the same nuclear background, and can be generated with heteroplasms across the full spectrum of mutant load. The authors measured a diverse array of features including RNA expression, protein expression, cell volume, growth rates, mitochondrial morphology and mtDNA content. The sheer diversity of data collected, across multiple levels of heteroplasmy, makes this an important dataset in understanding the threshold effect and mitochondrial dysfunction (despite the fact that the experiment consists of a single biological replicate, albeit with several technical replicates). Under the interpretation of the threshold effect presented above, one might expect a monotonic response to heteroplasmy, as spare capacity is depleted and the cell seeks alternative means of energy provision. However, Picard *et al.* observed complex multiphasic responses across numerous physiological readouts as heteroplasmy was increased [37].

The authors of that study identified four distinct transcriptional phases in the gene expression profiles of MELAS 3243A>G cells: 0%, 20-30%, 50-90% and 100% mutant load. They argue that continuous changes in heteroplasmy results in discrete changes in phenotype, because there exists a limited number of states that the nucleus can acquire in response to progressive changes in retrograde signalling [37]. In this work, by considering a distilled subset of the data from [37], and using simple, physically motivated arguments, we attempt to provide a simplified account of this dataset to gain better understanding of the consequences of this mutation and the threshold effect.

Our mathematical model suggests that cells attempt to maintain homeostasis in wild-type mtDNA density at low heteroplasms, through reduction of cell volume and therefore cellular power demand. We propose the existence of a single critical heteroplasmy, where cells are no longer able to maintain wild-type mtDNA density homeostasis, and toggle from power demand

reduction to supply increase. In this regime, energy supply pathways are upregulated. Our model also identifies an additional bioenergetic transition, in excess of 90% mutant load, as cells become fully homoplasmic. We explore the possibility of reduced transcriptional activity in mutant mtDNAs/mitochondria, limited tRNA diffusivity, and a connection between cellular proliferation rate and cell volume, finding all of the above to have explanatory power. We propose new experiments to verify the novel hypotheses made here, to drive forward understanding of the threshold effect.

3.3 MATERIALS AND METHODS

3.3.1 DATA NORMALISATION

We considered a core subset of 6 physiological features from the dataset of [37] for mathematical modelling: mitochondrially-encoded ETC mRNA pool size, ETC protein levels, glycolysis mRNA pool size, cell volume, cell proliferation rate (or cell growth rate as termed in [37]) and maximal respiratory capacity. We performed a number of normalisations to the data, the result of which is shown in Fig. 3.1.

OXPHOS, as well as glycolysis, have multiple different mRNAs participating in their respective pathways. To have some measure of the overall expression level of a pathway ($\bar{E}$), we used the mRNA concentration (in RPMK, reads per kilobase of transcript per million mapped reads), for each gene corresponding to enzymes of the pathway ($e_{i,k}(h)$, for gene i and technical replicate k at heteroplasmy h) and took a normalised sum as follows:

$$\bar{E} = \frac{1}{N} \frac{1}{n_r} \sum_{k=1}^{n_r} \sum_{i=1}^N \frac{e_{i,k}(h)}{1/n_r [\sum_{l=1}^{n_r} e_{i,l}(h=0)]} \quad (3.1)$$

where n_r = number of technical replicates, and N = number of genes in the pathway of interest. This quantity normalises the expression level of each gene to mean $h = 0$ levels, to avoid effects from consistently highly-expressed genes. The factor of $1/N$ results in $\bar{E}$ having the value of 1 at $h = 0$, so may be interpreted as a fold-change in expression relative to $h = 0$.

The standard error of $\bar{E}$ over technical replicates k ($s_{\bar{E}} = \mathbb{V}_k(\bar{E})^{1/2}$, where $\mathbb{V}_k(x_k)$ is the sample variance over x_k) may be written as:

$$s_{\bar{E}} = \frac{1}{N} \frac{1}{\sqrt{n_r}} \left(\mathbb{V}_k \left[\sum_{i=1}^N \frac{e_{i,k}(h)}{1/n_r [\sum_{l=1}^{n_r} e_{i,l}(h=0)]} \right] \right)^{1/2}. \quad (3.2)$$

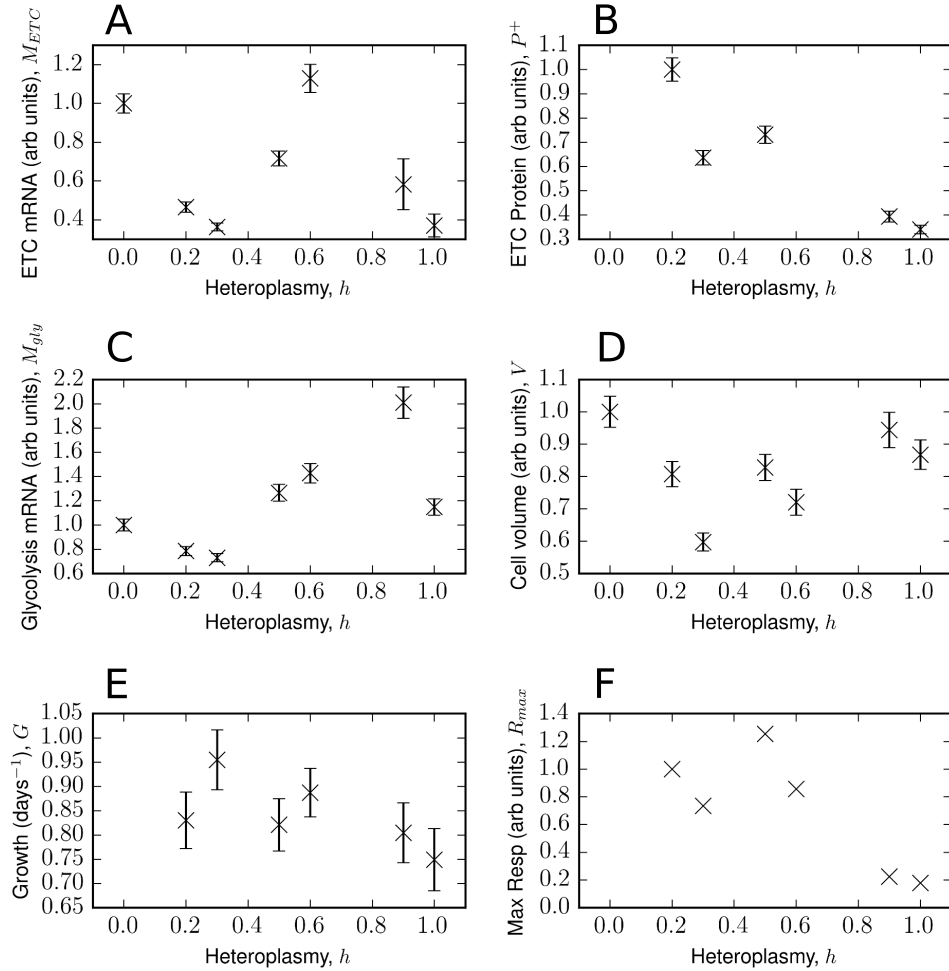


Figure 3.1: Multiphasic physiological response to increasing heteroplasmy – core data considered from Picard et al. renormalised to be in per-cell dimensions. Selected measurements of 143B TK⁻ osteosarcoma cells heteroplasmic in MELAS 3243A>G, from Ref. [37]. **(A)** mRNA levels for 11 mitochondrially-encoded electron transport chain subunits (*COX3*, *ND2*, *ND5*, *CYTB*, *ND3*, *ND6*, *COX1*, *ND4*, *COX2*, *ND1*, *ND4L*). **(B)** Protein levels for complexes I, III and IV. **(C)** Glycolysis mRNA levels, for genes (*PKM2*, *ENO1*, *PGAM4*, *PGK1*, *GAPDH*, *ALDOA*, *PFKP*, *GPI*, *HK2* and *SLC2A1*). All errors in (A-C) are the standard error of the transformed renormalised mean (see Eq.(3.2)). **(D)** Mean cell volume of an asynchronous population of growing cells $\pm$ SEM. **(E)** Growth rate, determined by linear regression. Error is the standard error in the slope from linear regression. **(F)** Maximum respiratory capacity. Details of the data normalisation and regression may be found in Section 3.3. See [37] for experimental protocols.

Eq.(3.1) and Eq.(3.2) are applied to glycolysis and ETC mRNAs in our main model, which yield dimensionless, normalised, measures of transcript levels for each biological pathway.

ATP synthase was excluded from both ETC mRNA and protein data, as it is expected to be regulated differently from other ETC proteins. This difference arises because mitochondrial membrane potential is required for cell growth [250], and glycolytic ATP may be used, even in cells without mtDNA, by ATP synthase to maintain membrane potential [251]. Thus, protein levels of ATP synthase may be expected to be regulated quite differently to those of the electron transport chain, and not generally indicative of respiratory activity.

For ETC protein, we simply used the sample mean of complexes I, III and IV, since the data given by Picard *et al.* [37] is already normalised.

3.3.2 DATA TRANSFORMATION TO PER-CELL DIMENSIONS

The data we considered of Picard *et al.* [37], consists of RNA-seq and Western blot measurements for mRNA and protein levels respectively. We wished to model the bioenergetic strategy of an average cell, so it is important that the data used to parametrise the model is of per-cell dimensions. We show in Section B.1 that it is appropriate to multiply protein and transcript data by cell volume to gain per-cell dimensions.

3.3.3 ERROR PROPAGATION

Our work focuses on describing mean behaviour with respect to heteroplasmy, so uncertainty in this mean must be quantified. For M_{gly} and M_{ETC} , we used error propagation on the normalised transcript levels $\bar{E}$ (see Eq.(3.1)) and V , to derive the volume-adjusted transcript uncertainties for the data in Fig. 3.1

$$\sqrt{\bar{E}^2 s_V^2 + V^2 s_{\bar{E}}^2} \quad (3.3)$$

where $s_{\bar{E}}$ is defined in Eq.(3.2) and s_V is the SEM for cell volume (raw data provided by Martin Picard). For the case of ETC protein data, since the corresponding experiments in Picard *et al.* [37] had only a single technical replicate, we derived an uncertainty by simply multiplying the normalised protein value (see Section 3.3.1) by s_V .

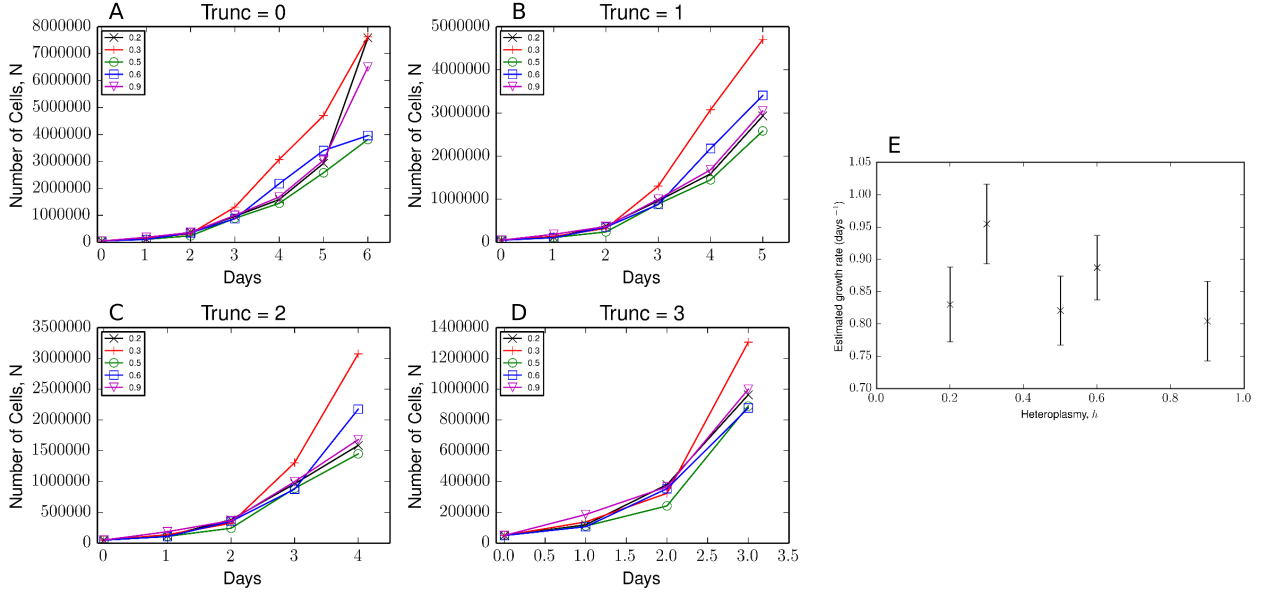


Figure 3.2: Cell proliferation data from Picard et al. (A-D) Number of cells (N) versus number of days of incubation, for different heteroplasmy levels, where a number of data points have been truncated (Trunc) from the right. Growth appears to be non-exponential by day 6, and is therefore removed subsequently. **(E)** Slope of linear regression with associated standard error, to derive the growth rate in dimensions of days⁻¹, as used in Fig. 3.1E. Raw data provided by Martin Picard.

3.3.4 GROWTH RATE DETERMINATION

The speed with which cells proliferate is dependent upon heteroplasmy, as can be seen in Fig. 3.2. However, by day 6 of growth, cell growth appeared to change its behaviour, with evidence of saturation; we therefore truncated the raw data to day 5 and calculated the exponential growth rate by linear regression in log-lin space.

3.3.5 MODEL INFERENCE

We used Bayesian inference to determine the parametric uncertainty in the mathematical model presented below, given the data in [37], for the 6 features considered here. We assumed Gaussian noise for the error in the normalised versions of each feature. For full details of the priors used for each feature, see Section B.2.

3.4 RESULTS

3.4.1 PER-CELL INTERPRETATION OF OMICS DATA HIGHLIGHTS MULTIPHASIC DYNAMICS IN RESPONSE TO HETEROPLASMIC LOAD

A simple interpretation of the threshold effect predicts the existence of spare capacity in the transcription, translation, enzyme complex and biochemical levels of the cell, in response to increasing heteroplasmy [35]. Under this interpretation of the threshold effect, we might expect all of these functions to have no more than one turning point with increasing heteroplasmy.

However, the data in Fig. 3.1, and indeed the dataset of Picard *et al.* [37] overall, shows a much more complex response. For instance, ETC transcripts clearly show two turning points, suggesting some kind of transient compensatory response. Across the features, these data also appear to be asynchronous in their turning points, for instance ETC transcripts peak at $h = 0.6$, but glycolysis transcripts peak at $h = 0.9$. This highlights the need for an extension in our understanding of the threshold effect, as well as the challenge in trying to parsimoniously model such a complex dataset.

We note that the measurement uncertainty, where reported in [37], for our selected features of interest are small relative to the variation with respect to heteroplasmy (see Fig. 3.1), justifying a non-linear fit to the data. It should be noted, however, that this uncertainty only reflects the technical variability in measurement, and does not include potential biological variability of these features. We use a Bayesian approach to appropriately account for this uncertainty, see Section B.2.

3.4.2 INTEGRATED OMICS DATA MOTIVATES A MODEL OF THE CAUSAL RELATIONSHIPS BETWEEN BIOENERGETIC VARIABLES

We present a qualitative description of our model in Fig. 3.3, which we will develop into a full quantitative description below. One of our central claims is the existence of a single transition in cellular behaviour, in response to increasing heteroplasmy of the 3243A>G mutation, over the 0-90% heteroplasmy range. *We propose that, at low heteroplasmy values, cells attempt to maintain homeostasis in wild-type mtDNA density by reducing their volume. This reduces biosynthetic and translational power demands, by the simplifying assumption that power demand scales directly with cell volume.*

Our model suggests that at a critical heteroplasmy, h^* , cells undergo a demand/supply toggle

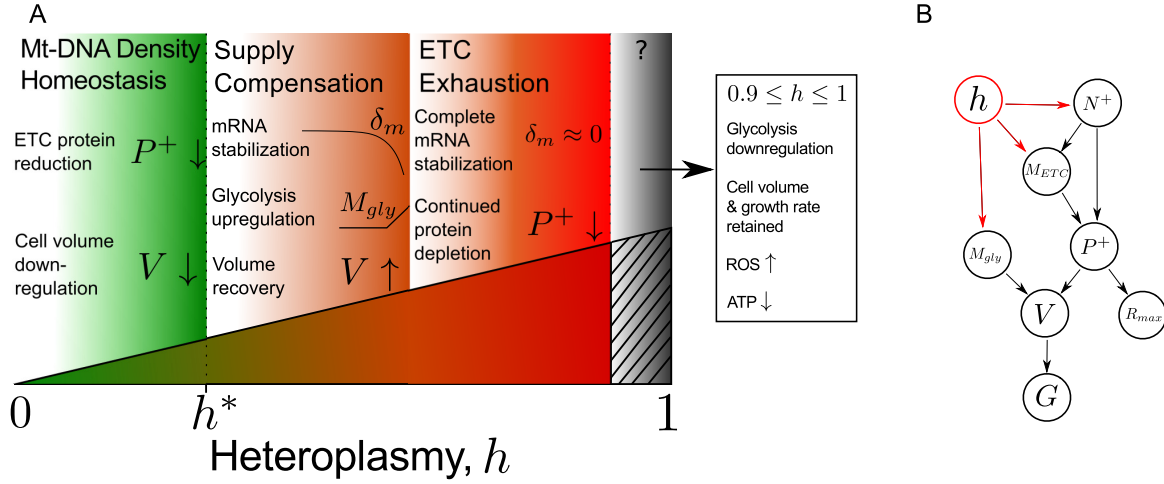


Figure 3.3: Qualitative description of continuous increase in MELAS mutant load. (A) At low levels of heteroplasmy (h), cells attempt to maintain homeostasis in wild-type mtDNA density by reducing their volume (V). This reduces power demands, allowing power supply/demand balance to be maintained despite rising heteroplasmy. Further increase in heteroplasmy triggers a power demand/supply toggle at a critical heteroplasmy h^* , where energy provision pathways are upregulated. This includes upregulation of both oxidative phosphorylation, by reduction in mRNA degradation (δ_m), and glycolysis transcripts (M_{gly}). Cell volume consequently recovers. Further increase in heteroplasmy exhausts ETC stabilisation, as $\delta_m \rightarrow 0$, and ETC protein (P^+) continues to deplete. In the transition to homoplasmy, glycolysis and ATP levels reduce, and yet cell volume and growth rate are retained. In this regime, the mode of energy production is unclear. (B) Flow of causality in our mathematical model. $N^+ =$ wild-type mtDNA copy number, $M_{ETC} =$ ETC mRNA, $R_{max} =$ maximum respiratory capacity, and $G =$ cellular growth rate. Heteroplasmy shown in red connecting to variables with explicit dependence.

where power supply is upregulated. Electron transport chain (ETC) transcripts are stabilised through reduced degradation, and glycolysis is increased. This bioenergetic compensatory behaviour at intermediate heteroplasmy allows cell volume to recover.

As heteroplasmy continues to increase, we claim that degradation of ETC transcripts becomes negligible. Thus, further increases in heteroplasmy results in reduction in ETC protein content and ETC exhaustion ensues.

These behaviours are captured in the mathematical framework of our model. However, as cells transition from 90% to 100% mutant mtDNA, another transition in cellular behaviour appears to occur, according to the data of [37]. Cells downregulate glycolysis, and yet retain cell volume and growth rate. The mode of energy production in this case is unclear, and opens new questions as to the most relevant energy supplies and demands in homoplasmic cells (see Section 3.4.5).

3.4.3 INTERACTIONS BETWEEN BIOENERGETIC VARIABLES CAN BE CAST AS A BOTTOM-UP QUANTITATIVE MODEL

We now present a quantitative description of our model, see Table 3.1, whose mechanistic interpretations will be more fully explored in Section 3.4.5. Our model attempts to unify the experimentally-measured features of [37] within a simple, physically plausible, bottom-up cell biological representation. We stress that our choice of model structure was not developed independently of the data in [37]; hence, at the level of choice of model structure, we have limited control of over-fit. However, the uncertainty in its parameters, given the data and a set of priors (see Section B.2), was computed using Bayesian inference. So whilst our parametrisation of the model has statistical control for uncertainty, we have not employed a statistical model selection framework. We believe this to be appropriate and practically unavoidable, as our objective is to yield a new reduced account for these heterogeneous data, present novel hypotheses and propose new experiments.

WILD-TYPE MTDNA SCALING A theme apparent in the data of Picard *et al.* is an overall downward trend of ETC mRNA and ETC protein with increasing heteroplasmy (h). We therefore use the hypothesis that these quantities scale with the amount of wild-type mtDNA (N^+). Total mtDNA copy number showed only a weak dependence on heteroplasmy, with only one heteroplasmy value displaying statistically significant deviation from $h = 0$ copy number (Fig. 11I of [37]). We therefore assume that $N = \text{mtDNA copy number} = \text{const}$ (set to 1 after normalisation, without loss of generality) which then defines N^+ , see Eq.(3.4). The successful performance of this simple model for N^+ is shown in Fig. 3.4. Note that this model has no free parameters, so was not associated with our inferred posteriors.

ETC mRNA Transcript copy number is determined by the balance of transcription (β) and degradation (δ_m) rates. Given our assumption of N^+ scaling, it can be shown (see Section B.3) that Eq.(3.5) may be used to model the ETC mRNA pool size (M_{ETC}). We further assume a constant mean transcription rate β for parsimony, and allow the degradation rate δ_m to vary with heteroplasmy in response to cellular signals. We require the degradation rate to be high for low heteroplasmies, and low at high heteroplasmies, to describe the ability of cells to upregulate their transcript copy number with rising heteroplasmy. A biologically-motivated choice of function which achieves this is a sigmoid, see Eq.(3.6), where k_{mRNA} , k_m and h_0 are constants.

Description	Equation	Equation number
Wild-type mtDNA, N^+	$N^+ = N(1 - h)$	(3.4)
ETC mRNA, M_{ETC}	$M_{ETC} = \frac{\beta}{\delta_m + 1} N^+$	(3.5)
	$\delta_m(h) = \frac{k_{mRNA}}{1 + \exp[k_m(h - h_0)]}$	(3.6)
ETC protein, P^+	$P^+ = \frac{N^+ M_{ETC}}{\delta_p}$	(3.7)
Glycolysis mRNA, M_{gly}	$M_{gly} = \begin{cases} c_1, & h \leq h^* \\ m_2 h + c_2, & h > h^* \end{cases}$	(3.8)
Cell volume, V	$\underbrace{k_o P^+ + k_g M_{gly}}_{\text{supply}} = \underbrace{V}_{\text{demand}}$	(3.9)
Growth rate, G	$G = \frac{k_{gr}}{V}$	(3.10)
Maximum respiratory capacity, $R_{\max}$	$R_{\max} = k_p P^+$	(3.11)

Table 3.1: Mathematical model of MELAS 3243 A>G mutation with progressive mutant load. See Table 3.2 for parameter descriptions. Heteroplasmy = h , mtDNA copy number = N .

Observables	Parameter	Description
M_{ETC} and M_{gly}	h^*	Critical heteroplasmy
M_{ETC}	k_{mRNA}	Maximum ETC mRNA degradation rate
	f_m	Relative ETC mRNA degradation rate at h^*
	k_m	ETC mRNA degradation steepness
	ζ	β/k_{mRNA} where β = ETC mRNA transcription rate
P^+	δ_p	ETC protein passive degradation rate
M_{gly}	c_1	Glycolysis mRNA pool size for $h \leq h^*$
	m_2	Glycolysis rate of increase with h , for $h > h^*$
V	k_o	ETC protein constant of proportionality
	k_g	Glycolysis mRNA constant of proportionality
G	k_{gr}	Constant of proportionality
R_{max}	k_p	Constant of proportionality

Table 3.2: Table of observables, and their corresponding parameters

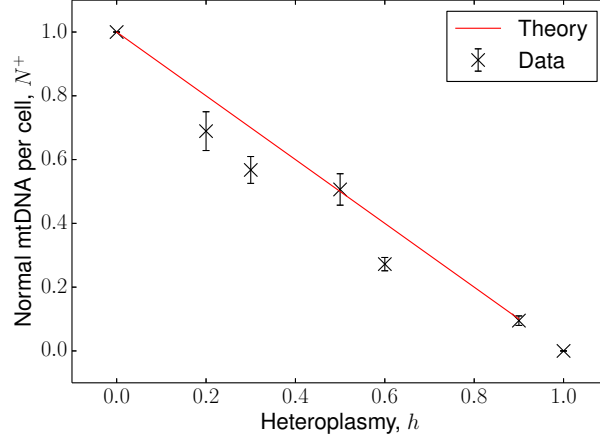


Figure 3.4: Model fit for wild-type mtDNA copy number Data for mtDNA copy number from Picard *et al.* [37] was multiplied by $(1 - h)$, as was the SEM. Displaying the model $N^+ = N(1 - h)$, for $N = \text{const} = 1$.

ETC PROTEIN It is intuitive to assume that mean protein levels scale with transcript levels, although this relationship may be noisy [252]. Following a similar assumption for ETC mRNA, we also assume that ETC protein (P^+) scales with wild-type mtDNA levels. Using analogous arguments to M_{ETC} (see Section B.3), we show that a reasonable model for ETC protein is Eq.(3.7), where $\delta_p = \text{const}$, denotes the baseline degradation of mitochondrial protein.

GLYCOLYSIS MRNA We assume that the glycolysis mRNA pool size (M_{gly}) is invariant to heteroplasmy, until a critical heteroplasmy h^* , where glycolysis is gradually upregulated as a result of cellular control. It is therefore parsimonious to assume that glycolysis regulation obeys a spline of a constant and linear model which toggles at h^* , see Eq.(3.8), where c_1 and m_2 are free parameters, and $c_2 = c_1 - m_2 h^*$, by continuity.

CELL VOLUME We propose that the power demands of the cell may be well approximated as scaling with cell volume, see Section B.3 for further discussion. As glycolysis and OXPHOS provide power supply to a first approximation, we assume that mean cell volume in an asynchronous population of cells (V) is effectively determined by a scaled sum of glycolysis and OXPHOS contributions, such that the cell obeys a power supply = power demand relationship, see Eq.(3.9), where $k_o, k_g = \text{const}$. We propose that changes in cellular volume are a long-term adaptation to heteroplasmy, and therefore power availability.

From Fig. 3.1, it is clear that this assumption fails at $h = 1$, where glycolysis levels and cell volume are comparable to $h = 0$ levels, and yet ETC proteins are only 30% of wild-type levels. As ATP levels are below wild-type levels in these cells (see Fig. S6E in [37]), the mode of energy

production is not clear and further metabolomic data may be required. We therefore exclude all $h = 1$ data, and limit the domain of our model to $0 \leq h \leq 0.9$.

GROWTH RATE We observe that the cellular proliferation rate (which we call growth rate, G , for consistency with [37]) varies with heteroplasmy (see Fig. 3.1). We hypothesise that there exists a relationship between mean cell volume and growth rate. It can be shown that, assuming individual cells increase their volume linearly through the cell cycle, growth rate varies inversely with mean cell volume (shown in Section B.3). This is shown in Eq.(3.10), where k_{gr} is a constant.

We show in Section B.3 that, under an exactly exponential model of cell growth, G is independent of V . However, given that there is presumably a wide class of cytoplasmic growth-vs-time profiles which cells may obey, we use a linear model as a parsimonious example of how cell growth may be connected to cytoplasmic volume.

MAXIMUM RESPIRATORY CAPACITY It has long been recognised that cells carrying the MELAS mutation experience a respiratory defect when heteroplasmic load exceeds approximately 90% [174, 253], and that this is due to a defect in protein synthesis [174, 249]. We therefore assume that maximum respiratory capacity ($R_{\max}$) is always determined by protein content. This yields a simple linear expression, see Eq.(3.11), where $k_p = \text{const.}$

MODEL SUMMARY In summary, our model of mean cellular behaviour with respect to heteroplasmy describes 7 features from Picard *et al.* [37] (N^+ , M_{ETC} , P^+ , M_{gly} , V , G , $R_{\max}$) and has 12 adjustable parameters (as discussed later this is fewer than the number required for 7 linear models with offsets), a table of which is shown in Table 3.2. In writing down this phenomenological model, we have attempted to account for a physiologically important subset of the data generated in [37], using bottom-up arguments wherever possible. In doing so, a number of novel, falsifiable, hypotheses are made.

3.4.4 PARAMETRISATIONS OF A SIMPLE BIOPHYSICAL MODEL ACCOUNT FOR COMPLEX OBSERVATIONS ACROSS A RANGE OF HETEROPLASMIC LOAD

The fit of the model described above is shown in Fig. 3.5. Between $0 \leq h \leq h^*$, h^* being the critical heteroplasmy where glycolysis is upregulated ($0.34 \leq h^* \leq 0.44$, 25-75% CI), our model reproduces the reduction in ETC transcript pool size. Similarly, we observe that ETC protein

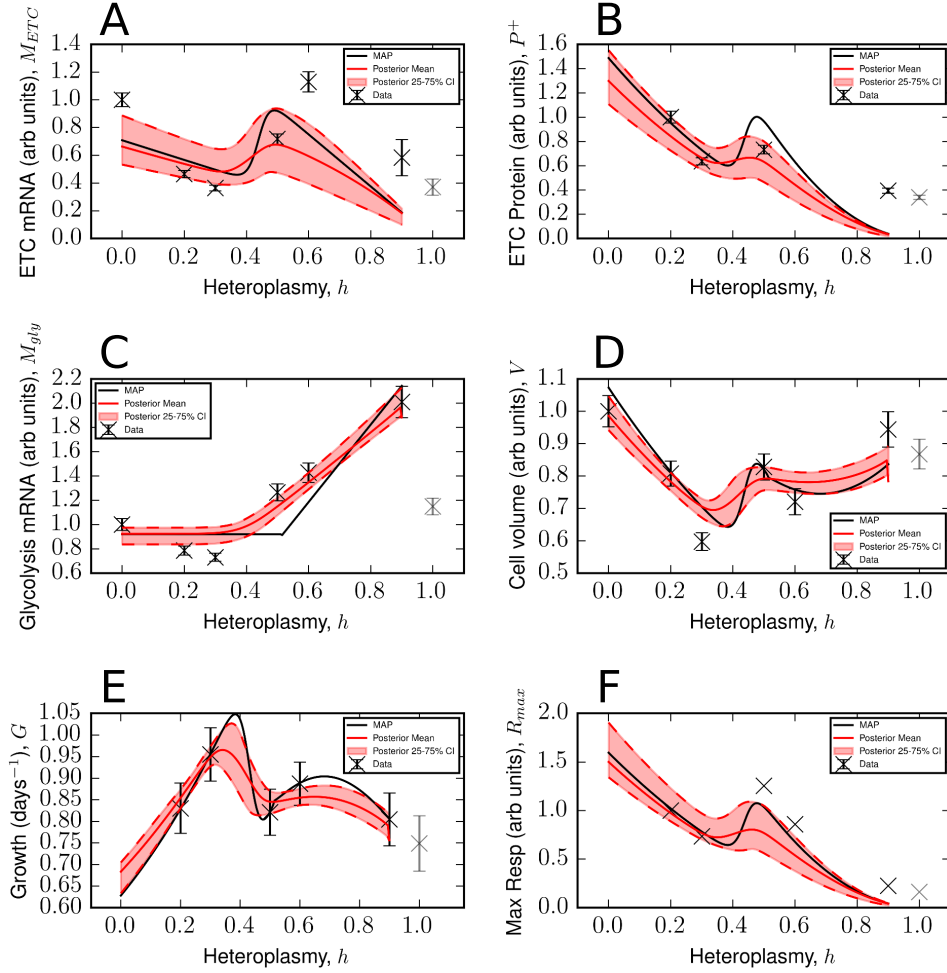


Figure 3.5: A simple biophysical model is consistent with complex observations across range of heteroplasmic load (A-F) Approximations for the maximum *a posteriori* estimate (black line), posterior mean (red line) and 25-75% Bayesian credible intervals (pink bands), for the model fits to selected data from [37]. The model makes predictions over the range $0 \leq h \leq 0.9$, see Section 3.4.3. Data for $h = 1$ have been plotted in grey, as they have not been used to train the model. Error bars are conservative and merely show the technical variability reported in [37], see Section 3.3.3.

pool size also reduces, as does cell volume and maximum respiratory capacity.

Our model successfully captures the transient compensatory responses in ETC mRNA, ETC protein and cell volume which begin around the critical heteroplasmy h^* . For heteroplasmies between $h^* \lesssim h \lesssim 0.5$, ETC mRNA degradation reduces causing ETC mRNA to be upregulated, along with ETC protein and maximum respiratory capacity. In this region, glycolysis becomes induced above wild-type levels, and cell volume can be observed to also recover.

In excess of $h \approx 0.5$, our model shows the observed reductions in ETC mRNA, ETC protein and maximum respiratory capacity. We see that continued upregulation of glycolysis mRNA allows cell volume to remain at an approximately constant value, although diminished relative

to a wild-type cell. Consequently, heteroplasmic cells between $0.2 \leq h \leq 0.9$ are predicted to proliferate at a faster rate than wild-type cells (see Fig. 3.5E).

3.4.5 KEY CLAIMS AND PREDICTIONS OF BIOPHYSICAL MODEL OF HETEROPLASMY

Here, we discuss the interpretations of our model in light of the mathematical description developed above, and explore the evidence for the biological insights it provides. We make experimental proposals to validate our claims, which are given in Section B.4. The set of mechanistic interpretations which follow from our mathematical model are:

- Wild-type mtDNA density is maintained homeostatically at low heteroplasmy;
- There exists a possible minimum cell volume which is approached at the critical heteroplasmy;
- Cells toggle from demand reduction (i.e. cell volume reduction) to supply increase (i.e. glycolysis and ETC mRNA upregulation), at the critical heteroplasmy;
- Mutant mtDNAs do not significantly contribute to the mitochondrial mRNA pool;
- Mitochondrial tRNAs remain moderately localised to their parent mtDNA;
- Maximum respiratory capacity is determined by ETC protein levels through a linear relationship;
- Cell growth rate is the reciprocal of mean volume, thus smaller cells proliferate faster.

WILD-TYPE MTDNA DENSITY HOMEOSTASIS IS MAINTAINED UNTIL A MINIMUM VOLUME IS REACHED NEAR THE CRITICAL HETEROPLASMY The parameter h^* determines the extent of mutant load, for which the cell begins to upregulate ETC mRNA and glycolysis mRNA. But what causes this change in behaviour, at this particular value of heteroplasmy? By examining the posteriors of our model fit (Fig. 3.5) we infer that cell volume takes its minimum value shortly before the most probable value of h^* (see Fig. 3.6). We hypothesise that an attempt to conserve wild-type mtDNA density (N^+/V) determines the position of h^* .

For $h \lesssim h^*$, wild-type mtDNA density is maintained despite increasing heteroplasmy, because cell volume diminishes. As a result of this reduced demand, the cell can tolerate diminished mitochondrial power supply. However, cell shrinkage cannot continue indefinitely and we hypothesise that the cell reaches a minimum cell volume at $h \approx h^*$. Once heteroplasmy exceeds

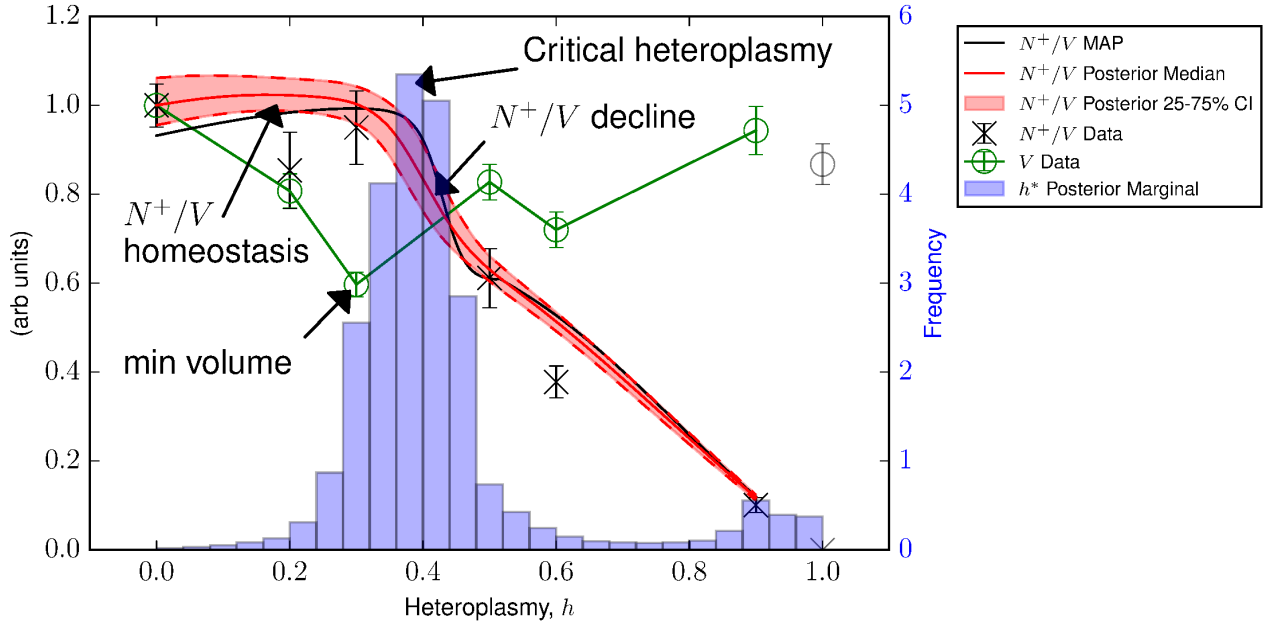


Figure 3.6: Wild-type mtDNA density (N^+/V) homeostasis may trigger supply/demand toggle Posterior statistics show an initial maintenance of N^+/V . When cell volume (V) takes its minimum value, the most probable value of h^* shortly follows. N^+/V then reduces. We suggest that the inability of the cell to maintain $N^+/V = \text{const}$, due to the existence of a minimum cell volume, causes cells to toggle in their strategy at h^* , from demand reduction to supply increase. Data from [37].

this value, the cell toggles its energy balance strategy from power demand reduction to supply increase, and the cell recovers in volume. We note that this observation was robust to removal of $h = 0.6$ data points where present, see Fig. 3.7.

There is evidence in the literature that wild-type mtDNA density is an important quantity. Bentlage and Attardi [253] observed that long-term culture of heteroplasmic MELAS cells resulted in an increase in mtDNA copy number, resulting in increased oxygen consumption. Whilst this was often accompanied by a decrease in heteroplasmy, some cell lines also exhibited this at constant heteroplasmy. This is consistent with the cell attempting to increase the absolute number of wild-type mtDNAs, perhaps to compensate for heteroplasmic load, and suggests that the absolute value of N^+ is a physiologically important quantity.

The density of mitochondrial content per unit cytoplasmic volume has been observed by many authors to be tightly regulated and physiologically predictive. The historical observations of Posakony *et al.* [44] showed that the mean ratio of mitochondrial content to cytoplasmic volume is kept relatively constant throughout the cell cycle in HeLa cells, occupying $\sim 10\text{-}11\%$ of cytoplasmic area throughout. Similar observations have been reproduced in more recent studies, in various other systems. Rafelski *et al.* found in budding yeast that mitochondrial content was proportional to bud size, and that all buds attain the same average ratio regardless of the

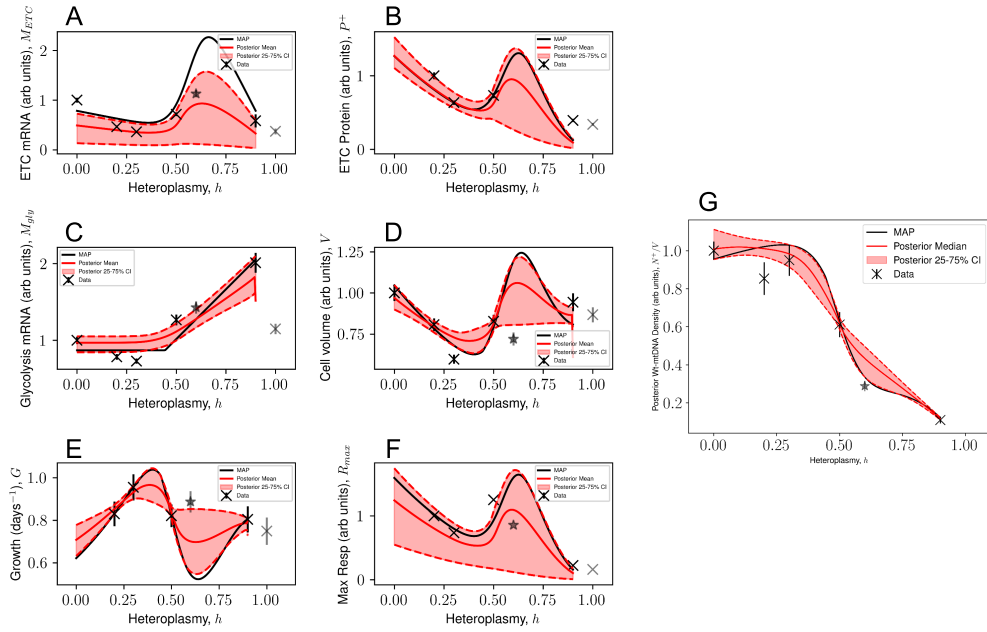


Figure 3.7: Model fits are largely robust to exclusion of data at 60% heteroplasmy. All data, if present, for $h = 0.6$ (grey stars) were excluded and the model was re-fitted as per Section B.2. **(A-F)** We observe the same qualitative behaviour as Fig. 3.5. M_{ETC} fits are expected to display a peak around $h = 0.6$ (see posterior mean), as is P^+ . M_{gly} is essentially unaffected, whereas V , G and R_{max} also show exaggerated behaviour relative to Fig. 3.5 at $h = 0.6$. **(G)** The mean posterior density of N^+/V is approximately constant between $0 \leq h \leq 0.3$, as seen in Fig. 3.6, suggesting that the hypothesis of wild-type density homeostasis is robust under the exclusion of this data point, given our model structure and priors.

mother's age or mitochondrial content [45], suggesting a stable scaling relation. Also, Johnston *et al.* [24] found that the density of mitochondrial mass was predictive of cell cycle dynamics, indicating that N/V (N = total number of mtDNAs) is physiologically relevant and potentially linked to cell power supply and growth dynamics. Indeed Jajoo, Paulsson and co-workers [49] found that the density of mitochondrial DNA tracks the quantity of cytoplasm inherited upon division in wild-type fission yeast. In Chapter 4, we will find that the assumption of constant mitochondrial density is explanatory in accounting for a recently published dataset [50], using mathematical modelling.

We may speculate as to the interpretation of a possible minimum cell volume. One straightforward interpretation is that a minimum cell volume corresponds to a mechanical constraint: a cell may only become so small because the machinery required to perform tissue-specific metabolic and structural tasks require a minimum amount of space. Alternatively, a bioenergetic cell volume may exist. For instance, appreciable power demands which scale with e.g. surface area or are invariant to cell size [254–256], may become large relative to power supply as cells become smaller. This is because cellular power supply (through glycolysis and oxidative phosphorylation) likely scales with cell volume. This mismatch in scaling behaviour between power supply and power demand may place a constraint on the minimum cell size.

ETC mRNA DEGRADATION DIMINISHES AT THE CRITICAL HETEROPLASMY CONTRIBUTING TO POWER DEMAND/SUPPLY TOGGLE The induction of glycolysis at the critical heteroplasmy is observed in our model by construction, see Eq.(3.8), since glycolysis is modelled to increase linearly when heteroplasmy exceeds this point. However, by observing the posterior distribution of the ETC mRNA degradation rate (see Fig. 3.8), we see that the critical heteroplasmy also coincides with the beginning of reduction in ETC transcript degradation with respect to heteroplasmy. Since ETC mRNA pool size varies with the inverse of this degradation rate (see Eq.(3.5)), ETC transcripts are consequently upregulated in tandem with glycolysis transcripts. This occurs until ETC degradation diminishes to negligible levels around $h \approx 0.5$, where this particular control mechanism becomes exhausted. Thus, the critical heteroplasmy coincides with a shift from power demand reduction, to supply increase from both glycolysis and OXPHOS contributions.

Since mtDNA is transcribed as a single polycistronic transcript [257], the relative stoichiometry of individual mRNA species must be controlled via active degradation. This is achieved by a balance between processes which stabilise and degrade mRNA [258]. The Picard data set can be explored further to seek corroborating evidence, by observing the ratio of ETC mRNA

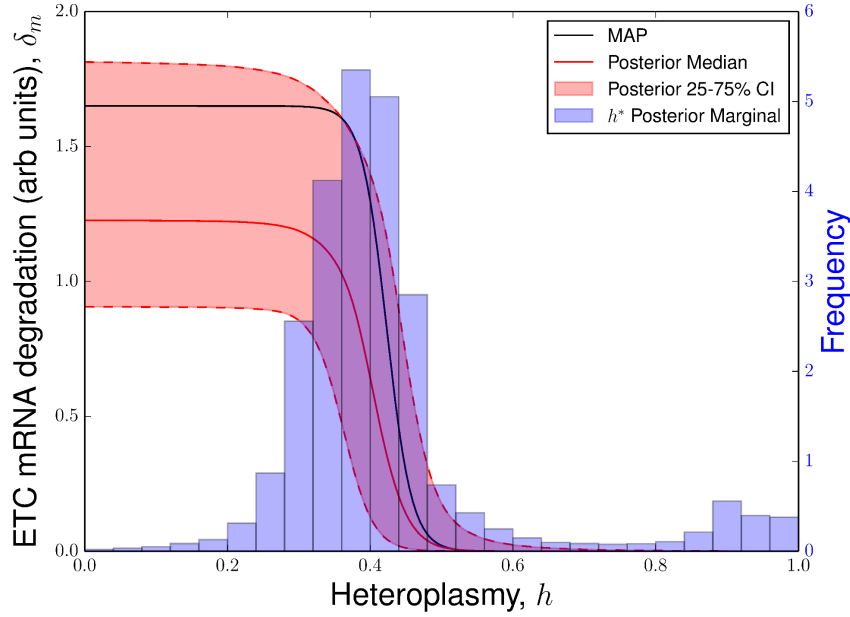


Figure 3.8: Critical heteroplasmy induces ETC mRNA stabilisation Posterior distributions for ETC mRNA degradation rate and critical heteroplasmy ($0.34 \leq h^* \leq 0.44$, 25-75% CI and $0.27 \leq h^* \leq 0.89$, 5-95% CI). It can be seen that the critical heteroplasmy coincides with the reduction in ETC degradation, signalling a power demand/supply toggle.

degraders to stabilisers. We find a qualitative similarity between this ratio (see Fig. 3.9) and the posterior distribution of the ETC degradation rate (see Fig. 3.8), both displaying a substantial reduction between $h = 0.3$ and $h = 0.5$.

MITOCHONDRIAL tRNAs ARE ENRICHED IN THE VICINITY OF THEIR CORRESPONDING PARENTAL mtDNA The co-location for redox regulation (CoRR) hypothesis states that mtDNA exists in the close proximity of respiratory units for the rapid and direct regulatory control of respiration [120, 121]. This hypothesis, which aligns with several leading theories on the origins of complex life [259, 260], has recently been supported by studying mitochondrial gene loss throughout the eukaryotic tree of life [261], as well as other experimental evidence reviewed in [121].

Given the CORR hypothesis, one might expect there to exist a functional link between the genotype of an mtDNA and the phenotype of its spatially-closest respiratory units. In other words, mitochondrial nucleoids (protein-DNA complexes which are thought to possess ~ 1.1 mtDNA molecules per nucleoid [181]) may each have a limited sphere of influence; this has been put forward as the ‘leaky-link’ hypothesis by other authors [119, 262].

The 3243A>G MELAS mutation causes amino-acid misincorporations, resulting in unstable

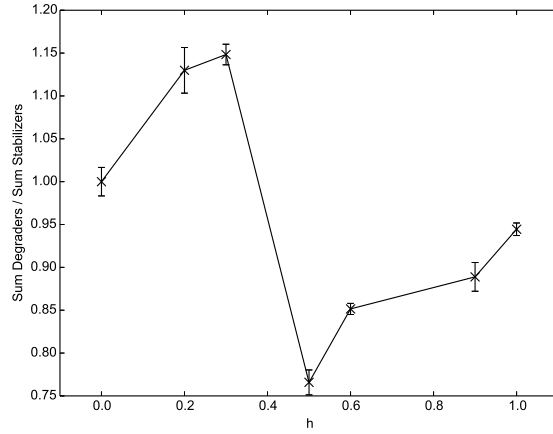


Figure 3.9: Corroborating evidence for ETC mRNA degradation from Picard et al. Ratio of mitochondrial mRNA degraders (*PDE12*, *PNPT1*, *SUPV3L1*) to stabilisers (*MTPAP*, *LRPPRC*, *SLIRP*), see Ref. [258] and references therein. We observe qualitative similarity in δ_m (see Fig. 3.8) and the ratio of normalised genes, both showing strong downregulation between $h = 0.3$ and $h = 0.5$. The numerator and denominator were normalised according to Eq.(3.1). Errors result from error propagation of a ratio, where the error for the numerator and denominator are derived using Eq.(3.2).

translation products [249]. It is reasonable to assume that mutant mtDNAs have a locally higher concentration of mutant tRNAs, since they are the source of these molecules, as well as the presence of cristae which may limit diffusion in the mitochondrial matrix [263]. ETC mRNAs in the vicinity of mutant mtDNAs will more often be translated into unstable protein due to the presence of defective tRNAs. Therefore, mutant mtDNAs may give a weaker contribution to the total functional ETC protein content of the cell than wild-type mtDNAs. Hence, low diffusivity inside the mitochondrial matrix is hypothesised to induce a local phenotype-genotype link [119, 262]. We take the limit of mutant mtDNAs providing zero contribution to the total ETC protein content of the cell, rather than a limited contribution, for parsimony – i.e. $P^+ \propto M_{\text{ETC}}N^+$. This is equivalent to assuming a strong local phenotype-genotype link. See Section B.5 for an alternative model of well-mixed tRNAs.

Evidence in the literature for this claim is mixed. It has been observed that mitochondrial mRNAs localise to mtDNA suggesting that mtDNA may be a site for mitochondrial translation [46, 186]. However, cybrid experiments by Ono *et al.* involving homoplasmic tRNA mutants 3243A>G and 4269A>G are able to recover their respiratory function by fusing such cells together to form hybrids [170]. Their recovery is presumably due to the diffusion of the healthy form of each tRNA, so that normal proteins may be translated. Yet, one might expect there to exist, at least limited, local coupling between mtDNA and their gene products under the CoRR hypothesis [120, 121, 261] and the leaky-link hypothesis [119, 262], as described above. It is noteworthy that the cybrid experiments by Ono *et al.* show recovered respiration after

a long adaptation phase of 10 days [170]. Since this is a much larger timescale than that of mtDNA replication, which is typically hours [264], it is conceivable that some kind of long-term adaptation occurred in this system which is still compatible with a strong local phenotype-genotype link; for example, the presence of nucleoids with an mtDNA of each kind of mutant, allowing the local transcription of both of the corresponding wild-type tRNAs. See Section B.4 and Table B.1 for experimental suggestions to determine the extent of tRNA diffusivity.

MUTANT MTDNAs HAVE A TRANSCRIPTIONAL DEFECT It has been hypothesised that, due to the limited diffusion of gene products, mutant mtDNAs may experience a local energy deficiency [119]. This local energy deficiency could conceivably reduce the ability of mutant mtDNAs to perform transcription through a local depletion of ATP. Since the concentration of ATP inside the mitochondrial matrix is low in the physiological setting [265], this may leave processes in the vicinity of mutant mtDNAs with increased susceptibility to local fluctuations in energy availability. If mutated mtDNAs are transcribed more slowly, they may have a lower contribution to the ETC transcript pool.

Alternatively, some other mechanism may generate a functional dependence between mitochondrial transcription and mtDNA genotype. For instance, mitochondrial transcription machinery (such as mitochondrial RNA polymerase, *POLRMT*) could be preferentially transported to functional mtDNAs, potentially via membrane potential sensing. Since dysfunctional mitochondria are observed to bud off from fused networks [164], it is possible that this mechanism allows the cell to sense the individual performance of mitochondria/mtDNAs [10], and hence target nuclear proteins accordingly. It is also possible that mitochondrial translational machinery could be preferentially targeted to functional mitochondria, which would also result in $P^+ \sim N^+$.

We adopted a model of no transcription for mutant mtDNAs, as opposed to limited transcription, for parsimony – i.e. $M_{\text{ETC}} \sim N^+$. See Section B.5 for an alternative model where mutant mtDNAs provide a non-zero contribution to M_{ETC} . We provide experimental suggestions for determining the extent of mutant mtDNA transcription in Section B.4.

CELL VOLUME IS NOT EXPLAINED BY CELL CYCLE VARIATIONS Our model predicts that cells, on average, change their size as heteroplasmy is varied, due to variation in power supply from OXPHOS and glycolysis. However, since cells vary their volume by a factor of 2 throughout the cell cycle, it is possible that cells with different heteroplasmies spend different durations at various stages of the cell cycle, explaining the observed variation in expected cell volume

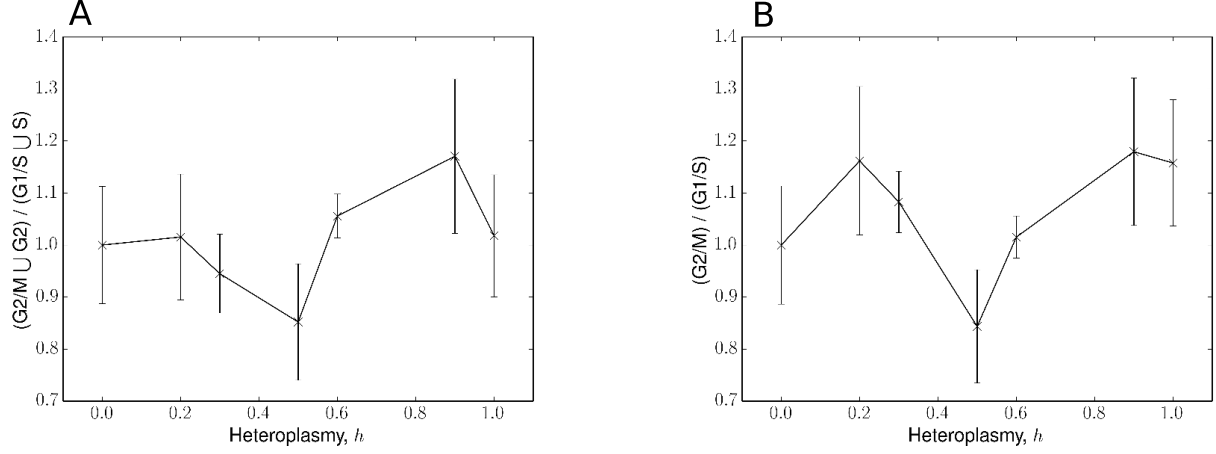


Figure 3.10: Variation of cell cycle markers with heteroplasmy A list of cell cycle markers, taken from [266] were normalised according to Eq.(3.1), yielding gene lists for G1/S, S, G2 and G2/M phases. **(A)** The ratio of G2/M to G1/S genes yielded no obvious trend with heteroplasmy. **(B)** G2/M and G2 gene lists were combined ($G2/M \cup G2$), as were G1/S and S ($G1/S \cup S$). This again yielded no obvious trend with heteroplasmy. Errors result from error propagation of a ratio, where the error for the numerator and denominator are derived using Eq.(3.2).

with heteroplasmy (see Fig. 3.1D). We sought evidence for this hypothesis by computing the ratio of the expression level for genes associated different stages of the cell cycle [266], see Fig. 3.10. However, we found little evidence to support the enrichment of cell cycle markers at any particular level of heteroplasmy.

OXPHOS CONTRIBUTIONS TO POWER SUPPLY ARE STABILISED AT THE CRITICAL HETEROPLASMY The relative contribution of OXPHOS to power supply, i.e. $k_o P^+ / (k_o P^+ + k_g M_{gly})$, is also interesting to observe as heteroplasmy is varied. We observed a transient stabilisation in OXPHOS contributions around h^* . A discussion of this is presented in Section B.6.

CELLS PROLIFERATE INVERSELY WITH THEIR SIZE Due to our reciprocal model connecting cell volume and growth (see Eq.(3.10)), our model suggests that wild-type cells proliferate more slowly relative to heteroplasmic cells due to their larger size.

MAXIMUM RESPIRATORY CAPACITY LINEARLY TRACKS ETC PROTEIN CONTENT It has long been suggested that cells above a particular threshold heteroplasmy experience a respiratory defect [174, 175, 253]. With a more classical interpretation of the threshold effect, we might have expected the need for a model which has switching behaviour in excess of 60% heteroplasmy [175] for maximum respiratory capacity, in analogy with glycolysis transcript levels (see Eq.(3.8)). However, in our model, we found that a simple linear relationship between ETC

protein and maximum respiratory capacity was sufficient to describe the data available (see Eq.(3.11)).

REACTIVE OXYGEN SPECIES MAY EXPLAIN THE TRANSITION TO HOMOPLASMY, BUT MODE OF ENERGY PRODUCTION REMAINS UNCLEAR In Eq.(3.9), we claim that cell volume is determined by the weighted sum of glycolysis transcripts and ETC protein. Over the range $0.9 < h \leq 1$, glycolysis transcripts reduce by 57%, whereas ETC protein and cell volume remain comparable, thus breaking the supply = demand relationship, as we have modelled it. Consequently, our model fails to describe the transition from $h = 0.9 \rightarrow 1$.

A potential explanation for the reduction in glycolysis transcripts over this range, comes from the fact that glycolysis provides substrate for oxidative phosphorylation. Damaged electron transport chain proteins may produce an excess of reactive oxygen species (ROS) [267], which can damage mitochondrial proteins, DNA and membranes. If, at high heteroplasmy, any flux through the electron transport chain causes high levels of ROS, then cells may attempt to reduce flux through glycolysis, to avoid production of these species.

Some evidence from Picard *et al.* supports this hypothesis, where superoxide dismutase (SOD) activity is largely constant with heteroplasmy, except for homoplasmic mutant MELAS cells, which have $\sim 20\%$ higher SOD activity than wild-type cells (see Fig. S7D of [37]). Furthermore, it is known that ROS can reversibly inhibit the activity of GAPDH, one of the enzymes involved in glycolysis [268, 269].

However, given that fatty acid oxidation (see Fig. 3.11) is strongly downregulated over this range, it remains unexplained how homoplasmic mutant cells maintain their cell volume (and growth rate), given their reduced reliance upon mitochondrial and glycolytic metabolism. Further metabolomic measurements may be required to uncover this mode of energy production.

ATP levels are also observed to decrease over this range (see Fig. S6E of [37]), which may even suggest that an alternative fuel currency besides ATP supports the growth and size of these cells. However, more careful investigation of this observation may be of value, since it is important to draw the distinction between ATP pool sizes and ATP fluxes, the latter perhaps being more indicative of ATP usage, and the former being indicative of only relative production/consumption rates.

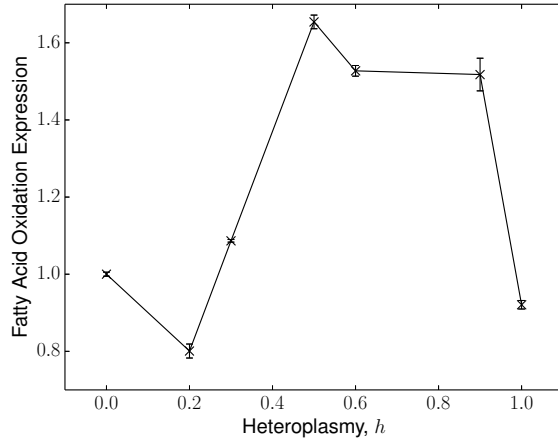


Figure 3.11: Expression of fatty acid oxidation enzymes with heteroplasmy. Showing variation of the genes (*ACADVL*, *ECHS1*, *HADH*, *ACAA2*) with heteroplasmy, normalised according to Eq.(3.1). It can be seen that these metabolites are downregulated between $h = 0.9 \rightarrow 1$, so fatty acid oxidation does not appear to be supporting the maintained cell volume and growth rates over this range of heteroplasmy. Errors are calculated using Eq.(3.2).

3.5 DISCUSSION

In this study, through use of a distilled subset of data from [37] and using minimal arguments, we have attempted to explore the apparent marked difference between the complex multiphasic observations of Picard *et al.*, and the classical step-like models associated with the threshold effect.

We note that the model presented here relies on data from osteosarcoma (143B TK⁻) cybrid cells [37]. MELAS is thought to affect muscle and the central nervous system most severely: i.e. quiescent, differentiated tissues. In contrast, 143B TK⁻ cells are highly proliferative. Hence, it is not obvious that the physiological consequences of MELAS heteroplasmy in the 143B TK⁻ system (for example the marked variation in cell volume, see Fig. 3.1D) is relevant for understanding the pathophysiology of MELAS (where variations in cell volume of tissue biopsies are not widely reported, although primary fibroblasts do have reported cell volume variation [270]). Whilst future studies in more pathologically-relevant nuclear backgrounds may be insightful, there is still potential value in understanding the physiological impacts of heteroplasmy in any given nuclear background, as it is plausible that at least some of our hypotheses may generalise to, and be testable in, a broader class of cellular systems.

Furthermore, cybrid cell lines are often unstable in the number of chromosomes that they possess, and the cybridization process itself may induce transcriptional [271] and metabolic changes [272]. A technical concern one may raise with the dataset of Picard *et al.* [37] is

therefore the extent to which their observations may be explained by random genetic drift in the nucleus. However, Picard *et al.* points out that the observations made with respect to heteroplasmy (at least for a subset of features) appear to be non-random with respect to increasing heteroplasmy: for example, the expression of heat shock proteins appeared to increase monotonically with heteroplasmy [37], which would not be expected under unbiased random perturbations at each level of heteroplasmy. We note also that cybrid systems are not without merit, as cells have the same nuclear background (although this background may be volatile), and have furthered progress in understanding mitochondrial dysfunction (reviewed in [273]). As is the case generally with model biological systems, it is important to seek corroborating evidence from multiple model systems before accepting the observations from any single experiment. Hence we have suggested a range of experiments to validate the hypotheses we have generated here.

We have argued that a single critical heteroplasmy, h^* , is sufficient to explain this subset of data over the heteroplasmy range $0 \leq h \leq 0.9$ and that other multiphasic behaviour arises naturally from the simple physical/biological assumptions of our model. Our model suggests that cells undergo a power demand/supply toggle at h^* , from demand reduction to supply increase. We hypothesise that homeostasis in wild-type mtDNA density is maintained via cell volume reduction, ensuring that the available functioning power sources are matched to a corresponding level of cellular demand, until a minimum cell volume is reached which coincides with h^* . This triggers the demand/supply bioenergetic toggle where energy production pathways are upregulated. We believe this re-emphasises the need for quantification of single-cell mtDNA content to be associated with volume measurements of the same cell: mtDNA density is a relevant physiological variable [24, 44, 45, 49, 274]. We find that the mode of energy production over the range $0.9 \leq h \leq 1$ is unclear, and that further metabolomic investigations may be required to determine this.

Our model further predicts that mutant mtDNAs have a reduced contribution to transcription and that tRNAs have low diffusivity. These hypotheses potentially follow from the ‘leaky-link’ hypothesis whereby mtDNAs have a limited extent of influence, such that there exists a phenotype-genotype link between an mtDNA and its spatially closest respiratory units [119, 262]. We also predict that a relationship exists between mean cell volume and cell growth.

A potential consequence of our predictions is that controlling either mtDNA copy number, wild-type mtDNA copy number, or wild-type mtDNA copy number density, to ensure optimal values of wild-type mtDNA copy number density could be valuable control axes in therapy.

Increasing mitochondrial DNA copy number, for instance through activation of the PGC-1 α pathway, may facilitate the increase of cell volume, deferring the critical heteroplasmy to higher values by delaying the approach towards a minimum cell volume. We might reason that this enhances a wild-type phenotype at higher heteroplasmy values, potentially deferring the full MELAS phenotype to higher heteroplasmy values, which typically appears between ~ 50 -90% mutant load [275]. Indeed, it has been found that increasing mitochondrial biogenesis can ameliorate mitochondrial myopathy in vivo [276].

We might also argue that as cells toggle from power demand reduction to supply increase, further bolstering of this compensatory response may have clinical significance. For instance, since we observe that cells switch to glycolytic metabolism to compensate for diminishing mitochondrial power supply, further encouragement of this energy mode may be therapeutic. This is supported by the recent observation that promoting the hypoxia response is protective against multiple forms of respiratory chain inhibition [277]. Alternatively, since we predict that cells innately downregulate ETC mRNA degradation, seeking to upregulate mitochondrial transcription may aid the cell in maintaining a sufficient mRNA pool size. Furthermore, promoting alternative energy production pathways such as fatty acid oxidation via the ketogenic diet may also aid in reducing the dependence on oxidative phosphorylation. This diet has been associated with increased mitochondrial transcripts [278], mitochondrial content [278, 279] and has been shown to slow mitochondrial myopathy progression in transgenic Deletor mice [279]. Indeed, the diet has recently been used in clinic as an adjunctive therapy for a patient suffering from MELAS, harbouring the 3260A>G mutation, which successfully decreased the frequency of seizures and stroke-like episodes [280].

4

MITOCHONDRIAL HETEROGENEITY, METABOLIC SCALING AND CELL DEATH

4.1 SUMMARY

Heterogeneity in mitochondrial content has been previously suggested as a major contributor to cellular noise, with multiple studies indicating its direct involvement in biomedically important cellular phenomena. A recently published dataset explored the connection between mitochondrial functionality and cell physiology, where a nonlinearity between mitochondrial functionality and cell size was found. Using mathematical models, we suggest that a combination of metabolic scaling and a simple model of cell death may account for these observations. However, our findings also suggest the existence of alternative competing hypotheses, such as a nonlinearity between cell death and cell size. While we find that the proposed nonlinear coupling between mitochondrial functionality and cell size provides a compelling alternative to previous attempts to link mitochondrial heterogeneity and cell physiology, we emphasise the need to account for alternative causal variables, including cell cycle, size, mitochondrial density and death, in future studies of mitochondrial physiology.

4.2 INTRODUCTION

The sources of mitochondrial heterogeneity are diverse, as discussed in Chapter 1 (see Section 1.2). In Chapter 3, we have seen how variation in heteroplasmy (the fraction of a variant mtDNA allele per cell) of a pathological mtDNA mutation may have profound effects upon cell physiology (see Section 3.4.4). In this chapter, we investigate the physiological impacts of another mitochondrial feature, namely mitochondrial functionality, upon cellular physiology. We do this by using data from a recently published study [50], which generated a rich dataset probing the relationship between mitochondrial functionality, cell size, apoptosis, cell cycle and metabolism.

We propose a set of quantitative arguments that could account for the dataset in Ref. [50], and invoke several models for bioenergetic scaling, the dynamics of cell proliferation and death, and the coupling between cell physiology and mitochondrial content. Amongst the insights gained from these models, we find that:

- Cell size scaling of cellular power demand implies a reduction in mitochondrial functionality with increasing cell size, in large cells;
- The cell’s characteristic (or “optimal”) size may be largely set by a particular form of intrinsic cell growth dynamics, rather than mitochondrial functionality;
- Cell death alone may induce a non-linearity in mitochondrial functionality with cell size, if mitochondrial functionality is considered a passive indicator of the health of a cell;
- A non-heritable, nonlinear relationship between mitochondrial functionality and cell size, is compatible with a wider set of single-cell data [23] and provides a compelling alternative to link mitochondrial heterogeneity and cell physiology.

We highlight the importance of considering variables such as cell cycle, volume, mitochondrial density and cell death in future studies of mitochondrial heterogeneity. We also propose further experiments to distinguish the models proposed here, and further elucidate the coupling between mitochondria and cell dynamics.

4.3 LINKING MITOCHONDRIAL HETEROGENEITY AND CELL SIZE

Mitochondrial membrane potential is often considered a measure of mitochondrial functionality, since a higher membrane potential may support a larger respiratory rate and ATP/ADP ratio

[152, 281], as well as its role in quality control [164] (although this correspondence may not hold under all conditions, since inhibiting ATP synthase can cause membrane potential to increase [152, 153]). It is important to distinguish between total $\Delta\Psi$, which is expected to scale with a characteristic size of the system (for example, the size of the mitochondrial population, total cellular protein, or some measure of the size of the cell itself e.g. the radius), and a measure of $\Delta\Psi$ which is normalised to account for this scaling. We denote a relative measure of $\Delta\Psi$ with $r\Delta\Psi_m$, $r\Delta\Psi_p$ or $r\Delta\Psi_r$ for normalisations by mitochondrial mass, total cellular protein and cell radius respectively.

The main probe which Miettinen & Björklund [50] used for normalising $\Delta\Psi$ was the lipophilic cationic dye JC-1 [282]. JC-1 can pass cellular membranes and accumulates in regions of negative potential, such as the mitochondrial matrix. JC-1 is green-emitting in its monomeric form; however, if its concentration exceeds a threshold value then it forms aggregates which are red-emitting [282]. By taking the ratio of red to green signals, JC-1 is considered as a semi-quantitative measure of $r\Delta\Psi_m$. Perry *et al.* [153] have suggested several complexities in using this probe. Since aggregates form according to threshold effects based on the concentration of monomers, JC-1 may be ill-suited to detecting subtle gradations in $\Delta\Psi$ and has been recommended to be used as a binary indicator of high/low $\Delta\Psi$. Furthermore, it has been shown that aggregate formation is sensitive to loading concentration [283], suggesting that cellular surface-to-volume ratio may have an effect on aggregation rate. This is of particular relevance to the data of Miettinen & Björklund [50] since variations in cell size are of interest. Finally, equilibration times tend to be long for JC-1, as can be seen in Ref. [50] Fig. S2E (although the qualitative observations of this particular experiment appeared to hold throughout the staining kinetics).

Despite these complexities, many of the key insights of Miettinen & Björklund [50] were reproduced with other means of normalising mitochondrial membrane potential. Throughout their experiments, the authors used flow cytometry to measure cell size using forward scatter (FSC). Calibration of FSC is important for its rigorous interpretation [284]. Upon binning cells by diameter, with bin width of ~ 100 nm, the authors found good correlation between FSC and cell radius across the binned populations. $\Delta\Psi$ was measured with TMRE (another lipophilic cationic dye) and total cellular protein content using CellTrace FarRed stain [285] which covalently binds to cytosolic proteins. The ratio of these signals yields $r\Delta\Psi_p$. The authors found that $r\Delta\Psi_p$ initially increased and subsequently decreased (henceforth referred to as ‘turning behaviour’) with increasing cell radius, which was qualitatively similar to $r\Delta\Psi_m$ found via JC-1 (Ref. [50] Fig. S2K, 1C). The authors also used MitoTracker Red as a measure of $\Delta\Psi$ and Mi-

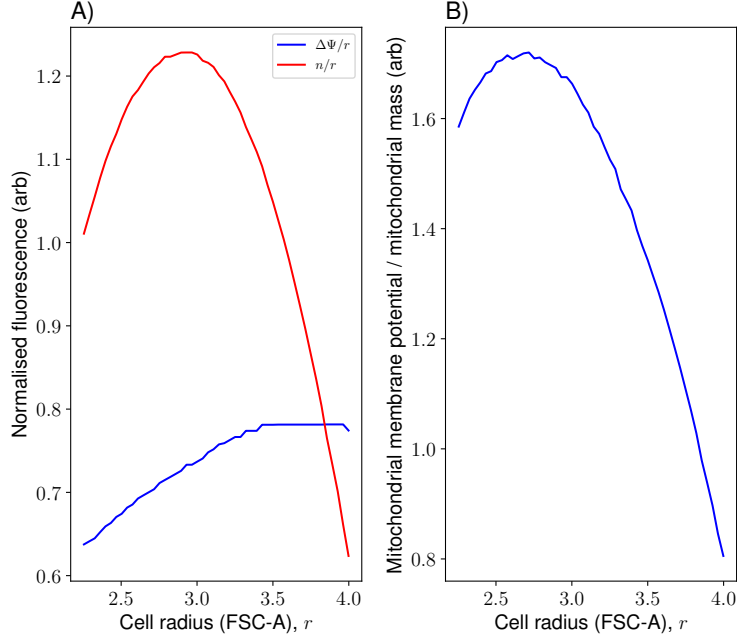


Figure 4.1: Median mitochondrial membrane potential normalised by median mitochondrial mass shows turning behaviour with increasing cell size using MitoTracker dyes (A) Median trends (using loess) of mitochondrial membrane potential / cell radius ($\Delta\Psi/r$, red) and mitochondrial mass / cell radius (n/r , blue) using MitoTracker Red and Green respectively. Data from Ref. [50] Fig. 1D. (B) The curves in (A) were multiplied by r to obtain $\Delta\Psi$ and n , and their ratio was plotted versus r . This yields turning behaviour with r . Since cell radius is presumably connected to cell volume via a purely increasing function, this turning behaviour should also hold for cell volume. Note that it is unclear whether the location of the turning point in (B) coincides with that of Ref. [50] Fig. 1C due to the presence of optical drift (meaning that the absolute values of FSC may change over time), making inter-experimental comparisons difficult.

toTracker Green as a measure of mitochondrial mass (Ref. [50] Fig. 1D). Their data suggested that $r\Delta\Psi_r$ shows turning behaviour with increasing cell radius; re-analysis of these data also suggests that $r\Delta\Psi_m$ shows turning behaviour with either increasing cell radius or volume (see Fig. 4.1).

While the precise interpretation of JC-1 ratio remains somewhat unclear due to the complexities above, the empirical observation that median JC-1 monomer fluorescence closely tracks median total cell protein suggests that it is appropriate for normalisation by cell size [50]. The orthogonal validation experiments suggest some degree of qualitative similarity between $r\Delta\Psi_m$, $r\Delta\Psi_p$ and $r\Delta\Psi_r$, all of which show turning behaviour with increasing cell radius. The qualitative agreement between median JC-1 ratio and these validation experiments suggest that JC-1 ratio is a qualitative measure of system-size-normalised mitochondrial membrane potential. We therefore use the notation $r\Delta\Psi$ to denote JC-1 measurements, which the data of [50] suggests has qualitative similarity to $r\Delta\Psi_m$, $r\Delta\Psi_p$ and $r\Delta\Psi_r$. This turning behaviour was found with the JC-1 probe in immortal cell lines from human (Jurkat), fly (Kc167) and chicken (DT40) as

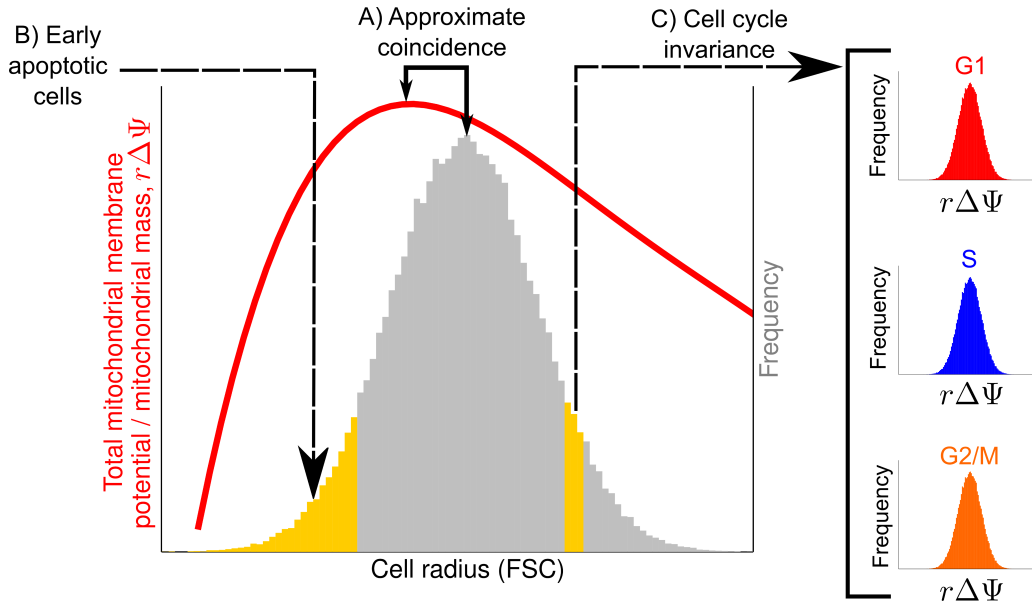


Figure 4.2: Summary of findings of Miettinen & Björklund. The authors binned cells by size of windows approximately 100 nm in diameter. For each bin, the authors measured total mitochondrial membrane potential, normalised by cell size, ($r\Delta\Psi$) using the ratiometric probe JC-1, and reported its median (red curve). **(A)** Approximate coincidence between the mode of the cell radius distribution (grey histogram), determined through forward scatter with flow cytometry, and median $r\Delta\Psi$. **(B)** Although cells were gated for cell death using propidium iodide, the smallest cells in the population were high in an alternative annexin-based sensor of cell death (pSIVA) after normalisation by cell protein content, suggesting that these were early apoptotic cells. Observations (A)&(B) do not exclusively rely on JC-1 measurements. **(C)** For approximately fixed cell size, the authors showed that the distribution of $r\Delta\Psi$ was approximately invariant across different cell cycle stages (red, blue, orange histograms), suggesting that the variation in $r\Delta\Psi$ is more strongly informed by cell volume than cell cycle stage. Note that this observation was not validated independently of JC-1. Based on figures from [50].

well as primary rat hepatocytes and human primary HUVECs (i.e. non-immortalised cells, Ref. [50] Fig. S2L), providing some evidence that this observation may hold beyond immortalised cell cultures.

Miettinen & Björklund [50] found that the maximum $r\Delta\Psi$ over cell size approximately coincided with the modal cell size – see Fig. 4.2. The distribution of $r\Delta\Psi$ was also invariant to cell cycle stage for cells of an approximately fixed size. *The authors therefore postulated that nonlinearity in mitochondrial functionality results in an “optimal” cell size* [50], independently of cell cycle stage.

4.4 CELL PHYSIOLOGICAL MODELS COUPLED TO CELL DEATH

Miettinen & Björklund speculate that an allometric decline causes reduction in mitochondrial functionality with increasing cell size [50, 286], citing the West, Brown and Enquist (WBE) model of allometric scaling of metabolism with body size [287]. Briefly, this model suggests that

transport of materials through fractal space-filling networks explains the relationship between organismal mass and metabolic activity (see Box 4.1 for a discussion of this hypothesis). In contrast, the authors speculate that small cells display low $r\Delta\Psi$ due to a “newborn effect” whereby daughter cells inherit the low mitochondrial functionality of their parents, which itself originates from the allometric decline in the final stages of growth in parent cells. Intermediate-sized cells were postulated to “reset” their metabolic activity, increasing it during early stages of growth [50, 286]. However, both the “newborn” and metabolic “reset” arguments suggest differences in $r\Delta\Psi$ at different stages of the cell cycle, which was not observed using JC-1 (Ref. [50] Fig 2D, S3C, D). We note that the lack of cell-cycle dependence in $r\Delta\Psi$ was not reproduced independently of the qualitative JC-1 probe, so the newborn/reset hypothesis remains possible.

Box 4.1: Allometric Scaling Hypothesis of Cellular Power Demands

The WBE model of allometric scaling [287] claims that the power-law relationship linking an organism’s metabolic rate (B) and mass (M) is $B \propto M^{3/4}$ and can be derived by considering the transport of materials through space-filling fractal networks of branching tubes, for instance the cardiovascular network of most vertebrates. The theory was subsequently extended to explain the metabolic requirements of single-cells in terms of mitochondrial output [288]. Assuming that cell mass and volume (v) are proportional, through constant cellular density, and that the power demand (D) of a cell equates to its metabolic output, the WBE model suggests that the power demand of single cells scales as $D \propto v^{3/4}$. We model mitochondrial power supply (S) in terms of mitochondrial mass density ($\rho = n/v$, for mitochondrial mass n) and mitochondrial functionality (f) as $S \propto \rho v f$. Miettinen & Björklund [50], as well as other authors [44, 45, 49], suggest that $n \propto v$ (i.e. $\rho = \text{const}$) and therefore $S \propto v(r\Delta\Psi)$ (see Section 4.4.1 for further discussion of ρ variation with v), where we have interpreted $f \propto r\Delta\Psi$. Using $S = D$ results in the prediction that $r\Delta\Psi \propto v^{-1/4}$, in contrast to the model we suggest in Eq.(4.3).

Over the relatively small dynamic range of cell volume in a population of cells, the $r\Delta\Psi \propto v^{-1/4}$ relationship of the WBE model and Eq.(4.3) are unlikely to make large differences in their predictions. However, there are some conceptual difficulties in accepting the WBE model above a simpler energetic scaling argument. Whilst mitochondria do form fractal-like networks [289], and the level of fusion of these networks might have non-linear effects on metabolic output [10], it is not clear that a system of static, hierarchical, tubes with fluid flow is an appropriate modelling framework to describe energy transportation within single cells. Because of this, and the criticisms sometimes levelled at allometric scaling [290–292], we have considered a scaling argument which focusses on the potential bioenergetic demands from different specific cellular compartments [293]. Nevertheless, the fusion status of the mitochondrial network remains an appealing potential explanation for the mechanism by which cells alter their mitochondrial functionality to maintain energetic homeostasis with varying cell volume. It therefore remains an interesting area of further study to quantify the extent of mitochondrial fusion and fission [229] as a function of cell volume to explain the $r\Delta\Psi$ variations observed by Miettinen & Björklund [50].

Here, we take the approach of first considering a combination of metabolic scaling and cell death as a possible link between mitochondrial behaviour and cell physiology. We will later consider an alternative purely involving cell death as the underlying causal variable of the observed nonlinearity.

4.4.1 A BIOENERGETIC PROPOSAL: A COMBINATION OF SCALING OF CELLULAR POWER DEMAND WITH CELL SIZE, AND CELL DEATH, MAY ACCOUNT FOR OBSERVED PATTERNS OF MITOCHONDRIAL FUNCTIONALITY

In Box 4.2, we propose that metabolic scaling with a particular model of cell death may induce turning behaviour in $r\Delta\Psi$ with cell radius. Our model suggests that, under non-pathophysiological circumstances, cells alter their mitochondrial functionality to maintain energy supply/demand balance at different cell volumes (v).

The model makes the non-intuitive assumption that cells do not alter their mitochondrial density in response to volume-dependent changes in cellular energy demands: one may speculate that this is due to the timescale of changes in $r\Delta\Psi$ being able to match cellular energy demands, whereas mitochondrial mass (n) may not. Proportionality between n and v has found some support in the literature [44, 45, 49]. Indeed, in Chapter 3 we found that a mathematical model of the cell-physiological consequences of heteroplasmy in a deleterious mtDNA mutation associated with MELAS [37] suggests that failure to maintain the cytoplasmic density of wild-type mitochondrial DNA is associated with a pathophysiological response *in vitro*. However, the data of [50] (neglecting JC-1 measurements) did not strongly constrain the precise relationship between n and v so other scaling relationships are feasible, see Section C.1 for further discussion.

The scaling argument culminating in Eq.(4.3) suggests that $r\Delta\Psi$ decreases with cell volume. This is because mitochondrial mass has been taken to scale proportionately with volume but power demands do not (Eq.(4.1)). Small cells have lower mitochondrial mass, and must therefore have higher mitochondrial functionality to satisfy the constant component of power demand placed by e.g. the nucleus (k_n). As cells become larger and have higher mitochondrial mass, power demand from the nucleus may be more easily satisfied, so mitochondrial functionality reduces towards a constant value k_v .

Box 4.2: Mathematical model of power demand scaling and cell death

We will relate mitochondrial functionality ($r\Delta\Psi(v)$) to cellular volume (v) by modelling $r\Delta\Psi$ as a combination of two models: the first describes the eventual decrease of $r\Delta\Psi(v)$ with v using metabolic scaling arguments, denoted by $r\Delta\Psi_s$. In contrast, the initial increasing part of $r\Delta\Psi(v)$ will be modelled using cell death arguments and denoted by $r\Delta\Psi_\mu$. A combination of $r\Delta\Psi_s$ and $r\Delta\Psi_\mu$ will be used to represent the turning behaviour observed in $r\Delta\Psi$ with cell volume and therefore also cell radius (as the two are related through a monotonic relationship).

We relate mitochondrial functionality ($r\Delta\Psi$) to cellular volume (v) by considering total cellular power supply (S) and demand (D). To represent D , we consider contributions which scale with cell volume (e.g. cytoplasmic protein content), v , surface area (e.g. work done at the plasma membrane), $v^{2/3}$, and constant terms which do not scale with cell size (e.g. replication of the genome). This may be written as

$$D = k_v v + k_s v^{2/3} + k_c \quad (4.1)$$

where k_i are positive real constants, which were loosely based on a study of the power demands of thymocytes [293] (see Section C.2 for further discussion). In general, we might expect S to be of the form $S \propto v\rho(v)f(v)$, where $\rho(v)$ is mitochondrial mass density ($\rho(v) = n(v)/v$, where $n(v)$ is mitochondrial mass) and $f(v)$ is a measure of mitochondrial functionality, in an analogous form to the bioenergetic driving term in [24]. We interpret $f(v) \propto r\Delta\Psi(v)$, since larger values of $r\Delta\Psi$ support a larger respiratory rate and ATP/ADP ratio [152, 281]. We take $\rho(v) = \text{const}$ for parsimony (see Section 4.4.1 for discussion). Hence,

$$S \propto v \cdot r\Delta\Psi. \quad (4.2)$$

We model power supply as being balanced by power demand, $S = D$, with cells varying $r\Delta\Psi$ to maintain this relationship for various v . Rearrangement for $r\Delta\Psi$ yields

$$r\Delta\Psi_s(v) = k_v + k_s v^{-1/3} + k_c v^{-1} \quad (4.3)$$

where the subscript s denotes this metabolic scaling model for $r\Delta\Psi$.

For the cell death component of mitochondrial functionality, we assume that the cell death rate ($\mu(v)$, the probability of cell death per unit time of individual cells) increases amongst cells with smaller volumes. Above a threshold volume (determined by continuity), cell death is modelled as negligible. We then model relative membrane potential as a passive indicator of the cellular death rate μ , since cells which are apoptotic tend to have low mitochondrial membrane potential[294]. We hence use the following linear functions,

$$\mu = \begin{cases} -m_\mu v + c_\mu, & \text{if } v < \frac{c_\mu}{m_\mu} \\ 0, & \text{otherwise} \end{cases} \quad (4.4)$$

$$r\Delta\Psi_\mu(v) = -m_\psi \mu + c_\psi \quad (4.5)$$

where c_μ , c_ψ , m_μ and m_ψ are positive real constants. With this, we recover an alternative description of $r\Delta\Psi$ which initially increases with volume because $r\Delta\Psi_\mu$, μ and v are each connected by negative linear relationships.

We combine $r\Delta\Psi_s$ and $r\Delta\Psi_\mu$ through a rule which ensures continuity in $r\Delta\Psi$. For mathematical convenience, we choose $r\Delta\Psi(v) = \min\{r\Delta\Psi_s(v), r\Delta\Psi_\mu(v)\}$ for all v . This function will show a turning point as cells switch from having their membrane potential governed by risk of cell death to power demand scaling, for increasing cell size. $r\Delta\Psi$ is then observed with Gaussian noise $\mathcal{N}(0, \sigma_\psi)$ with standard deviation (SD) σ_ψ , which accounts for technical and biological sources of noise.

We used a stochastic hybrid systems approach [295] to simulate a population of exponentially growing cells undergoing “adder” dynamics [296, 297], whereby cells add a constant amount of volume (K_v) to their initial volume (v_0) each cell cycle, which has found recent support in bacteria [298]. Cells are modelled to then divide by stochastically partitioning their volume, with the first daughter inheriting volume $v_1 = \mathcal{N}((v_0 + K_v)/2, \sigma_v)$ (with SD σ_v), and the second inheriting $v_2 = v_0 + K_v - v_1$ through conservation of mass. Cell death events occur stochastically according to an inhomogeneous Poisson process with rate $\mu(v)$ (Eq.(4.4), see Section C.2 for simulation details). The results of this model are shown in Fig. 4.3.

This model alone suggests that $r\Delta\Psi$ is a purely decreasing function of cell volume: however, evidence exists for a turning point in the relationship between $r\Delta\Psi$ and v [50]. The data in Ref. [50] shows increased cell death in the smallest cells in the population (Ref. [50] Fig. 3C), and potentially a subtle increase in the largest cells also. Furthermore, the authors observed a depolarised subpopulation of cells at low JC-1 monomer values (Ref. [50] Fig. S2G), potentially indicative of a subset of cells with low mitochondrial functionality at low cell sizes. Cell shrinkage is associated with apoptosis [299], as is loss of mitochondrial membrane potential [294]. Hence, we suggest that smaller cells have a higher cell death rate (although our model has no explicit shrinkage mechanism – simply that smaller cells produced through cell division are at higher hazard of cell death), and that mitochondrial functionality may be a passive indicator of cell death under these circumstances (Fig. 4.3A). The result of simulating this model in a proliferating population of cells undergoing “adder” dynamics [296, 297], whereby cells add a constant amount of volume each cell cycle, is presented in Fig. 4.3.

As can be seen in the distribution of cell volumes (Fig. 4.3C), adder dynamics imposes a characteristic cell size associated with the mode of the distribution, which is often referred to as an “optimal” cell size. The width of the final distribution is set by the variability in partitioning of volume between daughters at division (although other sources of variability may be envisaged such as noise in the growth rate or the amount of volume added per cell cycle [300], which are neglected here). Under these simulation dynamics, the modal cell volume and the maximum value of $r\Delta\Psi$ are essentially independent (Fig. 4.3C) and can be tuned relatively to each other through appropriate choices of model parameters. The lack of alignment between the modal cell size and maximal $r\Delta\Psi$ could also be achieved with “sizer” dynamics, where cells divide once exceeding a maximal cell size. The coefficient of variation (CV) in $r\Delta\Psi$ also showed a minimum (Fig. 4.3D), in qualitative agreement with [50], although the minimum CV was constrained to coincide with the minimum $r\Delta\Psi$ (see Section C.3 for further discussion). *It is therefore largely compatible with the data of [50] to suggest that a cell’s characteristic size is set by a particular form of intrinsic cell dynamics, rather than requiring a direct link with $r\Delta\Psi$.*

In Section C.4, we explore an alternative hypothesis where cell death rate is a non-linear function of cell size, and that mitochondrial membrane potential is a passive readout of cell death rate *over all values of cell volume*. We found that this is also a compatible interpretation of the data presented in [50]. However, the causality of this model may also be inverted, where mitochondrial functionality is a non-linear function of cell volume, and cell death rate is linearly related to mitochondrial functionality. This highlights the importance of excluding cell death

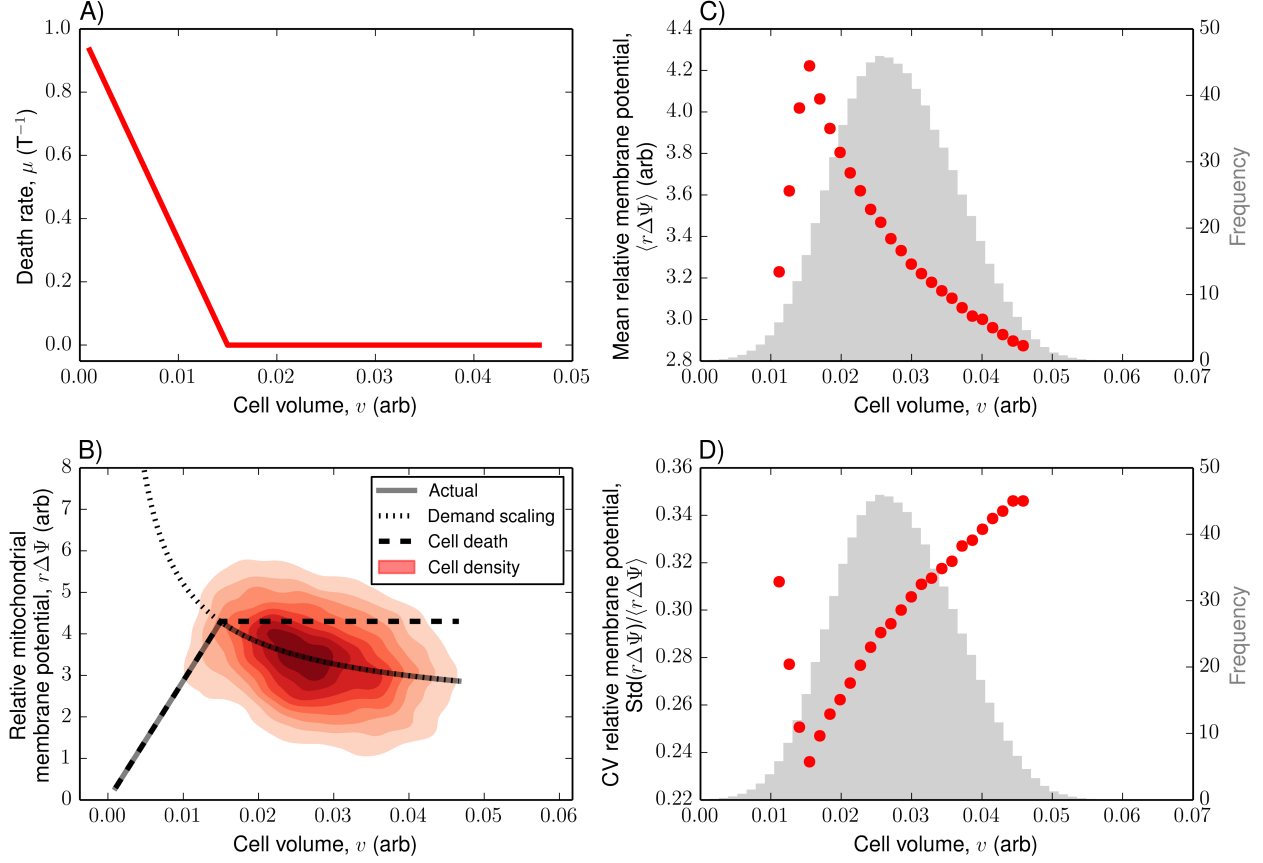


Figure 4.3: Power demand scaling, in conjunction with a simple cell death model, may account for observations of Miettinen et al. Cells obeying adder dynamics, initialised with random uniform volumes, were simulated with a volume-dependent death rate. **(A)** Model for death rate (probability of cell death per unit of simulation time T) as a function of cell volume. Death rate was assumed to be zero after cells reach a threshold size for computational convenience. **(B)** Power demand scaling (black dotted) was determined through a power supply = demand relationship in Eq.(4.3). Cell death scaling (black dashed) was determined through a piecewise linear relationship in volume (Eq.(4.5)). The mean $r\Delta\Psi$ (grey) was determined by taking the minimum of the demand scaling and cell death curves at every point in v (to model switching behaviour which is continuous in v). Addition of Gaussian noise to the mean relationship recovers a joint density of v and $r\Delta\Psi$ (red shaded region) similar to the findings in Ref. [50] (Fig. 1B therein). **(C-D)** Cells were binned by volume in silico, and for each bin the mean and standard deviation of $r\Delta\Psi$ and v were computed. **(C)** Mean $r\Delta\Psi$ (red points) shows a maximum with respect to volume which is not aligned with the modal cell volume (grey histogram of cell volumes), recapitulating the observation in Ref. [50] (Fig. 1C therein). **(D)** The coefficient of variation of $r\Delta\Psi$ (red points) shows a minimum with respect to volume, as observed in Ref. [50] (Fig. 3A therein), although the location of the minimum CV and maximum average $r\Delta\Psi$ are constrained to coincide in our model (C&D) – see Section C.3 for further discussion. For simulation details, see Section C.2.

as a potential confounding variable in future confirmatory studies so that these hypotheses may be disambiguated.

Future verification studies may include gating cells in FACS analysis based on the pSIVA cell death probe used in [50]. Alternatively, fluorescence labelling of upstream apoptosis markers, such as cytochrome-c, may allow detection of cells in early stages of cell death [301]. Combining this with fluorescence microscopy and tracking mitochondrial membrane potential, mitochondrial mass and cell volume, may allow early-apoptotic cells to be excluded and verification of the nonlinearity between cell volume and relative mitochondrial membrane potential.

4.4.2 A PARABOLIC FUNCTION OF MITOCHONDRIAL FUNCTIONALITY AGAINST CELL SIZE IS COMPATIBLE WITH A MATHEMATICAL MODEL OF MITOCHONDRIAL VARIABILITY

A previous model, accounting for the data in [23], described the cellular consequences of mitochondrial variability in HeLa cells [24]. That model phenomenologically described mitochondrial functionality as a stochastically inherited quantity, determined through a mean-reverting process (an AR(1) process). The biological mechanism for the heritability of mitochondrial functionality was unclear from that study, although one may speculate that the inherited concentration of mitochondrial fusion proteins may be a feasible explanation for this. Indeed, recent evidence suggests that the proliferation rate of yeast is heritable, and that slow-growing cells have higher oxidative stress whose proliferation rate may be rescued with an anti-oxidant, suggesting a link to mitochondrial functionality [302]. Further studies are required to determine whether this is true for a larger class of systems, including mammalian cells.

The data in Ref. [50] immediately suggests an alternative nonlinear relationship between function and dynamics (see Fig. 4.2). We hypothesised that this relationship may be sufficient to account for the data in [23] without invoking a phenomenological mechanism for the stochastic inheritance of mitochondrial functionality (though the stochastic inheritance of mitochondrial mass remains supported). Briefly, the model of Johnston *et al.* [24] predicts the relationship between cell volume, mitochondrial content and mitochondrial functionality with a pair of coupled ordinary differential equations. Cells undergo sizer dynamics, and the population size is controlled to be static over time.

Here, we replace the stochastic inheritance of mitochondrial functionality in Ref. [24] by approximating the relationship between mitochondrial functionality (f , which we interpret as proportional to $r\Delta\Psi$) and cell volume (v) phenomenologically as a non-heritable parabolic

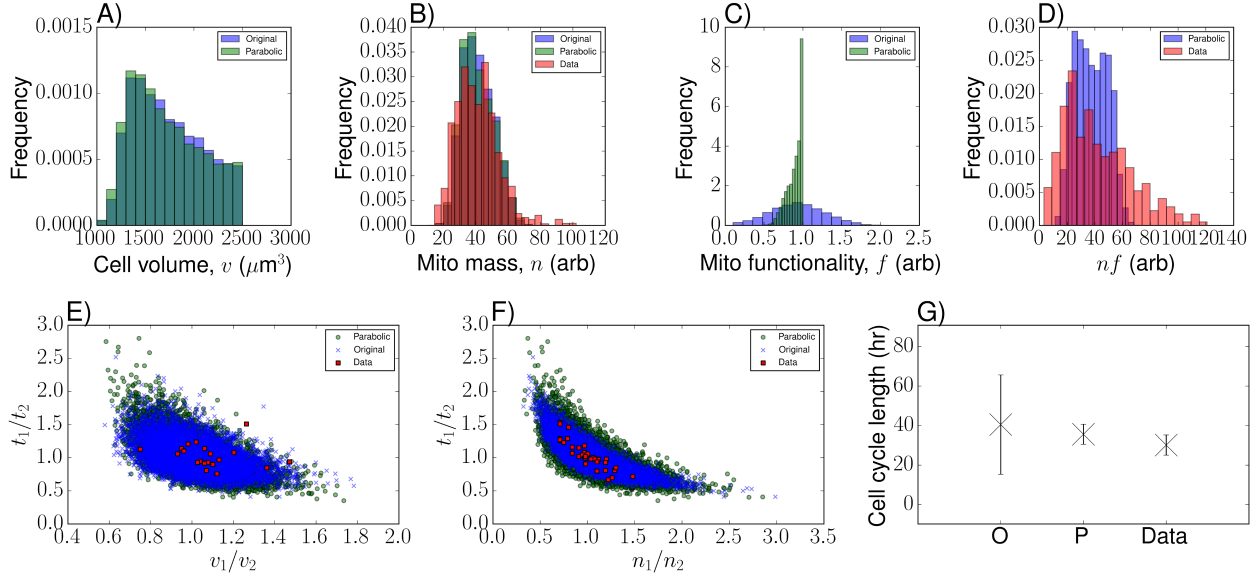


Figure 4.4: Nonlinear connection between mitochondrial functionality and cell volume is consonant with a wider set of single cell data. Replacing the stochastic process determining mitochondrial functionality (Original) with the parabolic relationship in Eq. (4.6) (Parabolic), and comparisons with data from [23] where available. (A) Distribution of cell volumes is relatively unchanged by the modification of f . (B) Distribution of mitochondrial mass in both models are comparable with data. (C) Distribution of mitochondrial functionality. Since Eq.(4.6) is deterministic, f cannot exceed its maximal value f_{max} , causing the skewness observed. (D) The distribution of total mitochondrial functionality nf has a reduced variance in the parabolic model, which departs somewhat from the data of [23]. (E) Relationship between ratio of cellular volumes at birth and the ratio of cell cycle lengths for sister pairs. (F) Relationship between ratio of mitochondrial contents at birth and the ratio of cell cycle lengths for sister pairs. Both models are comparable with data in (E&F). (G) Mean cell cycle lengths $\pm$ standard deviation for the original (O) and parabolic (P) models, as well as the experimental observation in [23] (Data). The variance in cell cycle lengths is greatly reduced, since an additional source of variability is removed in P from the deterministic relationship Eq.(4.6), which has better agreement with data. $v_{opt} = 2000$, $f_m = 2.0$, $f_{max} = 1.0$. All other parameters appearing in the original model were kept the same as [24].

relationship

$$r\Delta\Psi \propto f = f_{max} - f_m \left(\frac{v - v_{opt}}{v_{opt}} \right)^2 \quad (4.6)$$

where f_{max} is the peak value of f , occurring when $v = v_{opt}$, and f_m is the steepness of the parabola. In implementing the original model from [24], it was found that constraining $f > 0.1$ from the AR(1) process (rather than $f > 0.01$ described in [24]) yielded a better comparison to experimental cell cycle lengths. The result of replacing the AR(1) process in [24] with the deterministic relationship in Eq.(4.6) is shown in Fig. 4.4.

We found that whilst the distributions of cell volume and mitochondrial mass were relatively unchanged (Fig. 4.4A&B), mitochondrial functionality had greater skewness under the parabolic model (Fig. 4.4C). The distribution of total mitochondrial content nf was narrower in the parabolic model (SD = 11.5) than the original model (SD = 17.4) (Fig. 4.4D) which has poorer agreement with data from [23]. However, the cell cycle length distribution in the parabolic

model had much better agreement in its standard deviation ($SD = 5.07$) than the original model ($SD = 25.2$) compared with data ($SD = 5.13$) (Fig. 4.4H). The low degree of correlation between cell cycle lengths between parent and daughter cells was also retained (Original: $R^2 = 0.11$, Parabolic: $R^2 = 0.05$) in agreement with the data of [23] (Fig. 3C of [24]). Overall, this shows that the non-heritable nonlinear mechanism between mitochondrial functionality and cell volume found in [50] may broadly account for the data in [23] without invoking a phenomenological stochastic model for the inheritance of mitochondrial functionality.

4.5 DISCUSSION

Heterogeneity in mitochondrial content is exerted along many axes, including intrinsic functional variability, network status and genetic variability (see Chapter 1). A host of cellular phenomena are affected by noise, a growing number of which are being directly attributed to mitochondrial variability. Recent work [50] has extended this picture by probing the link between mitochondrial content and cell volume, suggesting a nonlinear relationship between mitochondrial functionality and cell radius.

We have shown that a simple model of power demand scaling predicts a decreasing trend in relative mitochondrial functionality with increasing cell volume. By combining this with a simple cell death model, we were able to qualitatively account for a large number of observations in Ref. [50]. We also found that incorporating a nonlinear relationship between $r\Delta\Psi$ and cell volume into a model of mitochondrial heterogeneity [24] was broadly consonant with a wider set of single cell data [23]. We have found that it is also compatible with the data of Ref. [50] to suggest that cell death is the causal variable in determining the nonlinearity in $r\Delta\Psi$. Our analysis has shown the existence of multiple competing hypotheses which are able to account for the data in Ref. [50], thus highlighting the importance of accounting for cell cycle, volume, mitochondrial density and cell death in future single-cell studies of mitochondrial heterogeneity.

5

SOMATIC MOSAICISM DURING NEURODEVELOPMENT IN GENES ASSOCIATED WITH NEURODEGENERATION

5.1 SUMMARY

Errors in replication of DNA during the intense cellular expansions involved with development may give rise to mutations in fully-developed tissues, some of which may be pathological – this is termed ‘somatic mosaicism’. In this work, we seek to estimate the size and frequency of pathologically mutated regions in the adult human brain by taking a highly simplified approach of modelling neurodevelopment as a deterministic branching process. We used data from a recent study of 54 post-mortem individuals to infer a distribution of plausible neurodevelopmental mutation rates using approximate Bayesian computation. We found that the inferred mutation rates are compatible with other literature values for the nuclear DNA mutation rate. As a preliminary means to estimate the propensity of neuronal somatic mosaicism, we forward-simulated from our model and found that approximately 11% of simulated individuals harboured one or more pathologically mutated, spatially contiguous, regions consisting of approximately 10^5 neurons. We perturbed our simple model to investigate effects such as cell death, founder effects from progenitor cells, and asymmetric division, finding that many of our

extrapolations are robust. A simple order-of-magnitude calculation accounts for this robustness: assuming neurodevelopment is unbiased and spatially proximal once the brain consists of 10^6 cells, given a mutation rate comparable to literature values, individuals are expected to harbour one spatially contiguous pathologically mutated region consisting of 10^4 – 10^5 cells. This opens the possibility of somatic mosaicism induced through neurodevelopment contributing to neurodegenerative disease risk. The methods described here have the potential to allow exploration of somatic mosaicism in various contexts, such as pathological mutations in nuclear-encoded mitochondrial genes and the mitochondrial genome.

5.2 INTRODUCTION

The human body consists of approximately 10^{13} cells [303]: consequently an enormous number of cell divisions must occur in order to create a human being from a single cell zygote. As discussed in Chapter 1, a finite mutation rate implies that a set of closely-related DNA sequences are expected to occur which do not all necessarily maximise organismal fitness [60, 61]. In contrast to the previous chapters, here we will focus on the nuclear genome, rather than the mitochondrial genome, so we may therefore consider a cell as consisting of 2 genomes per cell (since humans are diploid organisms). Given that development, roughly speaking, may be thought of as a tree [304–306], where nodes are cells and edges denote parent-daughter relationships, mutations which occur early in development may give rise to large numbers of mutated daughter cells [307]: this is termed ‘somatic mosaicism’ whereby genetically distinct subpopulations of somatic cells may exist in a given organism, which may give rise to pathology [308]. In neurodegenerative disorders, such as Parkinson’s and Alzheimer’s disease, $\sim 5\%$ of cases are caused, or contributed to, by germ line mutations in one of only ~ 50 genes [309, 310]. It is therefore plausible that somatic mosaicism during neurodevelopment may contribute to the pathogenesis of non-familial cases of neurodegenerative disorders.

Mathematical exploration of mutational frequencies in exponentially growing populations dates back to the 1940’s in the experiments of Luria and Delbrück [311], which were in bacterial populations. The authors developed a model assuming the initial condition of a single cell, which divides to yield an exponentially growing population, and accumulates mutations according to a Poisson process [311]. The number of mutants existing at time t under this model has since been termed the Luria-Delbrück distribution, and much has since been learnt about its properties [312–314]. The Luria-Delbrück distribution has been used to estimate the propensity of mutations seeded during development to eventually contribute to cancer risk [304, 305]. Various

branching process models have been used to understand oncogenesis and cancer development [306, 315]. Somatic mosaicism has also been proposed to play a role in the neurodegenerative disease amyotrophic lateral sclerosis [305], which potentially may be seeded by a small focus of genetically altered cells in the brain. Recent single-cell studies have found evidence for somatic mosaicism in the brain [316–318]; however, demonstration of somatic mosaicism at a larger scale across the human brain has yet to be performed.

In this chapter, we attempt to demonstrate the large-scale prevalence of somatic mosaicism in the human brain by analysing a recent dataset which involved ultra-high depth sequencing of 56 genes associated with common neurodegenerative diseases in 173 brain tissue samples across 54 adult human individuals [199]. From these data, we sought to infer the mutation rate during neurodevelopment by applying model-based inference. The simple nature of the available data made us favour a heavily simplified approach to modelling neurodevelopment, which we modelled as a deterministic branching process (binary tree). We applied approximate Bayesian computation on this model and dataset to infer the approximate posterior distribution of mutation rates during neurodevelopment. Given the plausible distribution of mutation rates, we sought to determine the prevalence of pathologically mutated regions of brain tissue in the human population through forward-simulation from the model. Our simplified approach to modelling neurodevelopment allowed us to make a number of modifications to the model assumptions, and easily understand their implications through forward-simulation (in contrast to a more technical approach leveraging the literature relating to e.g. the Luria-Delbrück distribution [311–313]). We therefore modified the deterministic branching process to account for several important factors, such as cell death; founder effects from progenitor cells; asymmetric division; and cellular migration. We found that many of our observations were robust to such model perturbations. This robustness arises from a simple order-of-magnitude calculation: assuming that the mutation rate of DNA is on the order of 10^{-10} – 10^{-9} mutations per base per cell division [306, 319], given unbiased, spatially proximal, replication of daughter cells once the brain consists of approximately 10^6 cells, then we suggest that every individual is expected to possess 1 pathologically mutated region of brain tissue consisting of 10^4 – 10^5 cells. This argument is independent of the neurodevelopmental mechanism prior to the brain consisting of 10^6 cells. Our model serves as an initial means to explore the prevalence of these mutated foci, suggesting that such foci are possessed by almost all individuals, as has been suggested previously in the context of cancer [304–306, 320, 321]. We hypothesise that such regions have the potential to generate proteinaceous fibrillar structures which could spread through the human brain and cause neurodegenerative disease [322].

5.3 METHODS

Our analysis is based upon a dataset which sequenced the exons of 56 genes known to cause, or predispose to, common neurodegenerative disorders (132,617 base pairs) [199]. Sequencing was performed on 173 brain samples from 54 individuals, each consisting of 1 cm³ blocks of brain tissue. The samples of grey matter were manually dissected to remove macroscopically identifiable white matter and vascular tissue. Samples were homogenised, DNA was fragmented, and 1 μ g of genomic DNA per sample was sequenced, capturing the equivalent of 3415 cells per sample on average [199] (although, since the DNA was fragmented, each fragment could potentially originate from any cell in the sample, which is much greater than 3415 cells). Brain tissue was obtained from post mortem cases of Alzheimer’s disease (AD, $n = 20$ brains), Lewy body disease (i.e. Parkinson’s disease or dementia with Lewy bodies; LB, $n = 20$ brains), and age matched controls with no significant neuropathology ($n = 14$ brains) [199]. A bioinformatics pipeline to determine true single nucleotide variants (SNVs) from sequencing error was performed (see [199] for details).

We discounted one variant in a control subject who displayed a mutation in all measured samples; the mutation was thought to be hematopoietic in origin [199], and therefore not useful in understanding the prevalence of somatic mosaicism during neurodevelopment. All other detected variants occurred in one sample per individual. Fisher’s exact test showed no significant difference between the number of individuals displaying mutations in control and neurodegenerative (AD + LB) cohorts (p-value = 0.66).

We restricted our attention to the neurodegenerative cohort (consisting of AD and LB cohorts, 40 individuals) and neglected the control cohort (14 individuals). Although there is insufficient evidence to reject the null hypothesis that both patient and controls have different mutational propensities from Fisher’s exact test (see above), this is not the same as positive evidence for both groups having the same mutation rate. For simplicity, we restrict our attention to the larger case cohort ($n = 40$ patients), but the similar frequencies of mutation across these groups provides an initial suggestion that our findings could generalise to wider populations.

5.3.1 ORDER-OF-MAGNITUDE CALCULATION FOR THE MUTATION RATE DURING NEURODEVELOPMENT

Before proceeding with a description of our model-based inference procedure (see Section 5.3.2), we first present a simple argument to deduce the mutation rate given the data; doing so will

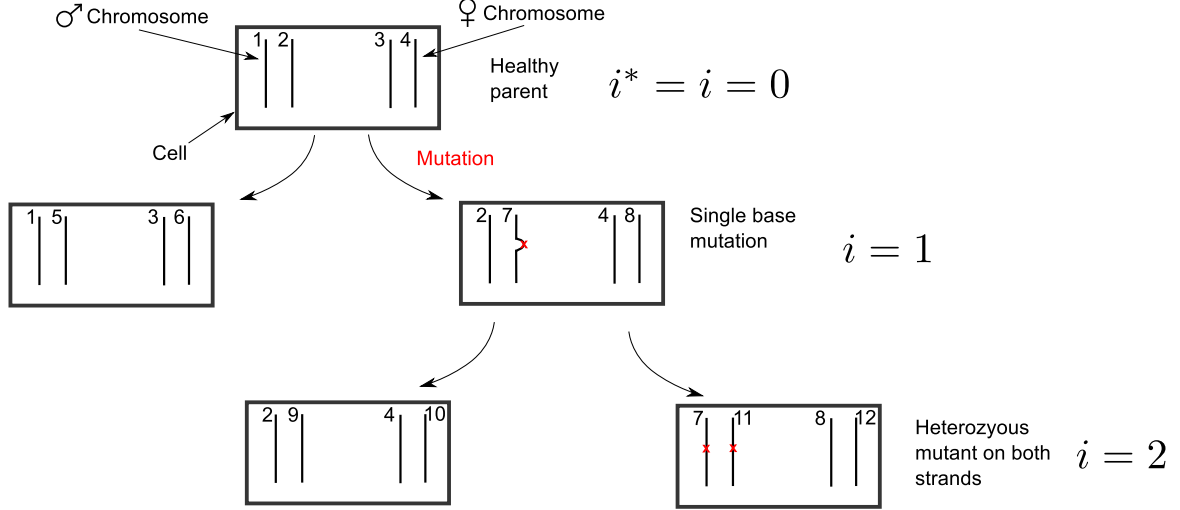


Figure 5.1: Branching process model of neurodevelopment. We idealise the cell as consisting of two chromosomes, one from each parent. During division each strand is copied. To illustrate this, strands are labelled with integers and new strands are labelled with a new integer. Neurodevelopment is modelled to begin with a single founder cell, the generation number beginning at $i = 0$. Consider a copying error which generates a SNV in strand 7 whilst copying strand 2 at generation $i^* = i = 0$. Strand 7 is subsequently copied in the following generation to create a heterozygous mutant present on both strands. Thus, a mutation occurring during copying at generation $i^* = 0$ causes 1 in 4 daughter cells at generation $i = 2$ (and all subsequent generations) to harbour a heterozygous mutant.

motivate our modelling choices in Section 5.3.2. First, let us assume that neurodevelopment may be modelled as a binary tree, where nodes are cells and edges are parent-daughter relationships (as will also be the case in Section 5.3.2). Neurodevelopment is assumed to begin with a single cell, and cells divide at discrete generations (labelled i), generating two daughters each. The number of mutations that arise at generation i^* (X_{i^*}) is

$$X_{i^*} \sim \text{Binom}(4N_b \cdot 2^{i^*}, \lambda_m) \quad (5.1)$$

where N_b is the number of bases of interest, and λ_m is the number of mutations per base per copying event. $\text{Binom}(N, p)$ is the binomial distribution with N trials and probability of success p . The factor of 4 counts for the ploidy of the human genome (2 genomes) and the number of strands in DNA (2 strands) (see Fig. 5.1).

Given a mutation event at generation i^* , we define the fraction of the adult brain affected by the mutation as f , where

$$f = 2^{-(i^*+2)}. \quad (5.2)$$

Note that f is independent of the total number of generations, and therefore independent of the final size of the adult brain. The factor of $(+2)$ arises from the lag associated with copying a single-stranded mutation into a daughter cell to become a double-stranded mutation, see

Fig. 5.1.

Let us compute the average total number of mutated cells originating from generation i^* in the final adult brain ($\bar{N}_{i^*t}$). Let N_{br} denote the total number of cells in the adult brain as a result of a given number of generations of the binary tree. Then using Eq.(5.1) and Eq.(5.2), and making the assumption that mutations are highly diluted in the adult brain such that no cell contains more than one mutation, then

$$\bar{N}_{i^*t} \approx \mathbb{E}(X_{i^*}) \cdot f N_{\text{br}} = N_b \lambda_m N_{\text{br}} \quad (5.3)$$

where, notably, $\bar{N}_{i^*t}$ is itself independent of i^* .

Now let us compute the total number of detectably mutated cells across the entire human brain, N_d . We do this by imagining that the brain may be split into ‘voxels’, each consisting of a fixed number of cells (N_v). Let $\bar{v}$ equal the average fraction of mutated cells in a detectably mutated voxel, and p_{mv} equal the probability that a voxel is detectably mutated. Then we may simply write

$$N_d = (\bar{v} N_v) \cdot \left(\frac{N_{\text{br}}}{N_v} p_{\text{mv}} \right) \quad (5.4)$$

where the first term is the average number of mutated cells in a mutated voxel, and the second term is the total number of mutated voxels.

Measurement of mutated cells within a sample has an associated sensitivity: if the fraction of mutated cells within a mutated sample is below the detection sensitivity then those mutations will not be detectable. Eq.(5.1) implies that mutations lower in the neurodevelopmental tree (larger i^*) are exponentially more common, whereas Eq.(5.2) implies that such mutation are also exponentially smaller in size. The existence of a sensitivity threshold suggests that the most common detectable mutations will correspond to a particular i^* which give rise to mutations corresponding to the detection threshold. Mutations with larger i^* are too small to detect by definition, whereas mutations with smaller i^* are exponentially less common. Detectable mutations therefore likely originate from a narrow band of i^* values around the detection threshold: we therefore assume that all detectable mutations come from a single i^* value, for parsimony. Consequently, we may equate Eq.(5.3) and (5.4) to yield

$$\lambda_m = \frac{\bar{v} p_{\text{mv}}}{N_b}. \quad (5.5)$$

Note that, strikingly, *this estimate for λ_m is independent of the voxel size, N_v* . Since $\bar{v}$ is likely to correspond to the detection sensitivity threshold of the experiment, which is an experiment-

intrinsic constant (see argument above), and λ_m and N_b are constants, this implies that p_{mv} is also a constant. Hence neither $\bar{v}$ nor p_{mv} have implicit dependence upon i^* or N_v .

Let us take N_v to be the number of cells in the 1 cm³ sample, i.e. we consider splitting each of the brains from the experiment of [199] into 1 cm³ voxels. 6 of the 173 measured samples were found to contain detectable mutations, so we may take $p_{mv} = 6/173$. We equate the experimentally determined average variant allele fraction (VAF) [199] with the average fraction of mutated cells in a mutated sample, such that $\bar{v} = 0.87 \times 10^{-2}$ (which is comparable to the detection threshold of approximately 0.5%) [199]. Note that this assumption is justified as each sample which contained a mutation, contained exactly one detectable mutation, therefore there likely exists a correspondence between the number of mutated sites and the number of detectably mutated cells amongst the captured cells of the sample. Therefore the average VAF may be thought of as an unbiased estimate of $\bar{v}$. The number of bases sequenced was $N_b = 132617$ bp. Thus, substituting these values into Eq.(5.5), this simple argument yields

$$\lambda_m = 2.3 \times 10^{-9} \text{ mutations per base} \quad (5.6)$$

which is similar in order of magnitude to the estimate of Tomasetti *et al.* [306], which found $\lambda_m = 0.64 \pm 0.11 \times 10^{-9}$ per base.

5.3.2 APPROXIMATE BAYESIAN COMPUTATION FOR A CRUDE BRANCHING MODEL OF NEURODEVELOPMENT

The genetic information we have allows us to make preliminary statements about the number of cells that might be affected by pathological mutations. The objective here is to place the empirical observations of Ref. [199] in a simplistic theoretical context, not to provide a definitive set of estimates. For more definitive estimates, more refined models, inference and data would be required. That noted, the inference outlined below can be decoupled from the collected dataset by simply using literature estimates for the mutation rates (e.g. [306]), or the estimate culminating in Eq.(5.6).

We consider neurodevelopment as a deterministic branching process where every cell has exactly two daughters per generation and the probability of death is zero (see Section 5.4.2.1 for an investigation of non-zero cell death). We assume that the brain begins with a single founder cell (see Section 5.4.2.2 for a relaxation of this assumption) which doubles exactly 37 times, yielding $2^{37} \approx 1.37 \times 10^{11}$ cells in the adult brain (both neuronal and non-neuronal cells).

We use an integer number of doublings for computational convenience, and chose 2^{37} as the closest number of cells to experimental measurements of the adult brain, being approximately $(1.7 \pm 0.14) \times 10^{11}$ cells ($\pm$ SD)[323]. This is a slight underestimate of the true number of neurons in the adult brain (falling 2.4 SD below the mean, 19% difference in mean estimates).

The way in which extracted tissue samples from Ref. [199] map to the simplistic tree structure of our model of the brain is clearly non-trivial. Samples consisted of 1 cm^3 blocks of tissue. Given that a human brain consists of approximately 10^{11} cells [323], and the volume of an adult brain is approximately 10^3 cm^3 , each sample consisted of approximately 10^8 cells, of which only 10^3 were sequenced [199]. The inference approach described below requires the full developmental history of all sampled cells to be simulated (see Fig. 5.2). However, simulation of 10^8 cells is computationally challenging. It is also unclear how mesoscopic fractions of the brain (such as $1/1000$) could reasonably be mapped onto a simple binary tree given the complicated nature of neurodevelopment. The argument presented in Section 5.3.1 suggests that the inferred mutation rate is independent of the size of the sample measured, and that the probability of observing a mutated sample is independent of sample size. We therefore make the strong assumption that the sampled DNA which was captured and sequenced originates from a set of cells which are spatially contiguous in the binary tree, i.e. sequenced DNA originates from a sample of cells consisting of $N_v = 3415$ cells, as opposed to $N_v = 10^8$ cells. Note that we subsume the fragmentation of DNA prior to sequencing into the assumption that $N_v = 3415$ cells; i.e. we assume that all sequenced DNA fragments originate from these $N_v = 3415$ cells. The argument in Section 5.3.1 provides an initial suggestion that our inferred mutation rates will not depend strongly upon this assumption. However, for more definitive estimates, more accurate models of neurodevelopment may be required or, alternatively, a tissue sampling strategy for which a 1D embedding in a binary tree may be more appropriate (such as sampling several thousand cells, as opposed to hundreds of millions, with a high sequencing capture efficiency). We again emphasise that the results presented below may be decoupled from the inference by simply using literature values for the mutation rate, such as Ref. [306], or by using Eq.(5.6) which is comparable in magnitude.

Our inference approach is a simple approximate Bayesian computation rejection algorithm, and proceeds by sampling a mutation rate (λ_m) from a prior distribution, which was motivated by the literature. A study by Roach *et al.* involved family-based whole genome sequencing in peripheral blood cells to deduce the human intergeneration mutation rate as $\sim 1.1 \times 10^{-8}$ per bp per generation per haploid genome [319]. In order to approximate the mutation rate per base per cell division, we estimate the number of cell divisions per generation to form the adult

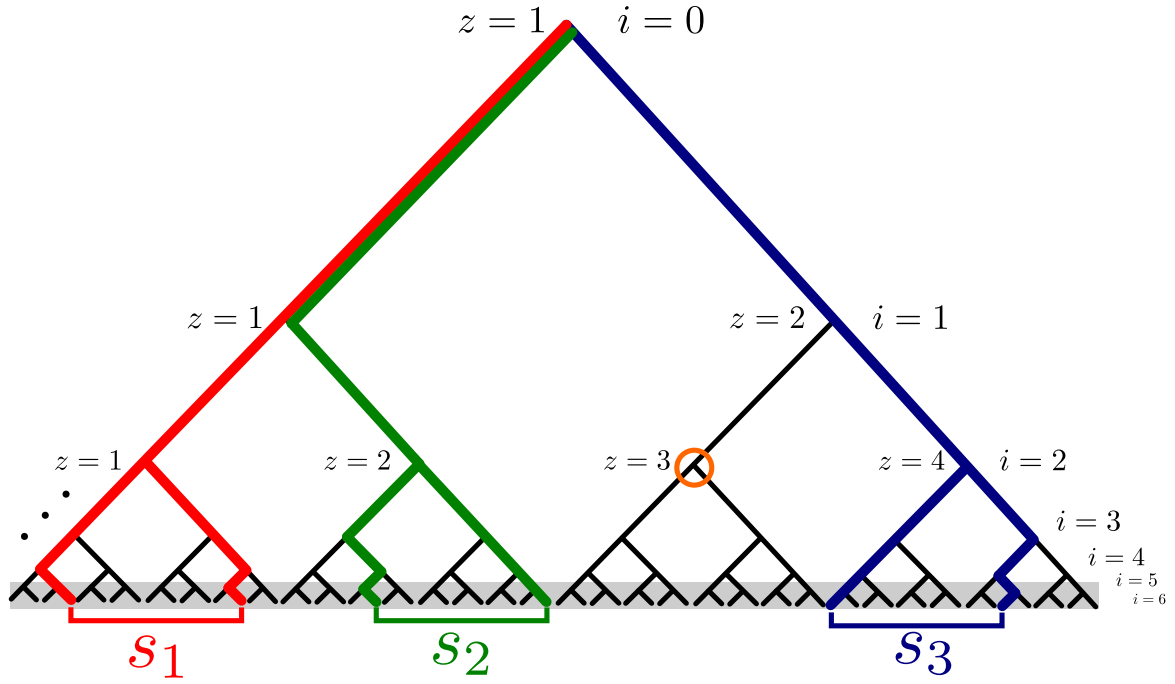


Figure 5.2: Inference strategy for approximate Bayesian computation of mutation rate. In this illustrative example, assume that the human brain consists of 2^6 cells and 3 samples (s_1 , s_2 and s_3) of 11 cells are taken from the brain. We may assign integer labels to cells from left to right for a given generation number ($z \in [1, 2^i]$ for integer z), giving each cell in the lineage a unique coordinate (z, i) . Given the coordinates of the sampled cells, we only wish to simulate mutation events in the lineage of parents which give rise to sampled cells: these are the cells which lie on or between the coloured boundaries. For instance, at generation $i = 1$, we must simulate mutations for both cells $z = 1$ and $z = 2$. However, cell $(3, 2)$ (circled) has no influence over the sampled cells, and therefore does not need to be simulated. Cells in the final generation ($i = 6$ in this case) form the adult brain and are not replicated. Replication of cells in the penultimate generation cannot give rise to mutants present on both strands of DNA (see Fig. 5.1), and therefore may also be neglected (grey shaded region) since mutations were checked for their presence on the reverse strand. Thus, in this example, we need only consider cells with $i \leq 4$.

brain as 36 (given our model). Since both strands of DNA are replicated during cell division, we estimate $\lambda_m \approx 1.1 \times 10^{-8} / (2 \times 37) \approx 10^{-10}$ per base per cell division. However, this order of magnitude estimate is subject to large uncertainty, since it is unclear the extent to which the germline mutation rate is representative of the mutation rate during neurodevelopment. Consequently, we chose to search ± 2 orders of magnitude around this estimate, using the following broad uninformative prior on λ_m :

$$P(\ln(\lambda_m)) = \text{unif}(\ln(10^{-12}), \ln(10^{-8})) \quad (5.7)$$

where $\text{unif}(a, b)$ is a uniform distribution between a and b . Note that the prior of λ_m is uniform on log-space, and therefore encodes our belief that λ_m is equally uncertain over 4 orders of magnitude

Given a sample of the mutation rate from the prior, we are then able to simulate neurodevelopment for an individual. Our strategy will be to simulate 40 such individuals (the number of individuals in the cohort with neurodegenerative disease) and, loosely speaking, if a simulation is ‘close enough’ to the data we will accept the draw from the prior as a sample from the approximate posterior distribution for λ_m , which captures all of our uncertainty in the parameter given the data. This is the essence of approximate Bayesian computation (ABC).

In Fig. 5.2 we show that it is only necessary to consider the set of parent cells which give rise to the cells in the sample (given the assumption of spatial contiguity of the sampled cells, see above). Once the set of parental cells have been determined, we seek to compute the locations of mutation events in the lineage. Given a draw of the mutation rate λ_m from the prior (see Eq.(5.7)), we assume that each base has a probability λ_m of mutating during replication. The total number of bases which are copied during development (B) is $B = 4N_bN_p$ where N_b is the number of base pairs to be considered (132617 bp in the case genes) and N_p is the number of parent cells which gives rise to the sampled cells. The factor of 4 counts for the ploidy of the human genome (2 genomes) and the number of strands in DNA (2 strands) (see Fig. 5.1). The total number of mutation events which occur in the lineage (M) is therefore modelled to be

$$P(M) = \text{Binom}(B, \lambda_m). \quad (5.8)$$

In order to find the locations of mutation events in the lineage, $R_1, R_2, \dots, R_m$, we wish to sample from the joint distribution $P(R_1 = r_1, R_2 = r_2, \dots, R_m = r_m, M = m)$ where upper case variables are random variables and lower case variables are realisations of the corresponding

random variable. Using the product rule of probability, $P(R_1 = r_1, \dots, R_m = r_m, M = m) = P(R_1 = r_1, \dots, R_m = r_m | M = m)P(M = m)$. $P(M = m)$ is given in Eq.(5.8). Since we have assumed that each base is equally likely to mutate during division, the conditional distribution is assumed to be

$$P(R_1 = r_1, \dots, R_m = r_m | M = m) = \prod_{k=1}^m \text{unif}\{1, B\} \quad (5.9)$$

where $\text{unif}\{a, b\}$ denotes the discrete uniform distribution over $[a, b]$ (note an analogous result is provably true in continuous space for Poisson-distributed events [324]). Thus, to sample from the joint distribution, we first draw a Binomial random variable for the total number of mutation events, and the location of each mutation event will be uniformly distributed across all of the bases in the lineage.

In doing this, we are assuming that mutations never affect the same cell at the same base more than once (although the same mutation may affect multiple different cells). This assumption is justified with the following calculation. Throughout the lineage history each base is assumed to be copied 36 times. If we assume a worst-case mutation rate of the upper-bound of our prior in Eq.(5.7), $\lambda_m = 1 \times 10^{-8}$, then the total number of times a single base is mutated is $X \sim \text{Binom}(37 - 2, \lambda_m)$ (the factor of (-2) is due to considerations discussed in Fig. 5.1). It follows that $P(X \geq 2) = 5.9 \times 10^{-14}$. We thus calculate an upper bound on the total number of times bases are mutated more than once across the entire brain of an individual by assuming that every base in every cell has an independent chance of being mutated, and therefore the total number of such events is $2^{37} \cdot 4N_b \cdot P(X \geq 2) \approx 4337$ multiple-mutated bases per individual. This is a strong overestimate since neurons share lineage history, so in effect the factor of 2^{37} is actually much lower. Given that the total number of mutation events in the brain is on average $4N_b \cdot \lambda_m \cdot \sum_{k=0}^{35} 2^k \approx 3.6 \times 10^8$ mutations per individual then, even for this overestimate, bases being mutated more than once per cell is an extremely rare event ($\mathcal{O}(10^{-5})$ of mutation events) and is thus neglected.

A mutation which occurs at generation i^* affects a fraction (f) of the brain (see Eq.(5.2)). Using this observation, once the position of mutations in the lineage have been determined using Eq.(5.8) and Eq.(5.9), the overlap between the daughters of the parental mutant cell and the sample can be determined to find the total number of mutations in sample s from patient p from mutational event j (given the assumption of spatial contiguity of the sampled cells, see above), in the cohort denoted as $X_{j,p,s}$. Note that $\sum_j X_{j,p,s}$ is not necessarily equivalent to the total number of mutated cells per sample, since a single cell may harbour multiple different mutations (e.g. consider the case of the founder cell of the brain containing a mutation, as well

as one of its daughters containing a different mutation, then a subset of the final population of cells harbour two mutations per cell).

To determine whether the sample from the prior λ_m may be accepted as a sample from the approximate posterior distribution, we use summary statistics to compare the simulation with data ($\mathcal{D}$). Experimentally, the threshold for detection of a SNV per sample was $\sim 0.5\%$. Using an average sample size of $N_v = 3415$ cells, this corresponds to a detection threshold of ≥ 17 cells in order for a mutation to be detectable amongst the captured cells of the sample. Before defining the summary statistics, we define the following indicator functions

$$\Phi(x) = \begin{cases} x & \text{if } x \geq 17 \\ 0 & \text{otherwise} \end{cases} \quad (5.10)$$

$$I(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5.11)$$

where $\Phi(x)$ is based on the experimental detection threshold. The three summary statistics $\mathbf{S} = (\Gamma_1, \Gamma_2, \Gamma_3)$ of the model are defined as:

$$\Gamma_1 = \sum_{p=1}^{40} \sum_{s=1}^{s_p} \sum_{j=1}^M \Phi(X_{j,p,s}) \quad (5.12)$$

$$\Gamma_2 = \sum_{p=1}^{40} I \left(\sum_{s=1}^{s_p} \sum_{j=1}^M \Phi(X_{j,p,s}) \right) \quad (5.13)$$

$$\Gamma_3 = \frac{1}{\sum_p \sum_s \sum_j I(\Phi(X_{j,p,s}))} \sum_p \sum_s \sum_j \frac{\Phi(X_{j,p,s})}{3415} \quad (5.14)$$

where s_p is the number of samples taken for patient p . Γ_1 measures the total number of detectable mutations across all samples and all patients, Γ_2 measures the total number of patients with any detectable mutations and $\Gamma_3 = \bar{v}$ is the average variant allele fraction (VAF) for all detectable mutations across all samples and patients. Using a hat to denote data, these summary statistics were found experimentally to be:

$$\hat{\Gamma}_1 = 6 \quad (5.15)$$

$$\hat{\Gamma}_2 = 6 \quad (5.16)$$

$$\hat{\Gamma}_3 = \bar{v} = 0.87 \times 10^{-2}. \quad (5.17)$$

Thus, the acceptance criteria we use are

$$|\Gamma_1 - \hat{\Gamma}_1| \leq \epsilon_{sa} \quad (5.18)$$

$$|\Gamma_2 - \hat{\Gamma}_2| \leq \epsilon_{pa} \quad (5.19)$$

$$|\Gamma_3 - \hat{\Gamma}_3| \leq \epsilon_v \quad (5.20)$$

where $\epsilon_{sa} = \epsilon_{pa} = 1$ and $\epsilon_v = 2\text{SEM}_v$ (SEM_v = standard error in the mean for the VAF across all positive experimental samples), and we demand that all three of the above criteria must be satisfied simultaneously. Given that these criteria are met, we accept λ_m as a draw from the approximate posterior distribution $P(\lambda_m|\rho(\mathbf{S}, \mathcal{D}))$ where ρ denotes the acceptance criteria given above. It is possible that a more compact and/or optimal set of summary statistics could be obtained; furthermore, poor choice of summary statistics may cause the approximate posterior to diverge from the true posterior [325].

5.3.3 FORWARD-SIMULATION TO ESTIMATE THE PREVALENCE OF SOMATIC MOSAICISM IN NEURODEVELOPMENT

After inferring the mutation rate given the data (see above), we will seek to generate estimates of the number of pathologically mutated neurons in the adult human brain. We do this by considering the development of just the neuronal population in the same manner as above, except with 36 doublings yielding $2^{36} \approx 6.9 \times 10^{10}$ neurons. This falls within 2 standard deviations of the estimate of the number of neurons in the brain by Azevedo *et al.* ($86.1 \pm 8.1 \times 10^9$ neurons), and is therefore broadly compatible with that study.

Using the draws from the approximate posterior distribution, we may then forward-simulate each individual using a single draw from $P(\lambda_m|\rho(\mathbf{S}, \mathcal{D}))$ as the mutation rate. We may therefore determine the distribution of mutant neuronal region sizes and the number of mutated neurons per individual. To simulate pathological bases, we found the set of sequenced bases which also appeared in The Human Gene Mutation Database (HGMD) [326]. We found $N_b = 298$ bp which appeared in the HGMD, and take this as the number of potential sites associated with neurodegenerative disease (see Eq.(5.1)).

5.4 RESULTS

5.4.1 A CRUDE MODEL OF NEURODEVELOPMENT PREDICTS THE PREVALENCE OF SOMATIC MOSAICISM IN THE HUMAN BRAIN

We performed approximate Bayesian computation on the deterministic branching process model of neurodevelopment (see Fig. 5.2 and Section 5.3.2). We determined the neurodevelopmental somatic mutation rate to be $\lambda_m = (4.8 \times 10^{-10} - 2.99 \times 10^{-9})$ per base per cell division (95% Bayesian credible interval, BCI), see Fig. 5.3A, which is comparable to the mutation rate found in mitotically active tissues ($\lambda_m = 0.64 \pm 0.11 \times 10^{-9}$ per base per cell division [306]), and the order-of-magnitude calculation presented in Section 5.3.1 (see Eq.(5.6)). We then simulated the neurodevelopment of individuals under the model, given the inferred distribution of mutation rates, to find the frequency and size of brain regions harbouring known pathological mutations in neurodegenerative disease genes [326] (see Section 5.3.3), see Fig. 5.3B-F. The distribution of the number of mutated regions seeded at each generation is shown in Fig. 5.3B. We found that each individual tended to have $10^5 - 10^6$ pathologically mutated cells (Fig. 5.3D, see Section D.3 for an explanation of the multimodality which is driven by the largest mutated region per individual). Each contiguous mutated region consisted of ~ 8 cells on average (Fig. 5.3C). We found that 10.8% ($\pm 0.1\%$) of simulated individuals possessed one or more regions of 2.62×10^5 spatially contiguous pathologically mutated cells (Fig. 5.3E). Furthermore, each individual brain was found to harbour 75-481 regions (95% BCI) of 128 pathologically mutated cells (Fig. 5.3F). A visualisation of a typical simulated individual is shown in Fig. 5.4

5.4.2 MODELLING OF NEURODEVELOPMENT AS A BINARY TREE IS ROBUST TO SEVERAL MODEL PERTURBATIONS

In the following subsections, we perturb the simple deterministic branching process to take into account cell death; founder effects from progenitor cells; asymmetric division; and cellular migration during neurodevelopment. We present a simple argument for why many of our observations are robust to such perturbations in Section 5.4.2.4.

5.4.2.1 BRANCHING MODEL AND CELL DEATH

In the branching model of neurodevelopment, we neglected cell death for simplicity. However, programmed cell death plays an important role in neurodevelopment (reviewed in Ref. [327]).

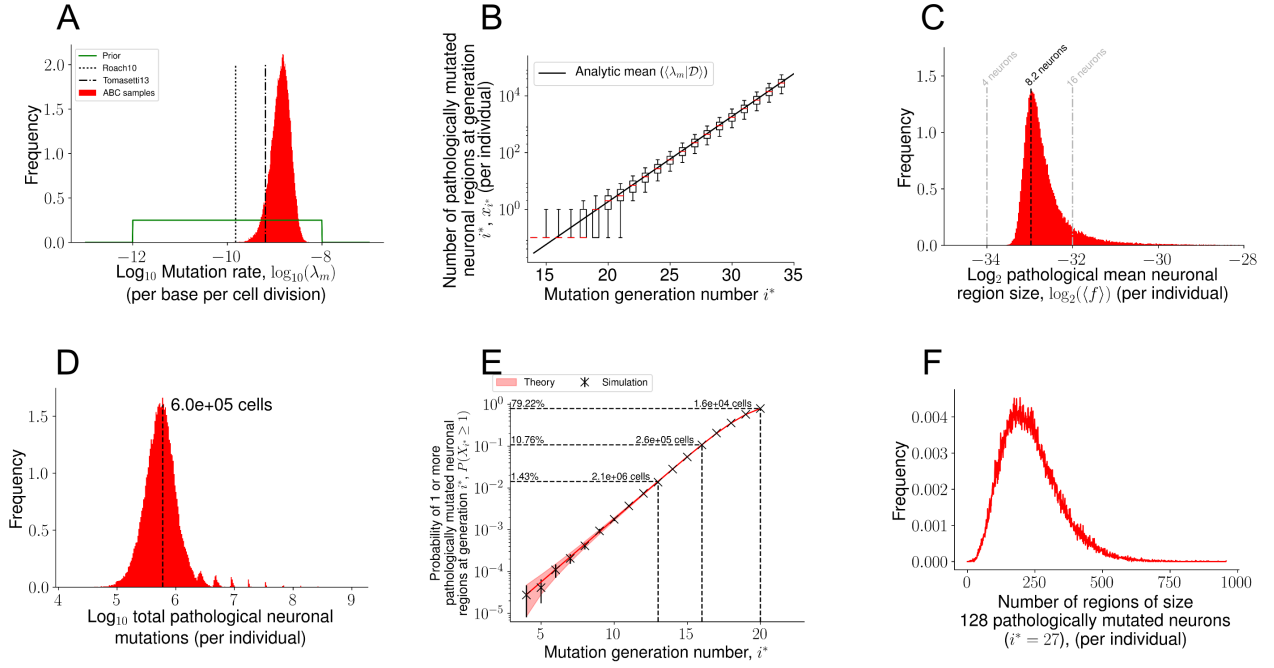


Figure 5.3: Estimates for the prevalence of somatic mutation during human neurodevelopment. (A) Prior (green solid) and approximate posterior (red bars) for the mutation rate using approximate Bayesian computation (5×10^7 iterations), with literature values (black lines). (B-F) Using the approximate posterior samples for λ_m we simulated brain development of the neuronal subpopulation (see Section 5.3.3). (B) Distribution of the number of pathologically mutated regions per individual seeded at generation number i^* . Boxplot whiskers show the 5th and 95th percentiles, median shown in red (when an individual has 0 mutations at generation i^* , the number of mutations is assigned to be 0.1, for representation on a log-scale). Showing the analytic mean of Eq.(5.1), where λ_m is the approximate posterior mean (black solid line). (C) Distribution of the pathological mean relative region size across individuals. (D) Distribution of total number of pathological mutations across individuals. Multimodality is induced by the largest pathological region, see Fig. D.1. (E) Fraction of simulated human individuals with at least 1 pathologically mutated region seeded at generation i^* . Comparison to theory shown in red, mean defined in Eq.(D.4) and standard deviation Eq.(D.5), see Section D.1 (error bars are a Bernoulli error model). (F) Frequency distribution over individuals of pathologically mutated regions of size 128 neurons.

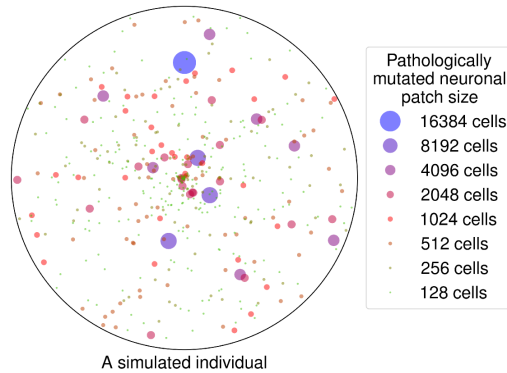


Figure 5.4: Representation of the size and frequency distribution of pathologically mutated regions in an individual human brain. Displaying mutations seeded between $20 \leq i^* \leq 27$. Whole brain area (black circle) is not to scale with the mutated regions (coloured circles). See Section D.2 for details.

We made the following modification to the model above as a preliminary means of investigating the extent to which programmed cell death may influence our estimations. We note that the following model of cell death is, itself, a marked simplification of real cell death during neurodevelopment, but benefits from its parsimony and computational tractability.

It is thought that 20-80% of neurons are lost during neurodevelopment, and that this wave of cell death often occurs at a relatively late stage in maturation [328]. To incorporate cell death into our model, and account for uncertainty in the level of cell death during development, we allow the brain to double in size for an additional number of generations, $g = 1, 2, \dots$, resulting in a brain which is a factor of 2^g larger than a normal adult brain; we will term this a “super-brain”. We then infer the mutation rate by taking a sample from the super-brain which is a factor of 2^g larger than the experimental sample size (3415 cells); we term this a “super-sample”. Note that we retain the assumption of spatial contiguity of the cells of the super-sample from the original model. We then sample a fraction of $1/2^g$ cells from the super-sample (without replacement) as cells which survived to adulthood. In doing this, we always sample 3415 cells, but the spatial location of these cells within the super-sample is random. In sampling the tree after a phase of additional growth, we are effectively placing cell death at the leaves of the tree: an approximation which is more appropriate assuming that the bulk of cell death occurs late in development. We then proceed with the approximate Bayesian computation as discussed in Section 5.3.2. Consequently, this results in a model where a fraction $(1 - 1/2^g)$ of all cells die (see Fig. 5.5A). By computing the posterior distribution of the mutation rate λ_m for different values of g , we may investigate the robustness of the mutation rate to this simple model of cell death.

We investigated values of $g = 0, 1, 2$ and 7 for death rates of 0% (Section 5.3.2), 50%, 87.5% and 99.2% respectively. In Fig. 5.5B, we show that the approximate posterior distributions for the mutation rate have large overlap for different magnitudes of cell death. This result has an intuitive explanation given the following two observations about the model. Firstly, a mutation occurring at generation i^* affects a *fixed proportion* of the brain ($f = 2^{-(i^*+2)}$), regardless of how large the brain subsequently becomes. Secondly, we grow the brain to be a factor of 2^g larger than a normal adult brain, but use a survival rate of $1/2^g$ to ensure that number of cells which survive to adulthood is always the number of cells observed in a real adult brain, on average. Since the fraction of the brain that is mutated, and the total number of cells in an adult brain, are invariant to g , we would expect that the total number of cells which are mutated in an adult brain to be invariant to g (and therefore are all compatible with the mutation rate from Tomasetti *et al.* [306] and Eq.(5.6)). It is therefore unsurprising that the inferred values of

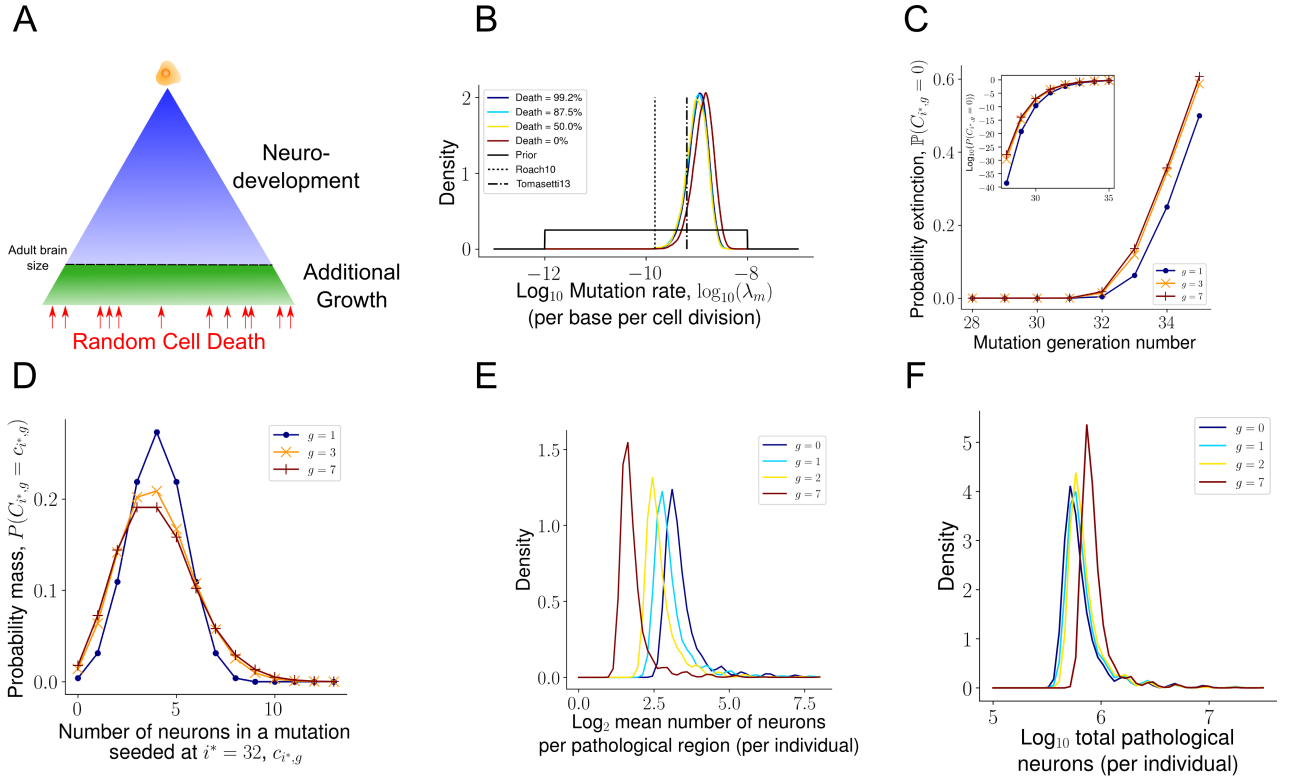


Figure 5.5: Investigation of a simple model of cell death during neurodevelopment. (A) We model cell death during neurodevelopment by allowing additional generations of growth ($g = 1, 2, \dots$) (green shaded), with stochastic cell death at the end of the additional growth (red arrows) such that the final number of cells is the number of cells in the adult human brain. (B) The inferred mutation rate was found to be broadly unchanged for $g = 1, 2, 7$ relative to the original model without cell death ($g = 0$). (C) The probability of extinction of a mutant region via cell death becomes negligible for mutation generation numbers $i^* \leq 32$. Inset displaying the same plot on a logarithmic scale for the vertical axis. (D) Distribution of the number of cells in a mutant region seeded at $i^* = 32$ for differing levels of cell death. (E) Mean number of neurons per pathological region for differing levels of cell death. (F) Total number of pathological neurons per individual for differing levels of cell death. In (E-F), we assume that mutant regions are diffuse in the adult brain.

λ_m are broadly invariant to g . The distributions are not identical because our crude model of cell death affects the size of mutational foci. Although the mean size of mutational foci given non-zero cell death is equivalent to the original, deterministic, case (see Eq.(5.23) below) these foci may fluctuate in size, as will be investigated below.

Given that the mutation rate is broadly invariant to the cell death rate, we suggest that many of our extrapolations to the human population (Figs. 5.3B-F) will also remain broadly invariant to this model of cell death. Whilst it is not straightforward to show this numerically, we provide some arguments for why this might be the case.

Firstly, consider a mutation seeded at generation i^* , which gives rise to $N_{i^*,g}$ daughter neurons prior to the cell death model, where

$$N_{i^*,g} = 2^{-(i^*+2)} \cdot 2^g \cdot N_{\text{neu}} \quad (5.21)$$

where N_{neu} = number of neurons in the brain $\approx 2^{36}$. When cell death is initiated in the model, N_{neu} cells are sampled without replacement from the super-brain of size $2^g N_{\text{neu}}$. These are the cells which survive to adulthood. We let $C_{i^*,g}$ be the number of surviving mutant daughter neurons seeded at generation i^* . $C_{i^*,g}$ may be modelled with a hypergeometric distribution. If $Z \sim \text{Hypergeom}(n, N, K)$, then Z is a random variable whose value is the number of successes in n draws taken without replacement from a population of size N where K objects have the feature of interest. Hence,

$$C_{i^*,g} \sim \text{Hypergeom}(N_{\text{neu}}, 2^g N_{\text{neu}}, 2^g N_{\text{neu}} 2^{-(i^*+2)}) \quad (5.22)$$

Using the properties of the hypergeometric distribution, the mean number of neurons which survive to adulthood is

$$\mathbb{E}(C_{i^*,g}) = N_{\text{neu}} 2^{-(i^*+2)}, \quad (5.23)$$

which is independent of g , as expected. This is equivalent to the number of mutations in the adult brain associated with a mutation event at generation i^* in the original model (see Section 5.3.2).

Given the above argument, we now examine each extrapolation in Figs. 5.3B-F and explore their robustness to this simple model of cell death:

NUMBER OF PATHOLOGICALLY MUTATED NEURONAL REGIONS SEEDED AT GENERATION i^* (FIG. 5.3B) We expect, according to our model of death and replication, that the distribution

of the number of pathologically mutated neuronal regions seeded at generations $i^* \leq 32$ (X_{i^*} , see Eq.(5.1) and Fig. 5.3B) remains unchanged. The dynamics of neurodevelopment under our cell death model are equivalent up to generation 36. These dynamics are determined by a single parameter, the mutation rate λ_m , the inference of which we have shown to be broadly invariant to cell death (see Fig. 5.5B), and comparable to Ref. [306] and Eq.(5.6). Hence, the number of mutations up to generation 36 in the cell death model (as determined by Eq.(5.1)) is expected to be similarly distributed in the cell death model, prior to the cell death process occurring.

However, the death process may stochastically eliminate mutational foci. In Fig. 5.5C, we show that for generations $i^* \leq 32$, corresponding to mutational foci of size $\gtrsim 4$ (see Fig. 5.5D for the distribution of sizes for regions seeded at $i^* = 32$), the probability of extinction of such foci through cell death is negligible under this model of cell death ($< 2\%$ for $g \leq 7$). Hence, the amount of probability mass for X_{i^*} in Fig. 5.3B which may be affected by cell death is upper-bounded by 2% for generations $i^* \leq 32$. As shown in the inset of Fig. 5.5C, the probability of extinction decays extremely quickly as i^* decreases.

MEAN SIZE OF PATHOLOGICALLY MUTATED NEURONAL REGIONS (FIG. 5.3C) We explore how the mean size of pathological neuronal mutations (Fig. 5.3C) is affected by this model of cell death, using the following model. If we assume that mutant regions in the brain are dilute enough such that no two mutated regions overlap (discussion below), we may write the total number of pathological cells (T_g) (Fig. 5.3D), for a given level of cell death, as

$$T_g = \sum_{i^*} \sum_{j=1}^{X_{i^*}} C_{i^*,g,j} \quad (5.24)$$

where $C_{i^*,g,j}$ is a random variable with the same distribution as $C_{i^*,g}$ (see Eq.(5.22)), with an additional index j to label each mutation event at generation i^* . X_{i^*} is the number of pathological mutations arising at generation i^* , see Eq.(5.1) where $N_b = 298$ bp is the number of pathological bases in the case genes [326]. Eq.(5.24) sums all surviving cells from every mutation event j , over all generations i^* , for a given level of cell death g . The mean size of all surviving pathological regions (S_g) is therefore

$$S_g = \frac{T_g}{\sum_{i^*} \sum_{j=1}^{X_{i^*}} I(C_{i^*,g,j})} \quad (5.25)$$

where $I(x)$ is defined in Eq.(5.11). In Fig. 5.5E, we show that (using a typical value of $\lambda_m = 1.5 \times 10^{-9}$ mutations per base per division, which is comparable to the literature value from

Tomasetti *et al.* [306] which is 0.64×10^{-9} mutations per base per cell division and also Eq.(5.6)) the mean number of neurons per pathological region reduces with increasing cell death. Given non-zero cell death, the additional generations may result in mutant foci which are on average smaller than 1 neuron (due to Eq.(5.23), where $N_{\text{neu}} = 2^{36}$ and $i^* > 34$ when $g > 0$). Hence, single-neuron mutational foci are more common given cell death, which reduces the expected size of mutant foci, and so Fig. 5.3C is sensitive to this model of cell death.

Eqs.(5.24)-(5.25) rely upon the assumption that mutations are sufficiently dilute such that no two mutations overlap (otherwise cell death may remove a cell which is e.g. doubly mutated: this is not accounted for in Eq.(5.24)). However, we observe in Fig. 5.5F that, for the ranges of cell death explored here, the majority of individuals have fewer than $\sim 10^6$ mutant neurons, corresponding to $\sim 0.0015\%$ of all neurons. Hence, our assumption of dilute mutant regions is reasonable.

TOTAL NUMBER OF PATHOLOGICAL NEURONS (FIG. 5.3D) We find that the distribution of the total number of pathologically mutated neurons per individual (see Fig. 5.3D) is significantly increased with non-zero cell death (see Fig. 5.5F, comparing $g = 7$ against $g = 0$ on a linear scale, $p < 10^{-12}$ one-sided Mann-Whitney U-test). This is also likely due to the additional rounds of growth introducing more single-cell mutations. However, the order of magnitude is largely unaffected ($\sim 10^6$).

PROBABILITY OF 1 OR MORE PATHOLOGICALLY MUTATED REGIONS AT GENERATION i^* (FIG. 5.3E) We expect the probability of 1 or more pathologically mutated regions at generation i^* , for $i^* < 20$, (see Fig. 5.3E) remains unchanged, see the above discussion of robustness for Fig. 5.3B.

NUMBER OF REGIONS PER INDIVIDUAL OF 128 PATHOLOGICALLY MUTATED CELLS (FIG. 5.3F) We expect the number of regions per individual consisting of approximately 128 pathologically mutated cells (see Fig. 5.3G), to remain unchanged given non-zero cell death. Using the properties of the hypergeometric distribution (Eq.(5.22)), the probability of a region seeded at $i^* = 27$ to go extinct when $g = 1$ (a plausible value, see [328]) is 10^{-78} , hence we expect all such regions to remain in the adult brain under this model of cell death. The size of such regions is 112-144 neurons (95% confidence interval), hence the distribution is sufficiently constrained for 128 neurons to be indicative of the size of regions seeded at this generation.

To conclude, our central claim that somatic mosaicism is expected to exist in the adult human brain given our data and modelling of neurodevelopment as a binary tree is expected to be robust when including this simple model of cell death.

5.4.2.2 INITIAL CONDITIONS FOR NEURODEVELOPMENT

In our deterministic branching model of neurodevelopment, we assumed that neurodevelopment begins with a single founder cell. This is a clear simplification, since the central nervous system is derived from the ectoderm of the embryo, whereas the other germ layers give rise to other tissues and organs [329]. As a preliminary means of understanding the effect of neurodevelopment beginning from an initial founder pool of cells, we may modify our original model of neurodevelopment (see Section 5.3.2), such that neurodevelopment begins from a set of e.g. $2^{10} = 1024$ cells which are assumed to be without mutation; we refer to this modified model as the “progenitor pool model”.

From Fig. 5.3E, we see that mutations occurring at generations $i^* \leq 10$ in the original model affect on the order of 1 in 1000 individuals, or fewer. The progenitor pool model differs from the original model only through enforcing that mutations in generations $i^* \leq 10$ occur with 0 probability (see Fig. 5.6A). Since mutations at $i^* \leq 10$ were rare in the original model, it is reasonable to assume that the distribution of inferred mutation rates for the progenitor pool model is highly similar to the original model (Fig. 5.3E). We therefore use the same distribution of mutation rates in Fig. 5.3A for the progenitor pool model (which are comparable to the mutation rate from Ref. [306] and Eq.(5.6)), and consider the consequences of this alternative model through forward-simulation.

- Number of pathologically mutated regions (compare to Fig. 5.3B): Since the progenitor pool model is identical to the original model from generations $i^* = 11$ onwards, the analogue of Fig. 5.3B for the progenitor pool model is identical.
- Mean pathologically mutated region size (compare to Fig. 5.3C): The distribution for the mean pathologically mutated region size is almost unchanged under the progenitor pool model (Fig. 5.6B). This is because there are exponentially more mutation events as the generation number i^* increases, which dominates the mean region size in both models.
- Total number of pathological mutations (compare to Fig. 5.3D): The modal value for the total number of pathological mutations per individual shifts from 6.0×10^5 cells per

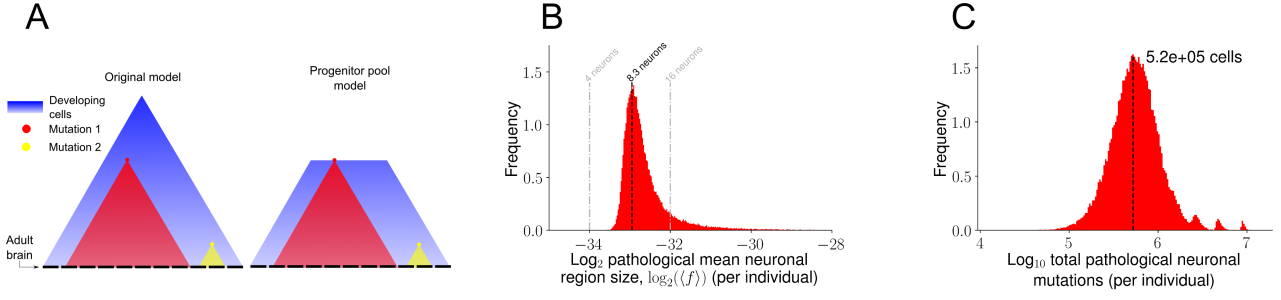


Figure 5.6: Investigation of the progenitor pool model. (A) In the progenitor pool model, neurodevelopment is assumed to begin with $2^{10} = 1024$ unmutated cells, which is followed by subsequent rounds of symmetric cell division. Note that, by construction of the model, a mutation which affects e.g. one of the progenitors in the progenitor pool model (red dot, right) has the same mutant region size in the final adult brain as a mutation occurring at generation 10 in the original model (red dot, left). This is the case for all mutations after generation 10 (e.g. yellow dots). (B) The mean pathological region size is almost unchanged relative to the original model (see Fig. 5.3C). (C) The total number of pathologically mutated neurons per individual is almost unchanged relative to the original model (see Fig. 5.3D).

individual in the original model (Fig. 5.3D) to 5.2×10^5 cells per individual (Fig. 5.6C) in the progenitor pool model; however, these distributions show broad overlap.

- Probability of 1 or more mutations at generation i^* (compare to Fig. 5.3E): The analogue of Fig. 5.3E for the progenitor pool model is for all mutations at generations $i^* \leq 10$ to be set to $P(X_i \geq 1) = 0$. However, the observation of e.g. 11% of individuals possessing 3×10^5 spatially-contiguous pathologically mutated cells is robust under this simple modification of the original model, since this corresponds to $i^* = 16$ and therefore follows the same distribution as the original model.
- Number of pathological regions of size 128 (compare to Fig. 5.3F): This observation is unchanged since mutations occurring at $i^* = 27$ are independent of generations $i^* \leq 10$.

Overall, assuming that neurodevelopment begins with a population of ~ 1000 unmutated cells (or fewer) results in similar predictions (both quantitatively and qualitatively) to the original deterministic branching process for neurodevelopment. This is because, in the original model, if a mutation has not occurred by generation 10, the progenitor model is entirely equivalent to the original model. This can be seen in Fig. 5.6A, where a mutation which occurs at the point where the developing brain is the same size between the two models affects an identical amount of the final adult brain (i.e. the red and yellow shaded regions have the same area in both models). In the original model, $\sim 1/1000$ individuals or fewer experience a mutation within the first 10 divisions (Fig. 5.3E). So, for $\sim 99.9\%$ of individuals, the progenitor model is equivalent to the original model. These 99.9% are in such a vast majority that they wash out

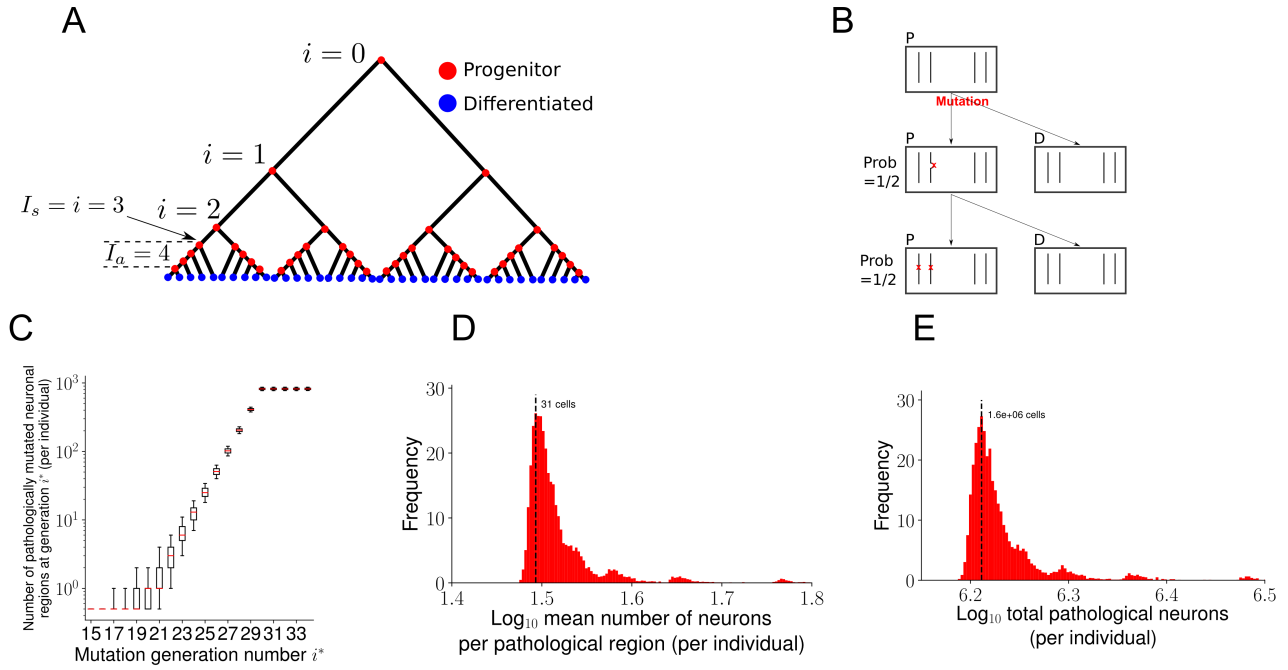


Figure 5.7: Investigation of a simple asymmetric model of neurodevelopment. (A) Neurodevelopment is modelled as consisting of a phase of symmetric division, followed by a phase of asymmetric division. (B) Asymmetric division induces a stochastic element to the inheritance of a mutation. Showing an example of a mutation event where a copying error occurs in a progenitor (P) cell (top) which is inherited by a daughter progenitor cell (middle); this occurs with probability $1/2$. The single-stranded mutation is copied, and a double-stranded mutation is inherited by another progenitor cell (bottom) (probability $1/2$). This mutation, which is now on both strands of DNA, will be inherited by all subsequent daughters (both P and D). Either of the preceding events could have been inherited by differentiated cells (D), which do not divide, and therefore affect a greatly restricted number of cells. (C) The distribution of pathologically mutated neuronal regions initially increases exponentially until the asymmetric division phase at generation 30, where the number of mutant regions remains constant thereafter. (D) Mean number of mutated neurons per mutated region, with a mode of 31 cells. (E) Total number of mutations per individual, with mode of 1.6×10^6 cells.

the statistical signal from the 0.1% who are differentially affected between the two models, as shown in Fig. 5.6B&C.

5.4.2.3 ASYMMETRIC DIVISION IN NEURODEVELOPMENT

In our original model of neurodevelopment (see Section 5.3.2), we assumed that every cell division resulted in two daughter cells which were both able to divide (symmetric division) until the final cell division. However, it is commonly thought that neurodevelopment proceeds in two phases: a proliferative phase mostly consisting of symmetric division, and a differentiative phase where each progenitor cell undergoes sequential rounds of asymmetric division [330]. In this case, we define asymmetric division as consisting of a progenitor cell dividing into another progenitor cell and a differentiated cell, where the differentiated cell ceases to divide.

We may modify the original model to include a simple model of asymmetric cell division, see

Fig. 5.7A. In this asymmetric model of neurodevelopment, let I_s be the generation number of the final symmetric division, and I_a the total number of divisions which give rise to differentiated cells (asymmetric divisions plus the final division, where progenitor $\rightarrow$ differentiated + differentiated). Again assuming that we are modelling only the neuronal subpopulation, then the total number of neurons (N_{neu}) is

$$N_{\text{neu}} = 2^{I_s}(I_a + 1). \quad (5.26)$$

In Fig. 5.7B we see that asymmetric division induces a stochastic element to the number of daughters affected by a mutation occurring at generation i^* . Since asymmetric divisions affect the tips of the neurodevelopmental tree, they are consequently the most numerous events: it is therefore important that these possible stochastic paths are properly accounted for. There are three cases to consider for when a mutation occurs, and the corresponding number of mutated daughters (N_{i^*}) which the mutation gives rise to. Firstly,

$$N_{i^*} = 2^{I_s - (i^* + 2)}(I_a + 1) \quad \text{if } i^* \leq I_s - 2. \quad (5.27)$$

In this case, a mutation is guaranteed to be inherited by a progenitor cell on both strands of one of its chromosomes (see Fig. 5.1), and affect all subsequent daughters. The second case corresponds to

$$N_{i^*} = \begin{cases} I_a & \text{w.p. } 1/2 \\ 1 & \text{w.p. } 1/2 \end{cases} \quad \text{if } i^* = I_s - 1 \quad (5.28)$$

(where w.p. stands for “with probability”). Upon copying the parental DNA, the base at which a mutation arises initially affects only a single strand of DNA: this mutation must be copied into the subsequent generation for it to affect both strands of a single chromosome (see Fig. 5.1). In the case denoted by Eq.(5.28), the single-stranded mutation is guaranteed to be inherited by a progenitor cell but, upon subsequent copying, the double-stranded mutation can either be inherited by a progenitor cell (affecting I_a subsequent daughters) or a single differentiated daughter cell. These two possibilities occur with equal probability. The third, and final, case is

$$N_{i^*} = \begin{cases} 0 & \text{w.p. } 1/2 \\ 1 & \text{w.p. } 1/4 \\ I_s + I_a - (i^* + 1) & \text{w.p. } 1/4 \end{cases} \quad \text{if } i^* \geq I_s. \quad (5.29)$$

In this case, a single-stranded mutation may be inherited by a differentiated cell w.p. $1/2$,

meaning that the single-stranded mutation is never subsequently copied (as it resides in a non-dividing cell). We do not count mutations which only appear on a single strand of DNA for simplicity, hence the number of mutated cells which arise from this event is 0, w.p. $1/2$. However, if the single-stranded mutation is inherited by a progenitor, upon subsequent copying, the double-stranded mutation may either be inherited by another progenitor or by a differentiated cell (each with probability $1/2$). If the double-stranded mutation is inherited by a differentiated cell, then we count 1 mutant. Otherwise, the mutation will affect all subsequent progenitors and differentiated cells from that lineage, affecting $(I_s + I_a - (i^* + 1))$ daughters in total. Since these final two possibilities are conditioned upon a single-stranded mutation being inherited by a progenitor (w.p. $1/2$), the overall probability of these two possibilities is $1/4$. Note that the three outcomes described above are therefore multinomially distributed.

In the monkey foetus, asymmetric cell division lasts for 30-60 days [331]. Based on this, we use a liberal estimate of 12 weeks for the duration of asymmetric division, and a cell division time of 24 hours, yielding $I_a = 84$ asymmetric divisions. Letting $N_{\text{neu}} = 86.1 \times 10^9$ [323] and rearranging Eq.(5.26) for I_s yields $I_s = 30$ (corresponding to a 6% overestimate in the number of neurons according to [323]).

We use the literature value of $\lambda_m = 6.4 \times 10^{-10}$ per base per cell division [306] as the mutation rate in the discussion below for where we have performed forward-simulation (10^4 simulated individuals). The mutation rate used from Tomasetti *et al.* [306] is broadly compatible with the distribution found in Fig. 5.3A, and also Eq.(5.6). We now investigate the implications of the asymmetric model of neurodevelopment relative to the original model presented above.

- Number of pathologically mutated regions (compare to Fig. 5.3B): In Fig. 5.7C, we find that the number of pathologically mutated regions increases exponentially up to generation I_s . This is in agreement with the original model (Fig. 5.3B), since the models are identical in structure up to generation I_s . The variance is smaller since we have used a single value for the mutation rate, rather than a distribution. After generation I_s , the number of pathological mutations at each subsequent generation is identically distributed (as shown by the plateau for $i^* \geq I_s$) because each generation number has an identical number of cells associated with it.
- Mean pathologically mutated region size (compare to Fig. 5.3C): The modal value for the mean number of neurons per pathological region is 31 neurons, as opposed to 8 neurons in the original model (Fig. 5.3C). This is expected because under the original model there are exponentially more cells in later generations, which give rise to smaller mutational

foci. However, in the asymmetric model, there are the same number of cells per generation during the differentiative phase, and therefore small foci are not as heavily represented in the asymmetric model.

- Total number of pathological mutations (compare to Fig. 5.3D): In Fig. 5.7E, we see that the most probable number of pathologically mutated neurons per individual is 1.6×10^6 , which is 2.7 times greater than the original model (6×10^5 , Fig. 5.3D); this can be accounted for, in part, by the observation that the asymmetric model has a value of N_{neu} which is 30% larger than the original model. We note that the predictions for the total number of mutations in both models are comparable in order of magnitude.
- Probability of 1 or more mutations at generation i^* (compare to Fig. 5.3E): Since the asymmetric model is identical to the original model up to generation $I_s = 30$, Fig. 5.3E also holds for the asymmetric model, albeit with different numbers of mutated cells corresponding to each value of i^* . Substituting into Eq.(5.27), 79% of individuals are expected to possess at least 1 region consisting of 2.2×10^4 spatially contiguous pathologically mutated cells, 11% of individuals possess a region of 3.5×10^5 cells, and 1% of individuals possess a region of 2.8×10^6 cells. These values are all comparable in magnitude to those displayed in Fig. 5.3E.
- Number of pathological regions of size 128 (compare to Fig. 5.3F): Since generation $i^* = 27 < I_s - 2$, the distribution for the number of regions seeded at generation $i^* = 27$ is identical to the original model (given the same set of mutation rates, see Fig. 5.3F). Note that, under the asymmetric model, mutations seeded at generation $i^* = 27$ correspond to mutant regions consisting of 170 cells.

Hence, under a simplified asymmetric model of neurodevelopment, the total number of pathological mutations is approximately doubled and the mean mutated region size is approximately tripled. Overall, this simple asymmetric model of neurodevelopment predicts a prevalence of somatic mosaicism which is similar in order of magnitude to the original model presented in Fig. 5.3.

5.4.2.4 INTUITION FOR MODEL ROBUSTNESS

In Section 5.4.2.1, Section 5.4.2.2 and Section 5.4.2.3, we found that many of our quantitative and qualitative observations regarding somatic mosaicism during neurodevelopment were robust to simple alterations in the model structure, which addressed issues of cell death, progenitor

founder effects and asymmetric division. Here, we provide an order-of-magnitude calculation for why this is expected to be the case. Our strategy will be to determine the size of the brain required for 1 pathological mutation to occur on average, and then understand how many daughter cells in the adult brain such a mutation gives rise to.

Suppose at some instance in time (t) during neurodevelopment there exist $N(t)$ cells. Suppose $N(t)$ is large enough such that, upon replication of each of the $N(t)$ cells once, 1 pathological mutation occurs on average. This scenario may still be modelled with a simple binomial distribution. We may therefore apply Eq.(5.1), where we substitute 2^{i*} with $N(t)$ and set $\mathbb{E}(X_{i*}) = 1$ (where X_{i*} is a random variable denoting the number of mutations which arise in copying $N(t)$ cells). This yields

$$1 = 4N_b N(t) \lambda_m. \quad (5.30)$$

Substituting $\lambda_m = 6.4 \times 10^{-10}$ mutations per base per doubling [306] and $N_b = 298$ bp (mutations in which are pathological [326]) yields $N(t) = 1.3 \times 10^6$ cells. We neglect subtleties relating to lags associated with copying each strand of DNA (see Eq.(5.2) and Fig. 5.1). We now assume the following

1. We require *unbiased* replication of all daughters *after* the mutation event when the brain consisted of $N(t)$ cells (unbiased meaning that each cell gives rise to the same number of daughter cells).
2. We require daughters to remain spatially proximal during this replication.

Given assumption (1), we expect the fraction $\sim 1/N(t)$ of all subsequent cells to carry this mutation. (Note that accounting for the copying lag will reduce this to $\sim 1/4N(t)$, but is on the same order of magnitude.) Using $N_{\text{neu}} = 86.1 \times 10^9$ neurons results in the prediction that every individual harbours, on average, 1 pathologically mutated region consisting of approximately N_{mut} spatially-contiguous cells, where

$$N_{\text{mut}} = N_{\text{neu}}/N(t) = 6.6 \times 10^4 \text{ cells}, \quad (5.31)$$

and spatial contiguity follows from assumption (2). Increasing the value on the left hand side of Eq.(5.30) (i.e. the number of mutation events at a particular stage of neurodevelopment) increases $N(t)$ and reduces N_{mut} according to Eq.(5.31). This is the qualitative essence of somatic mosaicism: mutations which are more common are smaller, and vice versa.

No details relating to the process which generated the $N(t)$ cells were assumed, so founder effects

are unimportant prior to when the cell consisted of $N(t)$ cells. All of the models explored above satisfy these two assumptions, so this order-of-magnitude calculation is expected to approximately hold.

We may check this argument by examining Fig. 5.3E at $i^* = 20$, since all of the models described in Section 5.4.2.1, Section 5.4.2.2 and Section 5.4.2.3 satisfy Fig. 5.3E at $i^* = 20$. This states that 80% of individuals are expected to carry mutations of size 1.6×10^4 cells. This observation is consonant with the above argument to within an order of magnitude. We therefore expect models of neurodevelopment which satisfy the above 2 properties to show somatic mosaicism where it is common for individuals to have mutant regions of $\sim 10^4$ – 10^5 cells, given a mutation rate between 10^{-10} – 10^{-9} mutations per base per division. This argument also applies to combinations of the above model perturbations, e.g. a model which begins with an initial set of 2^{10} cells (Section 5.4.2.2) with asymmetric division after generation 30 (Section 5.4.2.3), because such models still satisfy conditions (1) and (2) above. A notable example of where this argument does not apply is in the case of large cellular migration, which we now discuss.

5.4.2.5 CELL MIGRATION AND SPATIAL CORRELATIONS IN NEURODEVELOPMENT

In the branching model of neurodevelopment, we assumed that cells which were in close spatial proximity shared a similar developmental lineage. However, cells have the ability to migrate during development [332]. If cells have appreciable migration during development, then this reduces spatial correlations in the mutational profile of brain cells.

The extent of cell migration informs upon the degree to which our original model is perturbed from true neurodevelopment. It has been suggested that neurons are established by radial glial cells, forming functional neocortical columns [333]. These columns are between 200–1000 μm across [334]. If most migration were to occur on the micron scale, our experimental approach would be largely insensitive to this, since each sample was approximately 1 cm^3 in size and homogenised. However, tangential migration has also been observed [332], where cells move across different subdivisions of the forebrain.

In Section D.4, we provide an alternative perspective to neurodevelopment where the brain consists of independent regions of some characteristic size. Such a model better represents the diffusive aspects of neurodevelopment, although makes other assumptions which are arguably less plausible than the deterministic branching process presented above (see Section 5.3.2). We note that we find the total number of pathological mutations in all of the models investigated

here to be on the order of 10^5 - 10^6 cells (see Fig. 5.3D, Fig. 5.5F, Fig. 5.6C, Fig. 5.7E and Fig. D.2E).

5.5 DISCUSSION

In this work, we have modelled neurodevelopment as a deterministic branching process, and used data from a recent ultra-high depth sequencing study [199] to infer the distribution of nuclear DNA mutation rates. The inferred mutation rates were found to be broadly comparable to existing literature values. Through forward simulation from the model, we were then able to estimate the prevalence of somatic mosaicism in genes associated with neurodegenerative disorders. We found that assuming a tree structure for neurodevelopment with a constant mutation rate induced spatially contiguous islands of pathologically mutated cells in the adult brain, where smaller islands were more frequent and vice versa. We found that $\sim 10.8\%$ of simulated individuals possessed one or more regions of $\sim 10^5$ spatially contiguous pathologically mutated cells. We explored the robustness of our findings by perturbing our model to account for simple models of cell death, progenitor cells, and asymmetric division, finding that our results were largely robust on an order-of-magnitude scale. A simple argument assuming a mutation rate on the order of 10^{-10} - 10^{-9} mutations per base per cell division (which is supported by literature estimates), with unbiased, spatially proximal, replication of daughter cells suggests that every individual is expected to possess 1 mutated region associated with neurodegeneration, consisting of 10^4 - 10^5 cells in their brain.

A natural area for future development of our inference approach would be to relax the assumption of spatial contiguity of the captured DNA of the sample, allowing for the sample size to correspond to the full 1 cm^3 block of tissue. However, it is not obvious that altering the sample size will have any effect upon the inferred mutation rate, since the simple argument in Section 5.3.1 suggests that the mutation rate is independent of the sample size. Furthermore, the experimental data in Ref. [199] does not directly observe any instances of pathologically mutated cells, although large islands consisting of $10^8 \cdot \bar{v} = 8.7 \times 10^6$ mutated cells were directly observed when considering all of the case genes (since a sample consisting of 1 cm^3 of tissue contains 10^8 cells, and $\bar{v} = 0.87 \times 10^{-2}$). Direct observation of pathologically mutated brain regions would provide the strongest evidence for pathological somatic mosaicism. The calculation in Section 5.3.1 allows us to estimate the total number of samples required to directly observe a pathological mutation using simple arguments. Rearranging Eq.(5.5) for $1/p_{\text{mv}}$, where $1/p_{\text{mv}}$ is a characteristic number of samples required to measure 1 mutated region, and using

$\bar{v} = 0.5 \times 10^{-2}$ (measurement sensitivity), $\lambda_m = 0.64 \times 10^{-9}$ per base [306], and $N_b = 298$ bp [326], shows that the number of samples required to directly observe a pathologically mutated sample is on the order of 10^5 samples. Although this is an experimentally challenging number of samples, expanding the set of genes of interest (i.e. increasing N_b), or technological advances which increase the detection sensitivity (i.e. reduce $\bar{v}$) may allow for the possibility of direct observation of pathological somatic mosaicism in the brain.

Even given the direct observation of pathologically mutated regions of brain tissue from potential future studies, it remains open as to whether risk of developing neurodegenerative diseases can be explained by somatic mosaicism. Mechanistic studies investigating the size of a mutated region of neurons required to induce neurodegenerative disease, or neuronal dysfunction, would test the plausibility that somatic mosaicism can cause neurodegeneration in later life. The approaches discussed here can be extended to a variety of nuclear sites of interest, for various tissues. It is possible that somatic mosaicism in nuclear genes linked with mtDNA maintenance (e.g. *POLG*, *PINK1* and *PRKN*) could induce islands of mutated mtDNA within somatic tissue, potentially causing pathological phenotypes, as discussed in Chapter 1 (see Section 1.2.1.3).

6

SUMMARY

In this thesis, we have explored how sources of mitochondrial heterogeneity impact upon cell physiology in healthy and diseased settings. We have used mathematical modelling throughout this work as a means to generate self-consistent, quantitative, descriptions of experimental observations. Where data has been limited in quantity, we have favoured parsimonious models, whilst emphasising biophysical plausibility. Taking a stochastic perspective has also allowed us to gain deeper understanding of mitochondrial contributions to cellular noise.

In Chapter 2, we present the first analytically derived link between mitochondrial genetic and network dynamics via heteroplasmy variance (Fig. 2.1, Eq. (2.11)). For a chemical reaction network representing mitochondrial network and genetic dynamics, we show that mitochondrial fusion serves to rescale time. This is because only small mitochondrial fragments are susceptible to mitophagic degradation [164], thereby slowing the rate of random genetic drift. We show that our results hold for a range of plausible mtDNA control modes.

The potential neutrality of mitophagy and biogenesis of mtDNA with respect to genotype is particularly relevant for the work presented in Chapter 2. The extent to which it is reasonable to assume neutrality in mtDNA replication/degradation at the microscopic level is yet to be carefully explored. Visualisation of heteroplasmy in live cells, combined with single-nucleoid tracking, would allow determination of the selective pressure placed by mitophagy/biogenesis and validation of the assumption of neutrality. Perturbations of mtDNA copy number through

ddC treatment, or heteroplasmy through e.g. zinc-finger nucleases [110], could also constrain the state dependence of mitophagy and biogenesis.

In Chapter 3, we generated a simple mathematical model (Table 3.1) to summarise a diverse set of omics data [37] describing the cell-physiological consequences of progressive increase in heteroplasmy of the 3243A>G MELAS mutation. Our model suggests that cells attempt to maintain wild-type mtDNA density through cell volume reduction until a minimum cell volume is reached; thereafter, cells upregulate energy production pathways (Section 3.4.5), allowing cell volume to recover. Our work highlights the sensitivity of cell volume to the mitochondrial genotype, and the importance of mtDNA density in cell physiology.

The dataset analysed in Chapter 3 consists of aggregates of cells, and therefore heteroplasmy corresponds to the mean heteroplasmy of a cellular population. It is unclear the extent to which intra-population heterogeneity may impact the observations of Picard *et al.* Future single-cell studies are required to map the genotype-phenotype link between cellular heteroplasmy and its physiological consequences. Cells harvested from heteroplasmic animals, such as the heteroplasmic tRNA-Ala mouse [335], as opposed to cybrid cells [37], would also remove potential confounding effects from nuclear instability. Future work is required to clarify the bioenergetic importance of mtDNA density homeostasis, especially in quiescent cells given the importance of mitochondrial genetic state in determining ageing phenotypes in non-dividing tissues [7].

In Chapter 4, we provide a number of competing explanations for the recently observed non-linear relationship between cellular volume and (a normalised measure of) mitochondrial membrane potential [50]. We suggest that a simple model of cellular energy demand, consisting of cell volume and surface area contributions, and energy supply, consisting of mitochondrial mass scaled by mitochondrial functionality, explains cellular energy scaling relationships (Box 4.2). Our model of energy scaling, when combined with the hypothesis that smaller cells are at a higher hazard of cell death, may account for many of the observations of Miettinen & Björklund [50], and provides an alternative hypothesis to the proposition that maximal mitochondrial functionality is achieved at an optimal cell size [50].

To distinguish the hypotheses made in Chapter 4, stringent cell death gating on cells, and the use of quantitative measures of mitochondrial membrane potential and mass (as opposed to JC-1 [50]), would potentially be insightful. Our metabolic scaling argument (Box 4.2) predicts a monotonic relationship between mitochondrial functionality and cell size, whereas the explanation put forward in Ref. [50] (analogous to Eq.(4.6)) predicts a concave functional relationship, when early-apoptotic cells are removed. A key assumption in the model we present is

that mitochondrial density is constant across all cell volumes, independently of cellular power demand. Bayesian inference applied to a large dataset of joint single-cell measurements of cell volume and mitochondrial mass may be able to constrain the power-law relationship between these two variables.

In Chapter 5, we present a set of initial estimates for the prevalence of somatic mutations in genes associated with neurodegeneration in the human brain by assuming that neurodevelopment may be modelled as a deterministic branching process (Section 5.4.1). We find that islands of pathologically mutated cells consisting of on the order of 10^5 cells are potentially common. Our observations were broadly robust to incorporation of simple models of cell death, founder effects, and asymmetric division (Section 5.4.2). A simple order-of-magnitude calculation (Section 5.4.2.4), assuming unbiased, spatially proximal, replication during neurodevelopment provided intuition for the observed robustness of the branching process model.

Future work is required to i) validate the existence of such pathologically mutated regions, as hypothesised here; ii) verify whether such regions are capable of causing neurodegenerative disease. The power calculation presented in Section 5.5 suggests that on the order of 10^5 samples are required to detectably measure a single pathologically mutated sample. The number of samples required may be reduced by increasing the number of pathological sites of interest, or through increased within-sample detection sensitivity. An interesting area of future development is to apply tree-like models to understand the prevalence of somatic mosaicism for pathological mutations of the mitochondrial genome [336], and in nuclear genes associated with mitochondrial dysfunction [337], across human tissues.

The countable nature of mitochondrial DNA provides the opportunity for stochastic modelling to provide insight into mitochondrial (patho)physiology. Statistical inference allows us to connect such models with data, and refine our understanding. The quantitative treatments explored here have shed new light on the dynamics of mitochondrial mutations, as well as the influence of mitochondria upon cellular (patho)physiology. Continued interdisciplinary work to understand mitochondrial genetics and physiology holds the promise of allowing us to understand the foundations of cellular bioenergetics, and design rational interventions for the diverse set of diseases associated with mitochondrial dysfunction, as well as healthy ageing.



SUPPLEMENTAL INFORMATION

COMPLEMENTING CHAPTER 2

A.1 CONSTANT RATES YIELD UNSTABLE COPY NUMBERS FOR A MODEL DESCRIBING MTDNA GENETIC AND NETWORK DYNAMICS

We explored a simpler network system than the one presented in Section 2.3.1, but found that it produced instability in mtDNA copy numbers, which we regard as biologically undesirable. Consider the following set of Poisson processes for singleton (s) and fused (f) species



where Eq.(A.1)-Eq.(A.3) are analogous to Eq.(2.1)-Eq.(2.3) where mutant species are neglected. The parameter ρ is shared amongst all of the birth and death reactions in Eq.(A.4)-Eq.(A.7). ρ represents the intuitive assumption that, in order for a stable population size to exist, birth should balance death. However, for the network to have any effect at all, singletons should be at an increased risk of mitophagy relative to fused species. We represent the increased risk of singleton mitophagy with the parameter η . Since additional death is introduced into the system when $\eta > 1$, we include the parameter $\alpha > 1$ as an increased global biogenesis rate to balance the increased mitophagy of singletons.

We may write the above system as a set of ordinary differential equations

$$\frac{ds}{dt} = -\gamma s^2 - \gamma f s + \beta f + \alpha \rho s - \eta \rho s \quad (\text{A.8})$$

$$\frac{df}{dt} = \gamma s^2 + \gamma f s - \beta f + \alpha \rho f - \rho f \quad (\text{A.9})$$

where we have enforced the stochastic reaction rate to be equivalent to the deterministic reaction rate, and hence the s^2 term is proportional to γ rather than 2γ (justification of this is presented in Section 2.3.3.1, see Eq.(2.22)). In Figure A.1 we see that the system displays a trivial steady state at $s = f = 0$ and a non-trivial steady state. Computing the eigenvalues of the Jacobian matrix at the non-trivial steady state indicates that it is a saddle node, and therefore unstable. Initial copy numbers which are too small tend towards extinction with time, and initial copy numbers which are too large tend towards a copy number explosion. This simple example suggests that a system of this form with constant reaction rates is unstable, and therefore biologically unlikely to exist under reasonable circumstances. We hence consider analogous biochemical reaction networks with a replication rate which is a function of state, to prevent extinction and divergence of the total population size.

A.2 HETEROPLASMY VARIANCE RELATIONS FOR ALTERNATIVE MODEL STRUCTURES AND MODES OF GENETIC MTDNA CONTROL

Here we explore the implications of alternative model structures upon Eq.(2.65). Firstly, we may consider replacing Eq.(2.4) with

$$X_S \xrightarrow{\lambda} X_S + X_S. \quad (\text{A.10})$$

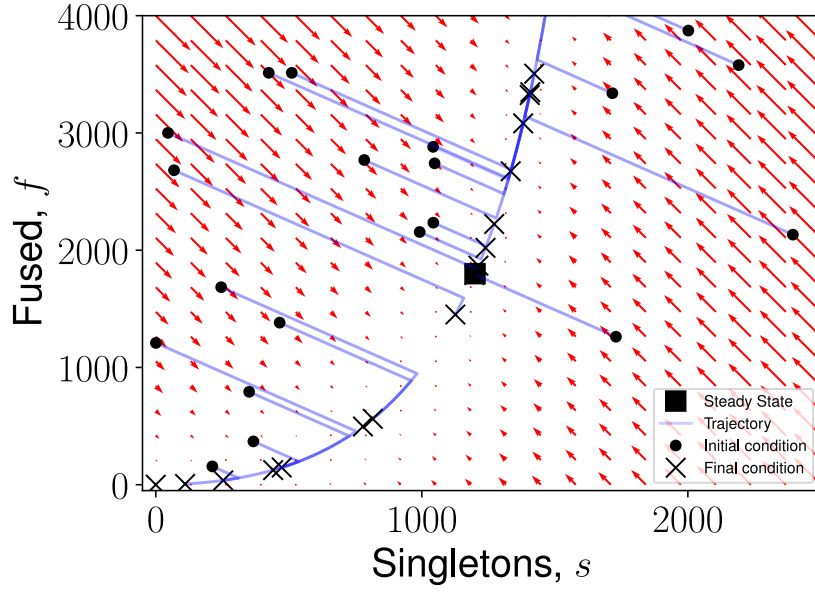


Figure A.1: Phase portrait for an ODE representation of a network system with constant rates. The system displays two steady states: a trivial steady state at $s + f = 0$, and a non-trivial steady state. At a steady state, both time derivatives in Eq.(A.8) and Eq.(A.9) vanish. Trajectories (blue) show the evolution of the system up to $t = 1000$. The direction and magnitude of the derivative at points in space are shown by red arrows. Trajectories can be seen either decaying to $s = f = 0$ or tending to infinity. $\gamma = 0.01656$, $\beta = 33.12$, $\rho = 0.023$, $\eta = 1.1$, $\alpha = 1.04$.

This corresponds to the case where replication coincides with fission, see [338]. Repeating the calculation in Section 2.3.3 also results in Eq.(2.65), so the result is robust to the particular choice of mtDNA replication reaction (see GitHub repository for Mathematica notebook).

Secondly, we may explore the impact of allowing non-zero mtDNA degradation of fused species. This could correspond to autophagy-independent degradation of mtDNA, for example via the exonuclease activity of POLG [339]. To encode this, we may add the following additional reaction



where $0 \leq \xi \leq 1$. We were not able to make analogous analytical progress in this instance. However, numerical investigation (Figure A.2) revealed that the following ansatz was able to predict heteroplasmy variance dynamics

$$\mathbb{V}(h) \approx \frac{2\mu t}{n(\mathbf{x})} h(\mathbf{x})(1 - h(\mathbf{x}))(f_s(\mathbf{x}) + \xi(1 - f_s(\mathbf{x}))) \Big|_{\mathbf{x}=\mathbf{x}_{\text{ss}}} . \quad (\text{A.12})$$

In other words, allowing degradation of fused species results in a linear correction to our heteroplasmy variance formula in Eq.(2.11). If fused species are susceptible to degradation at the same

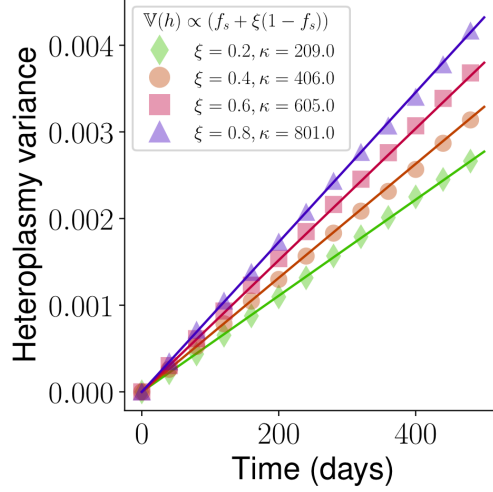


Figure A.2: Heteroplasmy variance for non-zero degradation of fused mitochondria. Allowing fused species to be degraded with relative rate ξ (see Eq.(A.11)), stochastic simulations for heteroplasmy variance (markers) and Eq.(A.12) (lines) are shown. As ξ is increased, κ in Eq.(2.10) is also increased to maintain steady-state copy number at approximately 1000 mtDNAs. Fused species degradation induces a linear correction to the heteroplasmy variance formula.

rate as unfused species ($\xi = 1$), then $\mathbb{V}(h)$ loses f_s dependence entirely and the mitochondrial network has no influence over heteroplasmy dynamics.

We also explored various different forms of $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$, which we label A-G after [43], and X-Z for several newly-considered functional forms, see Table A.1 and Figure A.3. Note that control D of [43] involves no feedback, which we do not explore – see Section A.1 and Figure A.1. The argument presented in the previous section requires the steady state solution of the system to be solvable, since we require the explicit form of $\mathbf{x}_{ss}$ in Eq.(2.59), Eq.(2.61) and Eq.(2.63). For controls B, C, E, F, G, Y and Z in Table A.1, the steady states are solvable and similar arguments to the above can be applied (see the GitHub repository for Mathematica notebooks). Controls B, C, E, F all satisfy Eq.(2.65); this can be shown numerically for controls A and X. However, controls G, Y and Z satisfy

$$dh = \sqrt{\frac{2\mu h(1-h)}{n(\mathbf{x})}} \Big|_{\mathbf{x}=\mathbf{x}_{ss}(h)} dW. \quad (\text{A.13})$$

Notably, Eq.(A.13) does not depend on f_s , unlike Eq.(2.65) (see GitHub repository for Mathematica notebooks). This is because control of copy number occurs in the degradation rate, rather than the replication rate, for controls G, Y and Z. A modified version of a Moran process (presented below) can provide intuition for why the diffusion rate of heteroplasmy variance depends on the network state when the population is controlled through replication, and does

not depend on network state when the population is controlled through degradation.

Interpretation	Replication rate	Degradation rate	Note
Linear feedback (see Figure 2.2, Figure 2.4, Figure A.3 A)	$\mu + b(\kappa - w_T - \delta m_T)$; see Table 2.1	μ ; see Table 2.1	Control E in (Johnston and Jones, 2016) and GitHub repository
Relaxed replication (see Figure A.3B)	$\alpha\mu(w_{opt} - w_T - \delta m_T)/(w_T + m_T)$; $\alpha = 1$, $w_{opt} = 1000$, $\delta = 1$	μ ; $\mu = 0.023$	Control A in (Johnston and Jones, 2016) and GitHub repository
Differential control for target population (see Figure A.3C)	$\alpha(w_{opt} - w_T)$	μ ; $\mu = 0.023$	Control B in (Johnston and Jones, 2016) and GitHub repository
Ratiometric control for target population (see Figure A.3D)	$\alpha(w_{opt}/w - 1)$; $\alpha = 1$, $w_{opt} = 1000$	μ ; $\mu = 0.023$	Control C in (Johnston and Jones, 2016) and GitHub repository
Production independent of wild-type (see Figure A.3E)	α/w_T ; $\alpha = 5$	μ ; $\mu = 0.023$	Control F in (Johnston and Jones, 2016) and GitHub repository
General linear feedback (see Figure A.3F)	$\mu + b(\kappa - \delta_1 w_s - \delta_2 w_f - \delta_3 m_s - \delta_4 m_f)$; see Table 2.1, $\delta_1 = 0.8$, $\delta_2 = 1.0$, $\delta_3 = 0.2$, $\delta_4 = 0.3$	μ ; see Table 2.1	Control X in GitHub repository
Ratiometric control through degradation (see Figure A.3G)	λ ; $\lambda = 0.023$	$\mu w_T/w_{opt}$; $\mu = 0.023$; $w_{opt} = 200$	Control G in (Johnston and Jones, 2016) and GitHub repository
Linear feedback in degradation (see Figure A.3H)	λ ; $\lambda = 0.023$	$\mu + b(w_T + \delta m_T - \kappa)$; $\mu = 0.023$, $b = 10^{-4}$, $\kappa = 1000$, $\delta = 1$	Control Y in GitHub repository
Differential control for target population in degradation (see Figure A.3I)	λ ; $\lambda = 0.023$	$\alpha(w_T - w_{opt})$; $\alpha = 1$, $w_{opt} = 1000$	Control Z in GitHub repository

Table A.1: Nominal parameterisations for the alternative feedback control functions explored. In all cases, the nominal fission and fusion rates were $\beta = 33.12$, $\gamma = 0.038$ respectively.

A.3 CHOICE OF NOMINAL PARAMETRISATION

In this section we discuss our choice of nominal parametrisation for the network system in Eq.(2.1)-Eq.(2.9), given the replication rate in Eq.(2.10). We will first discuss our choice of network parameters.

Ref. [216] found that the average fission rate in cortical neurons is 0.023 ± 0.003 fissions/mitochondria/min. Assuming that this value is representative of the fission rate in general, and converting this to

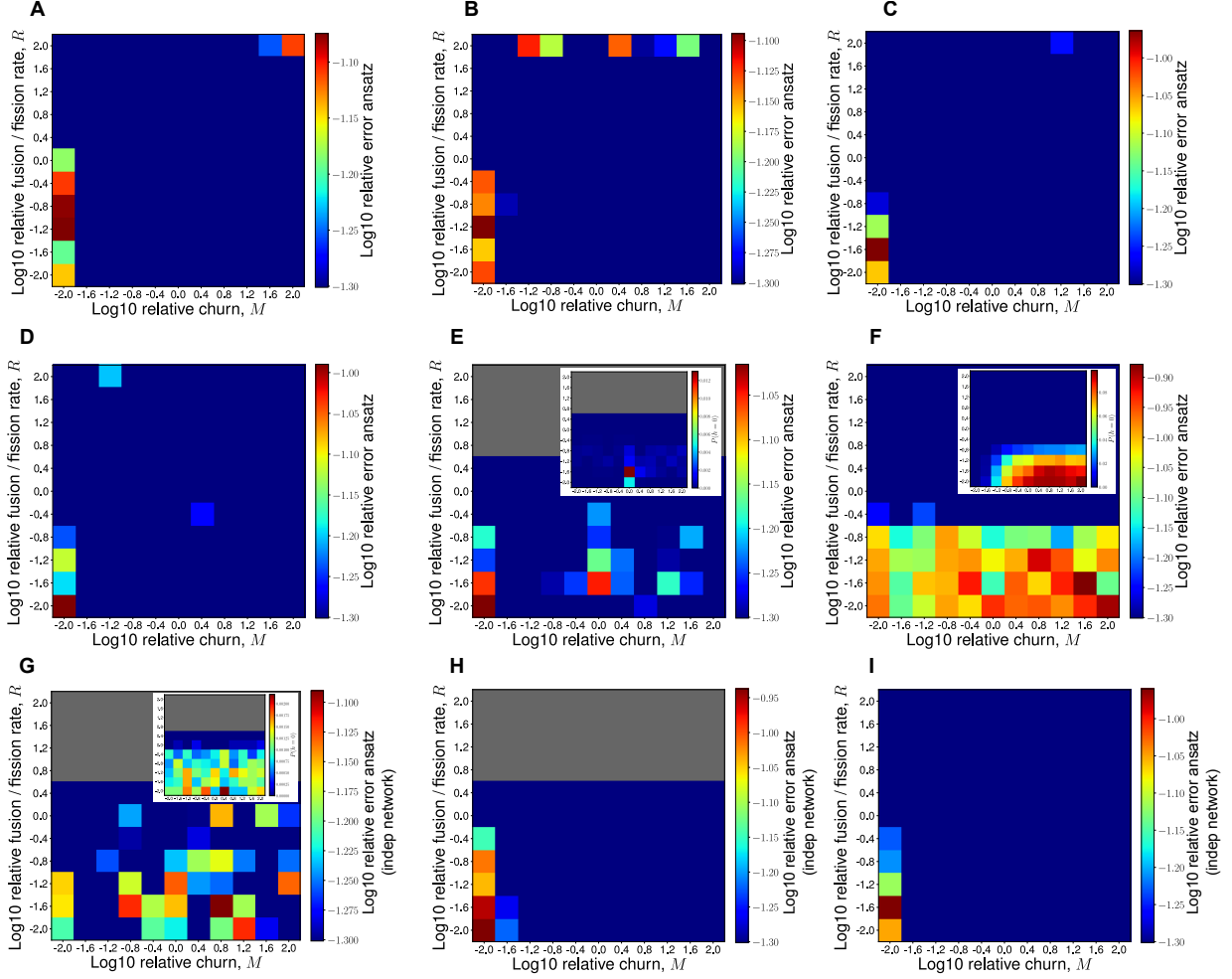


Figure A.3: Parameter sweeps of network fission-fusion rates for replication-based and degradation-based control modes show robustness of two heteroplasmy variance formulae respectively. Error in Eq.(2.11) (ansatz) and Eq.(A.15) (ansatz indep network) for sweeps of the fusion and fission rates for the corresponding feedback control functions in Table A.1. Equations are accurate to at least 5% (blue regions) across large regions of parameter space, for many control laws. Fusion and fission rates are redefined as $\gamma \rightarrow \gamma_0 MR$ and $\beta \rightarrow \beta_0 M$ where M and R denote the magnitude and ratio of the network rates, and γ_0, β_0 denote the nominal parametrisations of the fusion and fission rates respectively. Summary statistics for 10^4 realisations, with initial condition $h = 0.3$ and evaluated at $t = 500$ days. Errors in $\mathbb{V}(h)$ (see Eq.(A.28)) smaller than 5% are truncated and are shown as blue. Parametrisations where a deterministic steady state could not be found for an initial condition of $h = 0.3$ are shown in grey. Inset figures, where present, display the probability of fixation at $h = 0$. Where insets are not present, the probability of fixation is negligible. **(A-F)** When $\lambda = \lambda(x)$ and $P(h = 0)$ is low, Eq.(2.11) performs well in the high-churn limit. **(G-I)** When $\mu = \mu(x)$ and $P(h = 0)$ is low, Eq.(A.15) performs well in the high-churn limit.

units of per day, we may write the mitochondrial fission rate as $\beta = 33.12 \text{ day}^{-1}$.

Note that the dimensions of β are day^{-1} and not $\text{mitochondrion}^{-1} \text{ day}^{-1}$. This is because if the propensity (see Eq.(2.14), where $\Omega = 1$) of e.g. Eq.(2.3) is $\hat{a}_{\text{fis},w} = \beta w_f$ then the mean time to the next event is $1/(\beta w_f)$; therefore the dimension of β is per unit time and copy numbers are pure numbers, i.e. dimensionless. Similar reasoning constrains the dimension of the fusion rate, see below.

Evaluation of the fusion rate is more involved, since fusion involves two different chemical species coming together to react whereas fission may be considered as spontaneous. Furthermore, there are 7 different fusion reactions whereas there are only 2 fission reactions. For simplicity, assume that all species have a steady-state copy number of $x_i = 250$ (resulting in a total copy number of 1000, heteroplasmy of 0.5 and 50% of mitochondria existing in the fused state). Neglecting subtleties relating to bimolecular reactions involving one species (see Eq.(2.22)), each fusion reaction proceeds at rate $\hat{a}_{\text{fus},j} \approx \gamma x_i^2$. Since there are 7 fusion reactions (Eq.(2.1), Eq.(2.2), Eq.(2.7)-Eq.(2.9)), the total fusion propensity is $\hat{a}_{\text{fus}} \approx 7\gamma x_i^2$. Similarly, the total fission propensity is $\hat{a}_{\text{fis}} = \beta(w_f + m_f) = 2\beta x_i$. Since we expect macroscopic proportions of both fused and fissioned species in many physiological settings, we may equate the fusion and fission propensities, $\hat{a}_{\text{fus}} = \hat{a}_{\text{fis}}$, and rearrange for the fusion rate γ to yield $\gamma = 2\beta x_i / (7x_i^2) \approx 3.8 \times 10^{-2} \text{ day}^{-1}$. Note that the orders of magnitude difference between β and γ stems from the observation that fusion propensity depends on the square of copy number whereas the fission propensity depends on copy number linearly.

Given the network parameters, we then explored appropriate parametrisations for the genetic parameters: the mitophagy rate (μ) and the parameters of the linear feedback control (κ , b and δ , see Eq.(2.10)). mtDNA half-life is observed to be highly variable: in mice this can be between 10-100 days [65]. For consistency with another recent study investigating the relationship between network dynamics and heteroplasmy, we use an mtDNA half-life of 30 days [207].

The parameter δ in the replication feedback control (see Eq.(2.10)) may be interpreted as the “strength of sensing of mutant mtDNA” in the feedback control [68]. Assuming that fluctuations in copy number of mutants and wild-type molecules are sensed identically (as may be the case for e.g. non-coding mtDNA mutations) we may reasonably assume a model of $\delta = 1$ as the simplest case of a neutral mutation.

We are finally left with setting the parameters κ and b in the linear feedback control Eq.(2.10). In the absence of a network state and mutants, κ is precisely equal to the steady state copy number, since the degradation rate equals the replication rate when $\kappa = w$. However, the

presence of a network means that a subpopulation of mtDNAs (namely the fused species) are immune to death, resulting in κ no longer being equivalent to the steady state copy number. The parameter b may be interpreted as the feedback control strength, which determines the extent to which the replication rate changes given a unit change in copy number.

Given a particular value of b , we may search for a κ which gives a total steady state copy number (n) which is closest to some target value (e.g. 1000 as a typical total mtDNA copy number per cell in human fibroblasts [180]). We swept a range of different values of b and found that, for values of b smaller than a critical value (b^*), a κ could not be found whose deterministic steady state was sufficiently close to $n = 1000$. This result is intuitive because in the limit of $b \rightarrow 0$, $\lambda = \text{const.}$ From the analysis in Section A.1, we have shown that constant genetic rates (μ, λ) result in unstable copy numbers, and therefore a sufficiently small value of b is not expected to yield a stable non-trivial steady state solution. We chose $b \approx b^*$, and the corresponding κ , such that the steady state copy number is controlled as weakly as possible given the model structure.

A.4 RATE RENORMALISATION

In Eq.(2.1)–Eq.(2.10) we have neglected reactions such as



because they do not change the number of molecules in our state vector $\mathbf{x} = (w_s, w_f, m_s, m_f)$. One may ask whether neglecting such reactions means that it is necessary to renormalise the fission-fusion rates which were estimated in Section A.3. In estimating the nominal parametrisation, we began by using a literature value for the mitochondrial fission rate, and then matched the fusion rate such that the total propensity of fusion approximately balanced the total propensity of fission. This matching procedure is reasonable, since we observe a mixture of fused and fissioned mitochondria under physiological conditions: choosing a fusion rate which is vastly different results in either a hyperfused or fragmented network. We must therefore only justify the fission rate. Eq.(2.3) assumes that a fission reaction always results in a singleton, and a singleton is by definition a molecule which is susceptible to mitophagy (see Eq.(2.6)). Therefore, if fission reactions always result in mitochondria containing single mtDNAs which are susceptible to mitophagy, then we expect our model to match well to true physiological rates. If, on the other hand, fission reactions between large components of the network which are too large to

be degraded are common, then renormalisation of β by the fraction of fission events which result in a sufficiently small mitochondrion would be necessary, which would in turn renormalise γ through our matching procedure. We are not aware of experimental measurements of the fraction of fission events which result in mitochondria which contain a particular number of mtDNAs. Such an experiment, combined with the distribution of mitochondrial sizes which are susceptible to mitophagy, would allow us to validate our approach. Despite this, the robustness of our results over approximately 4 orders of magnitude for the fission-fusion rate (Figure A.3) provides some indication that our results are likely to hold in physiological regions of parameter space.

A.5 A MODIFIED MORAN PROCESS MAY ACCOUNT FOR THE ALTERNATIVE FORMS OF HETEROPLASMY VARIANCE DYNAMICS UNDER DIFFERENT MODELS OF GENETIC mtDNA CONTROL

We sought to gain insight into why control of population size through the replication rate ($\lambda = \lambda(\mathbf{x})$, $\mu = \text{const}$) results in heteroplasmy variance depending on the fraction of unfused mitochondria (see Eq.(2.11)), whereas control of population size through the degradation rate ($\mu = \mu(\mathbf{x})$, $\lambda = \text{const}$) results in heteroplasmy variance becoming independent of network state, where

$$\mathbb{V}(h) \approx \frac{2\lambda t}{n(\mathbf{x})} h(\mathbf{x})(1 - h(\mathbf{x})) \Big|_{\mathbf{x}=\mathbf{x}_{\text{ss}}} . \quad (\text{A.15})$$

We will proceed by considering an analogous Moran process to the set of reactions presented in Eq.(2.1)–Eq.(2.9).

First, consider a haploid biallelic Moran process consisting of wild-types and mutants, in a population of fixed size n . At each step in discrete time, a member of the population is chosen for birth, and another for death. Let m_t denote the copies of mutants at time t . Then,

$$P(m_{t+1} = j | m_t = i) = \begin{cases} i(n-i)/n^2 & \text{if } j = i \pm 1 \\ i^2/n^2 + (n-i)^2/n^2 & \text{if } j = i \\ 0 & \text{otherwise.} \end{cases} \quad (\text{A.16})$$

It follows that

$$\mathbb{E}(m_{t+1} | m_t) = m_t. \quad (\text{A.17})$$

Defining $h_t := m_t/n$ then from Eq.(A.16)

$$\mathbb{V}(m_{t+1}|m_t) = 2h_t(1 - h_t) \quad (\text{A.18})$$

and therefore

$$\mathbb{V}\left(\frac{m_{t+1}}{n} \middle| m_t\right) = \mathbb{V}(h_{t+1}|m_t) = \frac{2}{n^2}h_t(1 - h_t). \quad (\text{A.19})$$

Suppose that, instead of the process occurring in discrete time, the process occurs with continuous time, where each event is a simultaneous birth and death, and is modelled as a Poisson process. Suppose that events occur at a rate μ per individual. The waiting time between successive events (τ) is an exponential random variable with rate μn . Hence the expected waiting time between successive events is

$$\mathbb{E}(\tau) = \frac{1}{\mu n}. \quad (\text{A.20})$$

If we take the ratio of Eq.(A.19) and Eq.(A.20), we have

$$\frac{\mathbb{V}(h_{t+\tau}|m_t)}{\mathbb{E}(\tau)} = \frac{2\mu h_t(1 - h_t)}{n}. \quad (\text{A.21})$$

Heuristically, one could interpret Eq.(A.21) as a ratio of differentials as follows. If we were to suppose that n were large enough such that $\mathbb{E}(\tau)$ is very small, and h_t is approximately constant (h_0) after a small number of events, then

$$\frac{\Delta \mathbb{V}(h)}{\Delta \tau} \approx \frac{\mathbb{V}(h_{t+\tau}|m_t)}{\mathbb{E}(\tau)} \implies \frac{d\mathbb{V}(h)}{dt} \approx \frac{2\mu h_0(1 - h_0)}{n} \implies \mathbb{V}(h, t) \approx \frac{2\mu h_0(1 - h_0)}{n}t \quad (\text{A.22})$$

where we have replaced the inter-event time τ with physical time t . This result is analogous to Eq.(A.15) and Equation (12) of [43], and agrees with simulation (Figure A.4A).

Now consider the modified Moran process in Figure A.4B, which we refer to as a “protected” Moran process. Let $0 < f_s \leq 1$ be the fraction of individuals susceptible to death, which is a constant. $nf_s h_t$ mutants and $nf_s(1 - h_t)$ wild-types are randomly chosen to be susceptible to death, where n is large. In this continuous-time model, the inter-event time is $\tau \sim \text{Exponential}(\Gamma)$ where Γ will be defined below. Then an individual from the susceptible population is chosen for death, and any individual is allowed to be born. The birth and death events occur simultaneously in time.

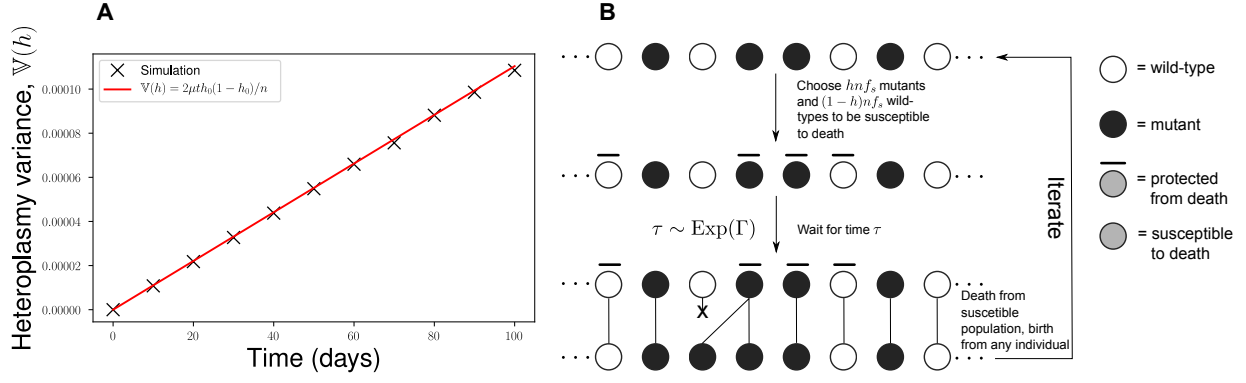


Figure A.4: Exploration of analogous Moran processes. (A) The original biallelic Moran process satisfies Eq.(A.15), where h_0 is the initial heteroplasmy, which is equivalent to $\mathbb{E}(h)$. (B) The “protected” Moran process. The population size is constrained to be fixed to some large constant, n . There exist two alleles, mutants (black circle) and wild-types (white circles). hnf_s mutant and $(1-h)nf_s$ wild-type molecules are susceptible to death, the rest are protected from death (denoted by a bar). An exponential random variable is drawn as the waiting time to the next event (see Section A.5 for discussion of the form of the rate Γ). Time is incremented by the waiting time, then a death event occurs from the susceptible population and a birth event from any individual simultaneously. The same individual is allowed to be chosen for both birth and death. The process is then repeated iteratively.

Again, using t as an integer counter of events, we have

$$P(m_{t+1} = j | m_t = i) = \begin{cases} \frac{nf_s h_t}{nf_s} \frac{n(1-h_t)}{N} & \text{if } j = i - 1 \\ \frac{nf_s(1-h_t)}{nf_s} \frac{nh_t}{N} & \text{if } j = i + 1 \\ \frac{nf_s h_t}{nf_s} \frac{nh_t}{N} + \frac{nf_s(1-h_t)}{nf_s} \frac{n(1-h_t)}{n} & \text{if } j = i \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.23})$$

$$= \begin{cases} h_t(1-h_t) & \text{if } j = i \pm 1 \\ h_t^2 + (1-h_t)^2 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.24})$$

which is equivalent to the definition of a Moran process in Eq.(A.16), meaning that Eq.(A.19) applies to the protected Moran process as well.

We consider two heuristic arguments for choosing the inter-event rate Γ , where the inter-event time $\tau \sim \text{Exponential}(\Gamma)$. Firstly, if the death rate per capita is constant (μ), then the rate at which a death event occurs in the system (Γ_{death}) is proportional to the number of individuals which are susceptible to death: $\Gamma_{\text{death}} = \mu n f_s$. If we assume that the overall birth rate is matched to the overall death rate so that population size is maintained, as is the case when $\lambda = \lambda(\mathbf{x})$ in the network system, then the overall birth rate (Γ_{birth}) must also be $\Gamma_{\text{birth}} = \mu n f_s$.

Hence,

$$\Gamma = \Gamma_{\text{birth}} + \Gamma_{\text{death}} = 2\mu n f_s \quad (\text{A.25})$$

where μ is a proportionality constant. Since, for a Moran event to occur, both a birth and a death event must occur, time effectively runs twice as fast in a Moran model relative to a comparable chemical reaction network model. We therefore rescale time by taking $\mu \rightarrow \mu/2$, and thus

$$\Gamma = \mu n f_s. \quad (\text{A.26})$$

As a result, $\mathbb{E}(\tau) = 1/(\mu n f_s)$ and therefore, using Eq.(A.19) and the reasoning in Eq.(A.22),

$$\frac{\Delta \mathbb{V}(h)}{\Delta \tau} \approx \frac{\mathbb{V}(h_{t+\tau}|m_t)}{\mathbb{E}(\tau)} \implies \frac{d\mathbb{V}(h)}{dt} \approx \frac{2\mu f_s h_0(1-h_0)}{n} \implies \mathbb{V}(h) \approx \frac{2\mu f_s h_0(1-h_0)}{n} t. \quad (\text{A.27})$$

This is analogous to when $\lambda = \lambda(\mathbf{x})$ and $\mu = \text{const}$ in the network system. *Hence, when $\lambda = \lambda(\mathbf{x})$ and $\mu = \text{const}$, the absence of death in the fused subpopulation means the timescale of the system (being the time to the next death event) is proportional to f_s .* We note that this argument is only a heuristic, since the Moran process is defined such that birth and death events occur simultaneously and therefore do not possess separate propensities (Γ_{birth} and Γ_{death}).

The second case we consider is when each individual has a constant rate of birth, hence $\Gamma_{\text{birth}} \propto n$. Then the death rate is chosen such that $\Gamma_{\text{birth}} = \Gamma_{\text{death}}$. In this case $\Gamma = \lambda n$, where λ is a proportionality constant. The same argument from Eq.(A.20) to Eq.(A.22) may be applied, with an appropriate rescaling of time, and we arrive at Eq.(A.15). This is analogous to when $\mu = \mu(\mathbf{x})$ and $\lambda = \text{const}$ in the network system. *Hence, when $\mu = \mu(\mathbf{x})$ and $\lambda = \text{const}$, the presence of a constant birth rate in the entire population means the timescale of the system (being the time to the next birth event) is independent of f_s .*

A.6 ERROR QUANTIFICATION

In Figure 2.5 and Figure A.3, we compare Eq.(2.11) and Eq.(A.15) to stochastic simulations, for various parametrisations and replication/degradation rates. In order to quantify the accuracy of these equations in predicting $\mathbb{V}(h, t)$, we define the following error metric ϵ

$$\epsilon = \left| 1 - \frac{\dot{\mathbb{V}}(h, t)_{\text{Th}}}{\mathbb{E}_t(\dot{\mathbb{V}}(h, t)_{\text{Sim}})} \right| \quad (\text{A.28})$$

where $\dot{\mathbb{V}}(h, t)$ is the time derivative of heteroplasmy variance with subscripts denoting theory (Th) and simulation (Sim). An expectation over time ($\mathbb{E}_t$) is taken for the stochastic simulations, whereas $\dot{\mathbb{V}}(h, t)$ is a scalar quantity for Eq.(2.11) and Eq.(A.15).

Simulations were performed using the Imperial College High Performance Computing Service.

B

SUPPLEMENTAL INFORMATION COMPLEMENTING CHAPTER 3

B.1 TRANSFORMATION TO PER-CELL DIMENSIONS USING CELL VOLUME

In this section we show that it is necessary to multiply measurements of protein and mRNA levels, when determined by Western blot and RNA-seq respectively, by cell volume to transform the data to mean cellular measurements.

Consider a Western blot experiment determining the levels of a gene (gene i) in two conditions (A and B). Denote the number of proteins per cell of gene i , as n_i^A , where the superscript denotes condition A. Let us also denote the total number of proteins per cell as N^A .

When one performs a Western blot, the protein of interest is stained with an antibody, cells are lysed, and a sample of fixed protein mass (m) is taken from the lysate. If we denote the number of proteins for the gene of interest in the sample as P_i^A , then we may write

$$P_i^A = \frac{n_i^A}{N^A}m, \tag{B.1}$$

since the proportion of protein i in the sample is determined by the proportion of protein i in the proteome (n_i^A/N^A). Western blot experiments also tend to be normalised by a loading

control (c), so the normalised measurement we have access to is

$$\frac{P_i^A}{P_c^A} = \frac{n_i^A}{n_c^A} \quad (\text{B.2})$$

which corresponds to the data given in Picard et al.[37].

Now, consider a perturbation in condition B, causing the amount of protein for gene i to be n_i^B , and the mean cell volume to experience a fold-change V_f , as in Fig. B.1. Consequently, $N^B = V_f N^A$, since total protein content scales with the volume of the cell. Using the reasonable assumption that the loading control is a gene whose expression also scales with cell volume (e.g. β -actin, as in Picard et al.[37]), then $n_c^B = V_f n_c^A$. It follows that

$$\frac{P_i^B}{P_c^B} = \frac{n_i^B}{V_f n_c^A}. \quad (\text{B.3})$$

Then, if we are interested in the relative fold-change expression of the protein between the two conditions, then we take the ratio

$$\frac{P_i^B/P_c^B}{P_i^A/P_c^A} = \frac{n_i^B/n_c^A}{n_i^A/n_c^A} \cdot \frac{1}{V_f} = \frac{n_i^B}{n_i^A} \cdot \frac{1}{V_f} \quad (\text{B.4})$$

Thus, the quantity on the left hand side of Eq. (B.4), which is what one usually measures in a Western blot, has a multiplicative-bias of $1/V_f$. Therefore, if one is interested in per-cell protein changes, the appropriate quantity of interest is

$$\frac{P_i^B/P_c^B}{P_i^A/P_c^A} \cdot V_f. \quad (\text{B.5})$$

Hence, we multiply each protein measurement by $V(h) \equiv V_f(h) = \langle V_k(h) \rangle_k / \langle V_k(h=0) \rangle_k$, where $\langle \dots \rangle_k$ denotes the sample mean over technical replicates k , such that $V(0) = 1$.

A similar argument applies to RNA-seq data, since a fixed mass of mRNA is extracted for an RNA-seq experiment, so an analogous pair of equations to Eq.(B.1) in conditions A and B holds. Using the assumption that $N^B = V_f N^A$, and denoting the number of mRNA molecules in each sample with M , it can be shown that

$$\frac{M_i^B}{M_i^A} = \frac{n_i^B}{n_i^A} \cdot \frac{1}{V_f}. \quad (\text{B.6})$$

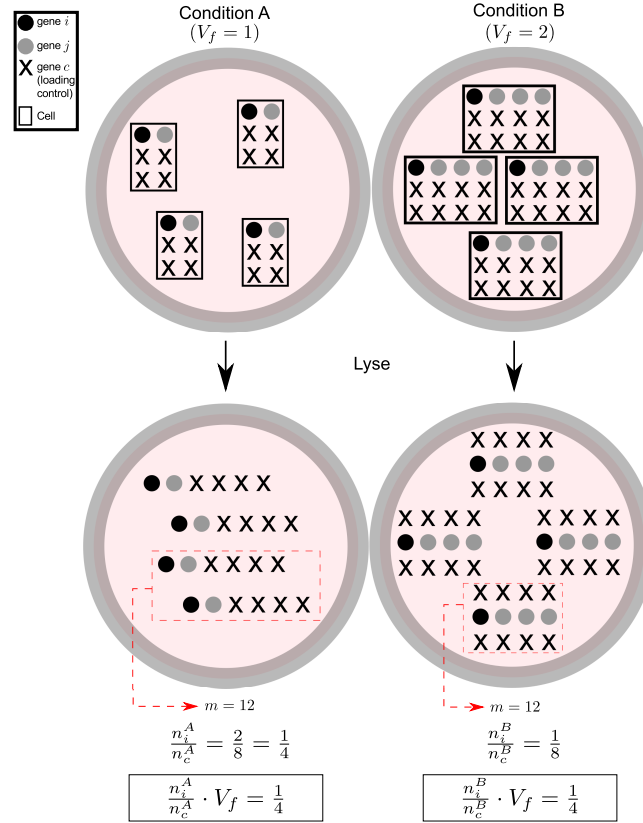


Figure B.1: Transformation of Western blot (or RNA-seq) data to per-cell dimensions. Consider a Western blot experiment, where we are interested in the fold-change expression of gene i per cell (n_i), and cell volume has a fold change $V_f = 2$ between conditions A and B. In this example, $n_i^A = n_i^B$. Taking an unbiased sample of size $m = 12$ from each condition, and dividing by the loading control, yields a quantity $1/V_f$ too small. It is necessary to multiply the ratio by V_f , to get an accurate measurement of gene i , in the context of strongly varying cell volume, as is the case in Picard *et al.* [37]. A similar argument holds for RNA-seq data, which also uses a fixed mass of RNA as the starting sample.

B.2 GENERATIVE MODEL DESCRIPTION

We used a Bayesian framework to find the supported parameter values given the data, using the Metropolis-Hastings algorithm [340]. To do this, we included an additional 6 noise parameters, for the features where parameter inference was performed (i.e. all of the features except N^+ , which has no free parameters, see Eq.(2.4)). For these 6 features ($M_{\text{ETC}}, P^+, M_{\text{gly}}, V, G, R_{\text{max}}$), we assumed that the data were generated subject to Gaussian noise.

Thus, the full statistical model contains 12 parameters (excluding 6 noise parameters for each feature), with 32 data points which enter the likelihood (after excluding $h = 1$ data). To summarise, counting the 6 features which have free parameters, the model consists of $12/6 = 2$ mean parameters per feature, on average. Note that simply fitting linear models to the 6 features in Fig. 3.1 would also require 2 parameters per feature. The model fit is shown in Fig. 3.5.

To connect our model of mean cellular behaviour $\mathcal{S} = \{M_{\text{ETC}}, P^+, M_{\text{gly}}, V, G, R_{\text{max}}\}$, to the data of Picard *et al.* [37], we assume that the sample mean of feature i ($y_{i,j}$) at a discrete value of heteroplasmy $h = j$ is generated via Gaussian noise ($\mathcal{N}(\mu, \sigma)$) whose mean corresponds to one of the models $\mathcal{S}$,

$$y_{i,j} = \mathcal{M}_i(h) + \mathcal{N}(0, \sigma_i), \quad (\text{B.7})$$

where $\mathcal{M}_i(h)$ is an element from the set of models $\mathcal{S}$. We stress that the data we train our model on, $y_{i,j}$, is the *sample mean*, rather than the raw data. This is a less common approach; however, we believe that it is appropriate as individual replicates only give us information on the technical variability measured in [37], whereas the total error is a combination of both technical and biological variability. Training our models on individual replicates would be likely to underestimate the true variability of the data, so we favoured training on the sample mean only. This raises the challenge of establishing an appropriately permissive model for our uncertainty in σ_i .

We can infer the distribution of the parameters (θ) of the models $\mathcal{S}$, given the data $y_{i,j}$, using Bayes rule and a prior distribution over θ ($P(\theta)$)

$$P(\theta|y_{i,j}) = \frac{P(y_{i,j}|\theta)P(\theta)}{P(y_{i,j})}. \quad (\text{B.8})$$

The log-likelihood in this case is

$$\log[P(y_{i,j}|\theta)] = \sum_{i,j} -1/2 \log(2\pi) - \log(\sigma_i) - \frac{1}{2} \left(\frac{y_{i,j} - \mathcal{M}_i(j)}{\sigma_i} \right)^2. \quad (\text{B.9})$$

We drop the constant $\sum_{i,j} -1/2 \log(2\pi)$ from our log-likelihood, since we will only be interested in differences in the log-likelihood to perform Bayesian inference using the Metropolis-Hastings algorithm [340].

We used exponential priors σ_i

$$P(\sigma_i) \propto \exp(-\lambda_i \sigma_i) \quad (\text{B.10})$$

as our error model. The constant λ_i was chosen such that the scale of decay of probability was on the same scale as the range of the data. Noting that $\langle P(\sigma_i) \rangle = 1/\lambda_j$, we chose

$$\lambda_i = \frac{\Omega}{\max_j \{\hat{y}_{i,j}\} - \min_j \{\hat{y}_{i,j}\}} \quad (\text{B.11})$$

where Ω is a hyper-parameter of the prior and $\Omega \geq 0$. Note that we may interpret $\Omega = 0$ as an improper uniform prior, since $P(\sigma_i) = \text{const}$ in this case.

We began with $\Omega = 0$ as the most permissive choice of prior possible, given the model in Eq.(B.10). We found that when $\Omega = 0$ *the maximum a posteriori estimates were qualitatively similar* to choosing $\Omega = 2$ (our final choice which we justify below) see Fig. 3.5 ($\Omega = 2$) and Fig. B.2A-F ($\Omega = 0$). However, we found that the posterior 25-75% credible interval supported model fits for M_{ETC} , P^+ and R_{max} which were relatively poor when $\Omega = 0$, compared to $\Omega = 2$ (see Fig. B.2A-F). We determined that large values of h^* were indicative of purely linear fits to the data, which is unlikely given the wider body of evidence demonstrating the nonlinearity of the threshold effect. This is seen in Fig. B.2G-L (high h^* , poorer fit) when compared with Fig. B.2M-R (low h^* , better fit). Comparison between Fig. B.2G and Fig. B.2M is particularly noteworthy, where the 25-75% posterior credible interval for high h^* sub-samples predicts $M_{\text{ETC}} \approx 0$ for all values of h , which is physiologically implausible, whereas low h^* sub-samples display non-linear fits which more faithfully track the data. Fig. B.3 shows that the high h^* mode is of comparable prevalence to the low h^* mode when $\Omega = 0$.

We therefore investigated the sensitivity to choice in Ω in Fig. B.3. We see that increasing Ω reduces the width of the marginal posterior distribution of h^* , constraining the posterior distribution to lie around the nonlinear solutions shown in Fig. 3.5. We found that the permissive prior $\Omega = 2$ was sufficient to strongly subdue this, physiologically implausible, large h^* mode.

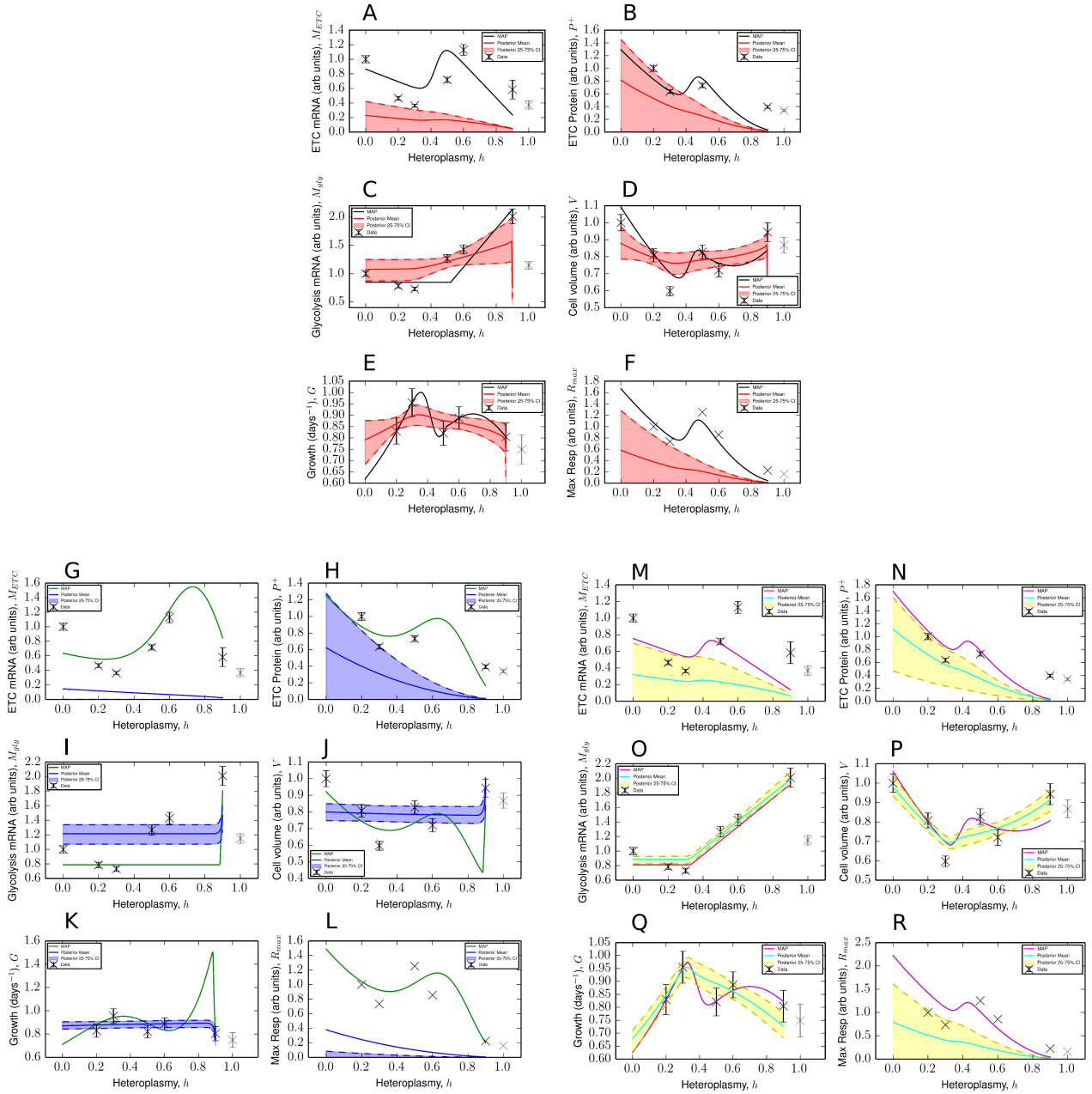


Figure B.2: Hyperparameter choice for error model (A-F) Model support for when $\Omega = 0$, corresponding to an improper uniform prior on the variance of each feature. (G-L) Sub-samples of posterior where $0.85 \leq h^* \leq 0.9$. (M-R) Sub-samples of posterior where $0.3 \leq h^* \leq 0.35$. We see a physiologically implausible fit for $0.85 \leq h^* \leq 0.9$, whereas when $0.3 \leq h^* \leq 0.35$ model fits were better able to describe the data (for instance, by comparing (G) to (M)).

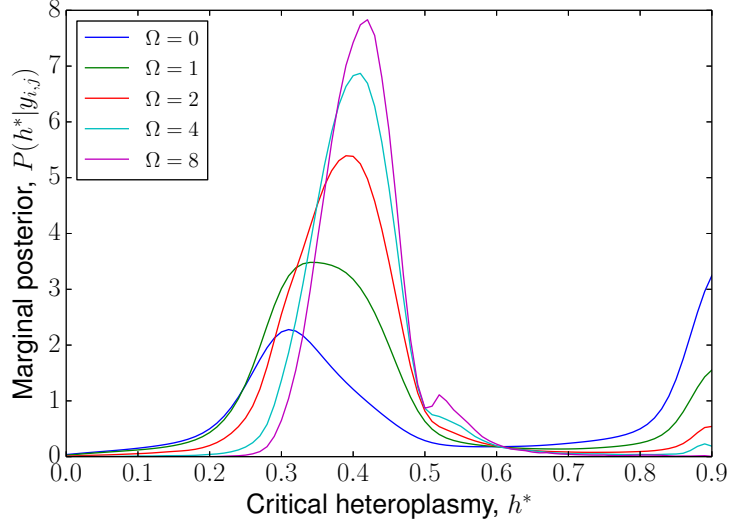


Figure B.3: Choice of uncertainty prior affects distribution of critical heteroplasmy and model fit. Larger values of Ω suppresses large values of model uncertainty (σ_i , see Eq.(B.10)), and consequently forces the model fit to more closely match the data. This corresponds to the h^* mode approximately between 0.3-0.4 see Fig. B.2.

This can be interpreted as a prior belief that our model uncertainty is, on average, 50% of the range of the data (since $\langle P(\sigma_i) \rangle = \lambda_i^{-1}$). We believe this to be a sensible prior choice, encoding our prior belief that the threshold effect is nonlinear while providing only a gentle constraint on parameters.

We favoured uniform priors on the remaining parameters so that the posterior would be dominated by the likelihood. However, a number of the parameters in the model were uncertain over orders of magnitude; in these cases, we allowed the log of these parameters to take uniform

distributions. Explicitly, our priors were chosen as:

$$P(h^*) = \text{unif}(0, 1) \quad (\text{B.12})$$

$$P(\ln(k_{\text{mRNA}})) = \text{unif}(-10 \ln 10, 2 \ln 10) \quad (\text{B.13})$$

$$P(f_m) = \text{unif}(0, 1) \quad (\text{B.14})$$

$$P(\ln(k_m)) = \text{unif}(-10 \ln 10, 2 \ln 10) \quad (\text{B.15})$$

$$P(\ln(\delta_p)) = \text{unif}(-10 \ln 10, 2 \ln 10) \quad (\text{B.16})$$

$$P(\ln(k_o)) = \text{unif}(-10 \ln 10, 2 \ln 10) \quad (\text{B.17})$$

$$P(\ln(k_g)) = \text{unif}(-10 \ln 10, 2 \ln 10) \quad (\text{B.18})$$

$$P(c_1) = \text{unif}(-10, 10) \quad (\text{B.19})$$

$$P(m_2) = \text{unif}(0, 100) \quad (\text{B.20})$$

$$P(k_{gr}) = \text{unif}(0, 100) \quad (\text{B.21})$$

$$P(\ln(k_p)) = \text{unif}(-10 \ln 10, 2 \ln 10). \quad (\text{B.22})$$

The ranges for h^* and f_m are justified since these quantities can physically only be between 0 and 1. c_1 and m_2 are parameters of linear models for M_{gly} (see Table 3.2 and Eq.(2.8)) for data which has been normalised to the scale of 1; therefore priors were chosen with suitably large ranges. Similarly for k_{gr} , a proportionality constant relating growth to cell volume ((see Table 3.2 and Eq.(2.10)), we expect k_{gr} to be of the order of 1, since the data has been normalised, and chose suitably relaxed priors. The ranges for all other parameters, which were sampled in log-space due to our greater uncertainty of their values, were chosen to be suitably large as to be unlikely to reach the boundary of the prior during sampling with MCMC.

The parameters β and k_{mRNA} from Eq.(2.5) were highly correlated. For more efficient chain mixing, we rearranged Eq.(2.5) into the form

$$M_{\text{ETC}} = \frac{\zeta}{\frac{1}{k_{\text{mRNA}}} + \frac{1}{1 + \exp[k_m(h - h_0)]}}, \quad (\text{B.23})$$

where $\zeta = \beta/k_{\text{mRNA}}$, and used the prior

$$P(\ln(\zeta)) \sim \text{unif}(-10 \ln 10, 2 \ln 10) \quad (\text{B.24})$$

such that the boundaries for the uniform prior were relatively relaxed.

We performed the Metropolis-Hastings algorithm [340] to sample from the posterior, using a

Gaussian random walk as our transition kernel, whose covariance matrix was determined from a trial run of the adaptive Metropolis algorithm [341]. All code was written in either `Python` or `C`, and is available upon request. The MCMC chain trajectory is presented in Figure B.4.

B.3 JUSTIFICATIONS FOR MODEL STRUCTURE

B.3.1 JUSTIFICATION OF ETC mRNA AND PROTEIN

Consider a *single* molecule of wild-type mtDNA which, when transcribed, generates mRNA for the electron transport chain (ETC), which we denote as m_{ETC} . Transcripts are generated according to a deterministic process (stochasticity in gene expression [18] is neglected in this picture) with rate (β) and also passively degrade at some basal rate (δ_m^b). We consider a controlled, active degradation process (δ_m^a) that acts in addition to the background level. Thus, at the single mtDNA level, we may write down the differential equation

$$\frac{dm_{ETC}}{dt} = \beta - (\delta_m^a + \delta_m^b) m_{ETC}, \quad (\text{B.25})$$

where we assume that β, δ_m^b are constant.

Control of the expression levels of different mitochondrial genes is manifest at the level of mRNA degradation [258], because mtDNA is transcribed as a single polycistronic transcript [257]. We therefore use the simplifying assumption that, in the pathogenic case, mitochondrial mRNA is also controlled at the level of degradation. Thus we allow the active degradation to vary with heteroplasmy $\delta_m^a = \delta_m^a(h)$, and assume the transcription rate to be constant.

Cells are measured at steady-state, so setting the derivative to zero yields

$$m_{ETC} = \frac{\beta}{\delta_m^a + 1}, \quad (\text{B.26})$$

where δ_m^b has been absorbed into the definitions of β and $\delta_m^a(h)$. Assuming N^+ scaling, whereby only wild-type mtDNAs contribute to the transcript pool, we arrive at an expression for *cellular* levels of ETC mRNA:

$$M_{ETC} = \frac{\beta}{\delta_m^a + 1} N^+.$$

Dropping the a superscript yields Eq.(2.5).

Note that, in our Bayesian inference, we chose to express the constant h_0 in Eq.(2.6) in terms of the critical heteroplasmy h^* using the expression $h_0 = h^* - \ln[(1 - f_m)/f_m]/k_m$, where

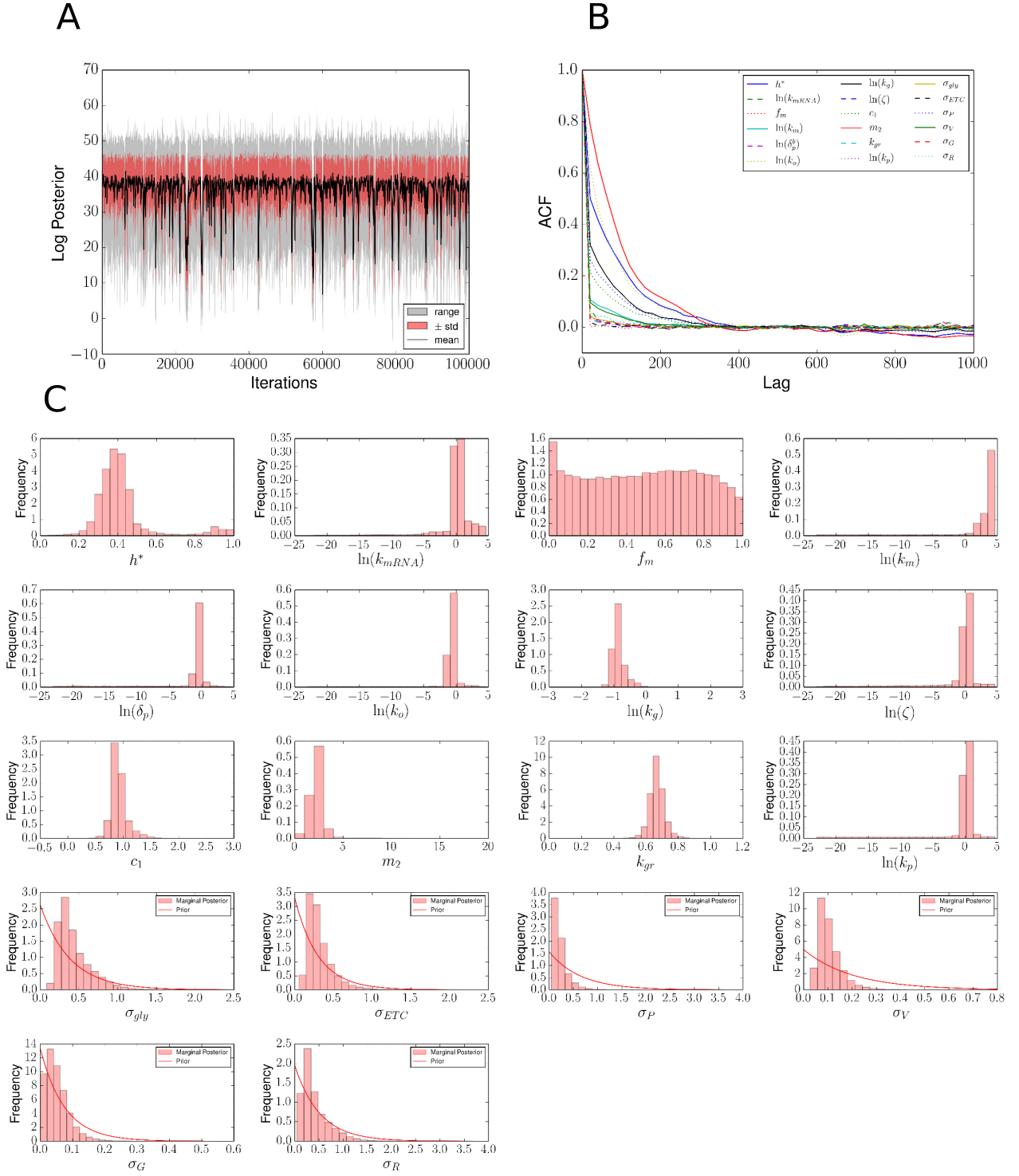


Figure B.4: Metropolis-Hastings posterior samples for main model 10^{10} iterations of Metropolis-Hastings were performed, which were thinned to 10^5 samples. The hyperparameter for model uncertainty, Ω , was chosen as $\Omega = 2$ (see Section B.2). (A) Trajectory of the unnormalised log posterior after thinning. Samples are split into bins of 100 iterations; displaying mean, standard deviation and range of each bin. (B) Autocorrelation function for each parameter, on thinned samples. (C) Approximate marginal posterior distributions for each parameter in the model. Exponential priors for model uncertainties are plotted, all other priors are uniform (see Section B.2).

$0 < f_m \leq 1$. This simply expresses the location of h^* in terms of the fraction f_m of the sigmoid's maximal value. Intuitively, if f_m and k_m are sufficiently large, h^* signals the beginning of reduction in ETC degradation.

For ETC protein, we assume that the following equation holds at the cellular level

$$\frac{dP^+}{dt} = \lambda M_{\text{ETC}} N^+ - \delta_p^b P^+ \quad (\text{B.27})$$

where $\lambda = \text{const}$, and we assume there is no active degradation of ETC protein. At steady-state

$$P^+ = \frac{M_{\text{ETC}} N^+}{\delta_p^b} \quad (\text{B.28})$$

where λ is absorbed into the definition of δ_p^b . Dropping the b superscript yields Eq. (2.7).

B.3.2 JUSTIFICATION OF CELL VOLUME SCALING FOR POWER DEMAND

A reasonable general model for the way in which power demands of a mammalian cell scale with its volume (V) is $k_1 V + k_2 V^{2/3} + K + f(N, V)$, where k_i are proportionality constants. Each term may be interpreted as: $k_1 V$ are demands which scale with cell volume; $k_2 V^{2/3}$ scale with cell surface area; K are demands which are constant for a cell (for example, the cost of replicating the genome); and $f(N, V)$ is an unknown function corresponding to proton leak, which potentially depends upon mitochondrial mass (or alternatively mtDNA copy number, N) and cell volume. These are the dominant power demands of mammalian cells, as determined by [254–256].

Many energy-consuming processes in mammalian cells directly depend upon cell size; for example, a model system used by Buttgereit *et al.* found that $\sim 30\%$ of oxygen consumption corresponded to plasma membrane transporters, and $\sim 20\%$ corresponded to protein synthesis [254]. We assume that protein synthesis scales proportionally with V , because 60% of total cellular dry mass is protein [342] and mass scales with volume. Also, we may assume that plasma membrane power consumption scales with cell surface area, which scales with $V^{2/3}$. By using volume and surface area contributions alone, we may account for $\sim 50\%$ of the power demands of the cell, which is 63% of the accountable power demands for this model system since only 80% of total respiration rate could be attributed to particular processes in their study [254].

Thus, assuming that all power consumption is due to surface area or volume contributions then, using the data of [254], a reasonable model for power demand might be $f(V) = 0.4V + 0.6V^{2/3}$,

the volume parameter being $20/(20 + 30) = 0.4$. However, we see in Fig. 3.1D that the normalised volume data lie in the range $0.6 \lesssim V \lesssim 1$. In this region, the functions $f(V)$ and $g(V) = V$ are similar, with a difference of no more than $\sim 7\%$. Thus, $g(V) = V$ is a reasonable approximation for total power demands in this case.

We note, however, that the above proportions depend on the environment of the cell [254, 255] as well as the tissue type (reviewed in [256]), often showing variation on the order of tens of percent. In light of this uncertainty, and for the sake of parsimony, we make the simplifying assumption that total power demand scales purely with cell volume, see Eq.(2.9).

B.3.3 EXPECTED CELL VOLUME AND GROWTH RATE

Two of the simplest models for how cells may grow throughout the cell cycle are linear and exponential growth. We show below that a relationship exists between growth rate and the mean cell volume in an asynchronous population of cells under a linear model. Furthermore, assuming an exponential model, growth rate and mean cell volume are independent.

Firstly, we assume that the number of cells obey a pure-birth process, in other words the death rate of cells is negligible. If the initial number of cells (N_0) is large, then we can use a deterministic model of cell growth, $N(t) = N_0 \exp(Gt)$, where $N(t)$ is the number of cells at time t and G is the growth rate of cells, as described in Eq.(2.10). Assuming that the number of cells doubles every cell cycle period (t_d), then

$$G = \frac{\ln(2)}{t_d}. \quad (\text{B.29})$$

Under a linear model of cytoplasmic growth through the cell cycle, the volume of an individual cell may be written as

$$V_c(t) = V_0 + \lambda t \quad (\text{B.30})$$

where V_c is the volume of an individual cell, V_0 is the volume of a cell just after division (assumed to be constant for all cells) and $\lambda = V_0/t_d = \text{const}$ is the cytoplasmic growth rate. We note that, by our assumption of $\lambda = \text{const}$ and a linear growth model, smaller cells require less time to double in volume and consequently will proliferate faster (as shown below). From an energetics perspective, a simple model is that cellular power demand relating to proliferation scales with λ , i.e. more energy per unit time is required to increase the cytoplasmic growth rate. Since we model $\lambda = \text{const}$, we expect power demands associated with proliferation to also be constant regardless of the volume of the cell (and hence independent of heteroplasmy in the

model presented in Section 3.4.3). We use the parsimonious ansatz that this power demand is small relative to the power demands of maintaining cytoplasmic volume, and hence we neglect a λ -dependent term in Eq.(2.10).

In an asynchronous population, we assume that each cell is distributed uniformly through the cell cycle, in other words

$$T \sim \text{unif}(0, t_d) \quad (\text{B.31})$$

where T is a random variable describing the position in time, of a cell in its cell cycle.

We wish to find the expected value of cell volume, ($\mathbb{E}(V_c) \equiv V$, as described in Eq.(2.9)), given the assumption of Eq.(B.31). Eq.(B.30) can be viewed as a transformation of the random variable T . If X is a continuous random variable, then for any transformation $Y = r(X)$, $\mathbb{E}(Y) = \int r(x)P(x)dx$, where $P(x)$ is the probability distribution corresponding to the random variable X [343]. This implies that $\mathbb{E}(V_c) = 1/t_d \int_0^{t_d} V_c(t)dt$. For Eq.(B.30), this yields $\mathbb{E}(V_c) = 3V_0/2$, but since $V_0 = \lambda t_d$, then Eq.(B.29) yields

$$G = \frac{3\lambda \ln(2)}{2\mathbb{E}(V_c)}, \quad (\text{B.32})$$

i.e. $1/G \propto \mathbb{E}(V_c) \equiv V$.

If, however, we assume an exponential model of cell growth through the cell cycle

$$V_c(t) = V_0 \exp(\gamma t) \quad (\text{B.33})$$

then $t_d = \ln(2)/\gamma$ and $\mathbb{E}(V_c) = V_0/\ln(2)$ which cannot be written in terms of t_d and therefore G is independent of V .

The above makes intuitive sense: if a cell grows linearly, then a larger cell will need more time to double in size than a smaller cell, if their growth rates are the same. On the other hand, if a cell grows exponentially, then regardless of its initial size, the doubling time is constant, given a fixed cytoplasmic growth rate.

Since there is presumably a wide class of cell growth dynamics where cell size is dependent on growth rate, we favoured a linear model for its simplicity. Measurements by Tzur *et al.* show that, on average, under both a linear and exponential model of cytoplasmic volume growth, the rate constant varies with time [344]. However, the implication of this for the relationship between V and G remains unclear.

B.4 EXPERIMENTAL TESTS FOR MODEL OF MELAS HETEROPLASMY

Below, we will propose potential experiments to test the corresponding claims made in Section 3.4.5.

B.4.1 WILD-TYPE mtDNA DENSITY HOMEOSTASIS IS MAINTAINED UNTIL A MINIMUM VOLUME IS REACHED AT THE CRITICAL HETEROPLASMY

If N^+/V is a quantity kept under homeostasis, then under wild-type conditions, perturbations to mtDNA copy number may be expected to cause changes in cell volume. This might be testable by reducing mtDNA copy number with chemicals such as ddC, or increasing it through PGC-1 α overexpression, which, in the absence of other homeostatic effects, we expect to reduce and increase mean cell volume respectively.

A second testable prediction is that a minimum cell volume ($V_{\min}$) causes bioenergetic toggling at h^* . We should be aware of the two potential interpretations of $V_{\min}$ raised: (1) bioenergetic and (2) mechanical. If $V_{\min}$ is bioenergetic, then raising the power demands of the cell which do not scale with volume, may induce h^* to be encountered earlier. This could be achieved by increasing the amount of DNA in the nucleus which must be replicated, creating a one-time cost to the cell per cell cycle. This is not expected to affect any mechanical constraints since DNA content is not directly indicative of nuclear size [345]. Ideally, the amount of DNA introduced should be large (i.e. billions of base pairs), be replicated, and not interfere with normal functioning of the nucleus. This could be achieved by chemically inducing polyploidy, for instance by using Noscapine [346]. The location of h^* could again be determined by performing RNA-seq, and observing the upregulation of glycolysis with heteroplasmy.

If increasing the power demands of the nucleus yields no change in the distribution of h^* , then a mechanical constraint could be more relevant. This could be tested by perturbing cell volume. Reducing cell volume under this hypothesis is expected to shift h^* to lower values of heteroplasmy, which could be determined via RNA-seq.

B.4.2 MITOCHONDRIAL tRNAs ARE ENRICHED IN THE VICINITY OF THEIR CORRESPONDING PARENTAL mtDNA

The spatial distribution of mitochondrial tRNAs relative to mtDNA would be most directly determined by fluorescent labelling of mitochondrial tRNA and mtDNA. MtDNA labelling

could be achieved through picoGreen staining [347]. Labelling of processed tRNAs within mitochondria is more difficult, but methods exist for labelling mRNA in both fixed cells [348] and dynamically [186, 349] within mitochondria, which may be informative. The experimental suggestion put forward Busch *et al.* [119] to probe the strength of the genotype-phenotype link, via super-resolution microscopy [171], by determining the existence of focal deficiencies in ETC protein concentration would also be illuminating.

B.4.3 MUTANT mtDNAs HAVE A TRANSCRIPTIONAL DEFECT

If mutant mtDNAs have a transcription defect, then the abundance of mRNA encoding mutant tRNAs relative to mRNA encoding wild-type tRNAs would be expected to be smaller than h . This measurement could be performed by RT-qPCR, with probes which target the single-nucleotide polymorphism associated with the 3243A>G mutation, to probe the ratio of mRNA derived from mutant and wild-type mtDNAs. Superresolution microscopy to determine the existence of focal deficiencies in ETC protein would also potentially indicate the existence of a local genotype-phenotype link, which would support this hypothesis [119, 171].

B.4.4 CELL VOLUME IS NOT EXPLAINED BY CELL CYCLE VARIATIONS

To separate the potential confounding influence of the cell cycle on mean cell size, heteroplasmic cells could be transfected with Fucci markers [350], and relative enrichment of cell cycle stages determined.

B.4.5 CELLS PROLIFERATE INVERSELY WITH THEIR SIZE

To determine the dependence of growth rate on mean cell volume, wild-type cells could be synchronised, and sorted by their volume. These cells could then be plated and released from synchronisation, and the growth rate of cells measured similar to that described in Section 3.3. Synchronisation is necessary, because cell volume is expected to vary by a factor of 2 through the cell cycle, so any sorting would otherwise be strongly confounded by the cell cycle. A potential alternative to synchronisation, which can be stressful to cells, is to label genes associated with a particular stage of the cell cycle, and sort based on both this fluorescence signal and cell volume.

Claim	Experiment
MtDNA copy number affects volume	Perturb mtDNA copy number (ddC or PGC1- α), measure volume
Wild-type mtDNA density affects h^* (bioenergetic)	Increase nDNA content, perform RNA-seq
Wild-type mtDNA density affects h^* (mechanical)	Reduce cell volume, perform RNA-seq
Mutant mtDNAs have transcription defect	Measure normal/mutant tRNA abundance with RT-qPCR
Mitochondrial tRNAs have low diffusivity	Fluorescent labelling of mtDNA and tRNA or mRNA encoding tRNA
Cell cycle variation with heteroplasmy	Fucci markers in heteroplasmic cells
Mean cell volume affects growth rate	Synchronise, sort by volume, and measure growth rate
Maximum respiratory capacity $\propto$ ETC protein	Measure $R_{\max}$ at $h = 0$ and P^+ at $h = 0.0, 0.6$

Table B.1: Summary of experimental proposals, corresponding to the claims of the model.

B.4.6 MAXIMUM RESPIRATORY CAPACITY LINEARLY TRACKS ETC PROTEIN CONTENT

Measurements of maximum respiratory capacity at $h = 0$, as well as measurement of ETC protein levels at $h = 0.6$ and $h = 0$, may help determine whether a simple linear relationship is sufficient, or whether a more complex model is justified.

B.4.7 SUMMARY

A summary of the experimental proposals outlined are given in Table B.1

B.5 ALTERNATIVE MODELS

B.5.1 MUTANT MTDNA AND TRANSCRIPTION

Eq.(2.5) states that mutant mtDNAs do not contribute significantly to the transcript pool. We can relax this constraint by replacing Eq.(2.5) with

$$M_{\text{ETC}} = \frac{\beta}{\delta_m + 1} (N^+ + \mu N^-) \quad (\text{B.34})$$

where $0 \leq \mu \leq 1$ and $N^- = hN$, where N is the total number of mtDNAs, which we treat as a constant.

Using the uniform prior

$$P(\mu) = \text{unif}(0, 1), \quad (\text{B.35})$$

we sampled from the posterior, as described in Section B.2. The MCMC trajectory is shown in Fig. B.5. The marginal posterior density for μ in Fig. B.5, shows that $\mu \approx 1$ is the most likely value of the parameter, in other words mutant mtDNAs contribute equally to the transcript pool, compared to wild-type molecules. However, draws from the posterior distribution of M_{ETC} were often purely linear and thus inappropriate for understanding threshold effects (see Fig. B.6). We consequently rejected this model in favour of the model presented in Section 3.4.3 (see Fig. B.7 for model fit).

B.5.2 tRNA MISINCORPORATION MODEL

Eq.(2.7) states that ETC protein is generated when ETC mRNA is in contact with wild-type mtDNA, suggesting that tRNAs affected by the MELAS mutation, leucine-UUR, remain local to their parent mtDNAs. The alternative is that mitochondrial tRNAs are well diffused amongst mitochondrial mRNAs. If we assume that a mutant tRNA causes a misincorporation during translation with 100% efficiency, then the number of misincorporations per protein follows a binomial distribution. We assume that the probability of a single misincorporation is h . We further assume that proteins have a mutational tolerance of x misincorporations, or less, before they are considered mutated (and consequently degraded). With these assumptions, the expected proportion of mutant proteins (m_p) will be

$$m_p = 1 - F(x|N, h) \quad (\text{B.36})$$

where N is the number of leucine-UUR residues per protein, and $F(x|N, h) = P(X \leq x)$ is the cumulative distribution function of the binomial distribution, for N trials, x successes, and probability of success h . A plot of m_p is given in Fig. B.8, for different mutational tolerances x against heteroplasmy.

We can therefore use an analogous expression to Eq.(2.7) for P^+ , in the case of well-diffused tRNAs

$$P^+ = \frac{M_{\text{ETC}} F(x|N, h)}{\delta_p^b}. \quad (\text{B.37})$$

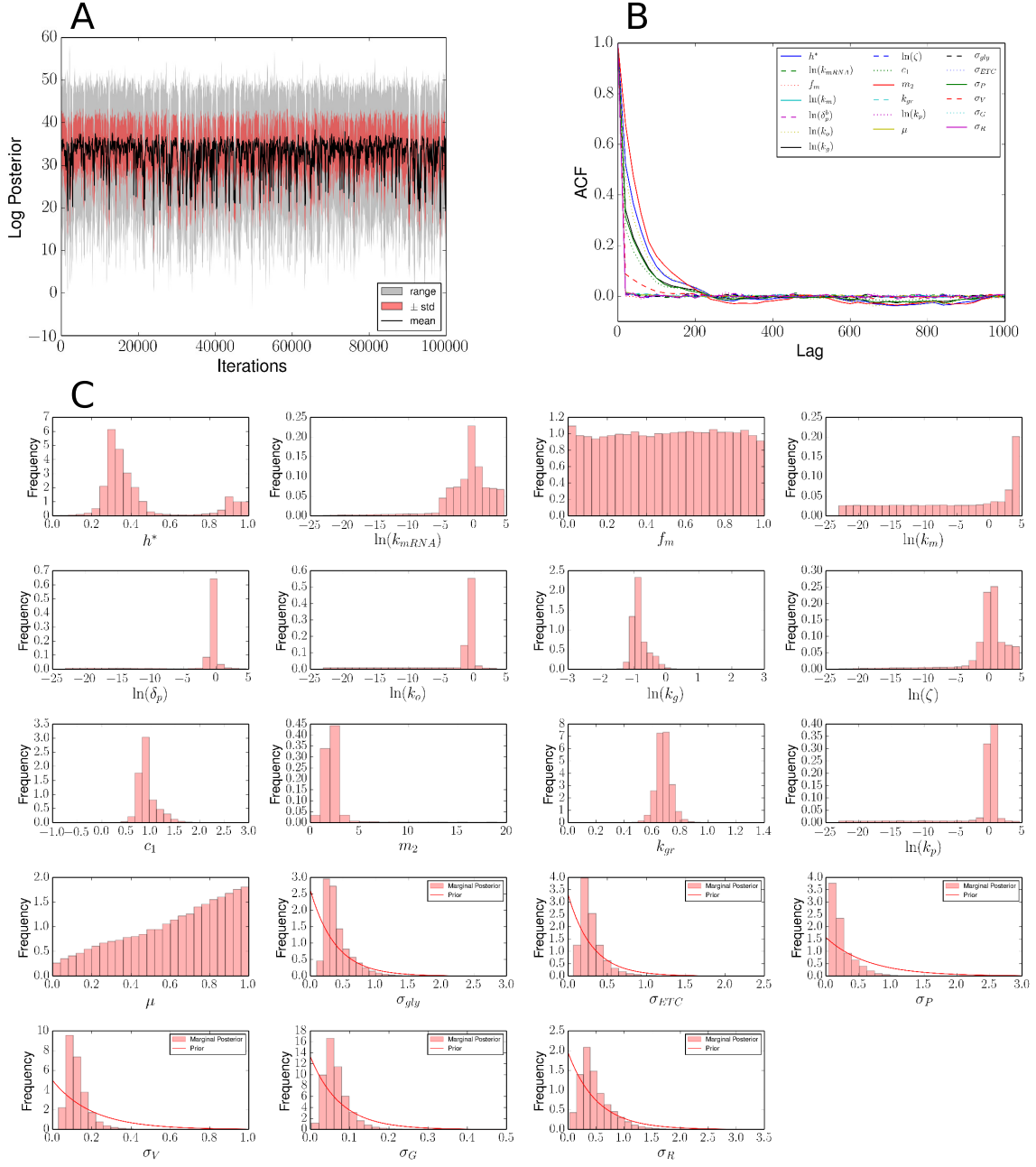


Figure B.5: Metropolis-Hastings posterior samples for mutant transcription model 10^{10} iterations of Metropolis-Hastings were performed, which were thinned to 10^5 samples. The hyperparameter for model uncertainty, Ω , was chosen as $\Omega = 2$ (see Section B.2). The prior for the additional parameter μ was chosen to be a uniform distribution between 0 and 1. **(A)** Trajectory of the unnormalised log posterior after thinning. Samples are split into bins of 100 iterations; displaying mean, standard deviation and range of each bin. **(B)** Autocorrelation function for each parameter, on thinned samples. **(C)** Approximate marginal posterior distributions for each parameter in the model. Exponential priors for model uncertainties are plotted, all other priors are uniform (see Section B.2). Note the marginal distribution of μ peaks near 1.

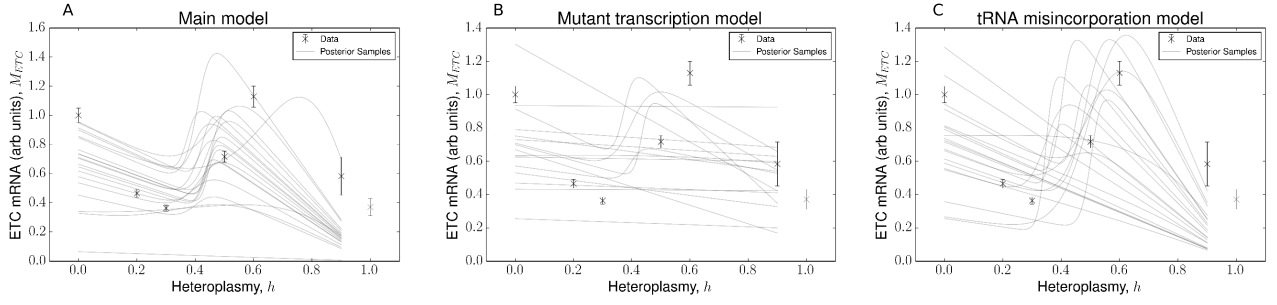


Figure B.6: Comparison of ETC mRNA levels for alternative models. Sample of 20 randomly sampled trajectories from the posterior distribution of M_{ETC} for (A) Main model (see Section 3.4.3); (B) Mutant transcription model (see Section B.5.1); (C) tRNA misincorporation model (see Fig. B.5.2). In models (B) and (C), we find that linear fits are more frequently selected when compared to our main model (A), despite being more complex in terms of number of parameters.

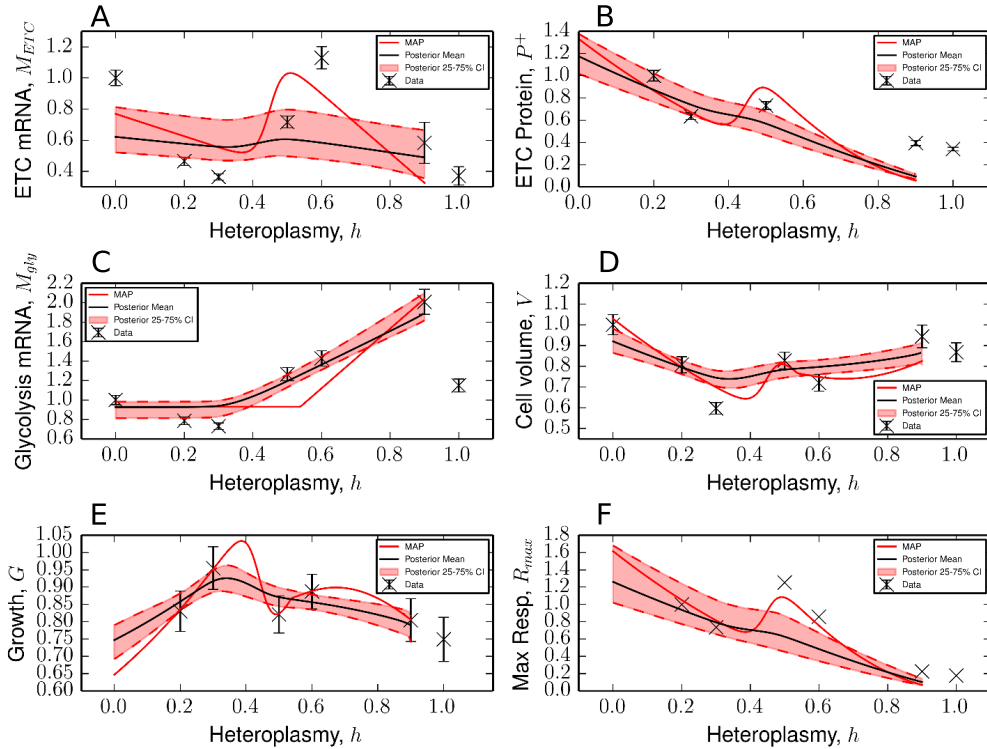


Figure B.7: Alternative mutant transcription model. (A-F) Model fit when Eq.(2.5) is replaced with Eq.(B.34). The 25-75% CI is flatter when compared with Fig. 3.5. We find this is due to the model more frequently selecting linear fits to the data, see Fig. B.6.

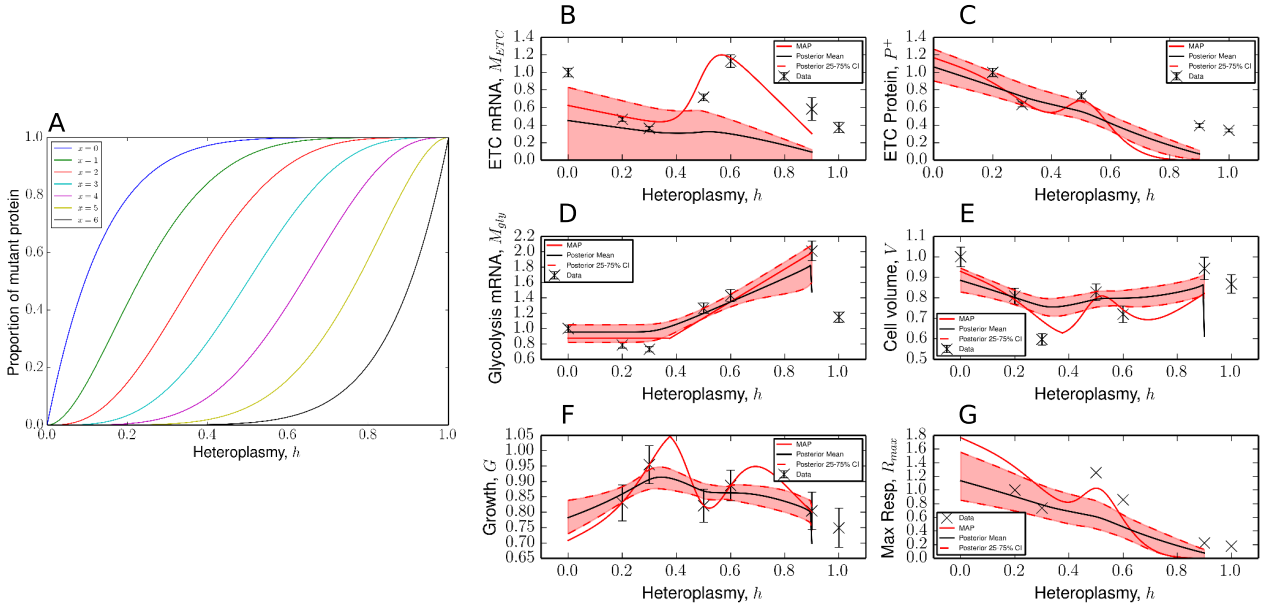


Figure B.8: Alternative tRNA misincorporation model (A) Expected proportion of mutant protein due to the MELAS mutation, given that each protein can tolerate x mutated residues. The chain length used is $N = 8$, which is the mean number of susceptible residues across all mitochondrially-encoded peptides, excluding ATP8 and ATP6 [249]. **(B-G)** Model fit when Eq.(2.7) is replaced with Eq.(B.37). M_{ETC} is qualitatively fitted more poorly than Fig. 3.5 (see also Fig. B.6), although the maximum a posteriori estimate is a closer fit, when compared with Fig. 3.5.

By replacing Eq.(2.7) with Eq.(B.37), we again sampled from the posterior as described in Section B.2. We chose $N = 8$, which is the average number of susceptible residues in the 11 mitochondrially-encoded subunits considered (see caption of Fig. 3.1) [249]. Our prior for the unknown tolerance to misincorporations, x , was chosen as a discrete uniform prior

$$P(x) = \begin{cases} \frac{1}{9}, & x = 0, 1, \dots, 8 \\ 0, & \text{otherwise.} \end{cases} \quad (\text{B.38})$$

The MCMC trajectory is shown in Fig. B.9, and the model fit in Fig. B.8. Draws from the posterior distribution of M_{ETC} were often purely linear and thus inappropriate for understanding threshold effects (see Fig. B.6). We consequently rejected this model, in favour of the model presented in Section 3.4.3. Furthermore, observing the marginal posterior distribution of the misincorporation tolerance x in Fig. B.9, we see that the most likely value of the parameter is $x = N = 8$. In other words, ETC proteins are immune to the MELAS mutation, which we believe to be incorrect [249]. For these reasons, we rejected this model in favour of the model presented in Section 3.4.3.

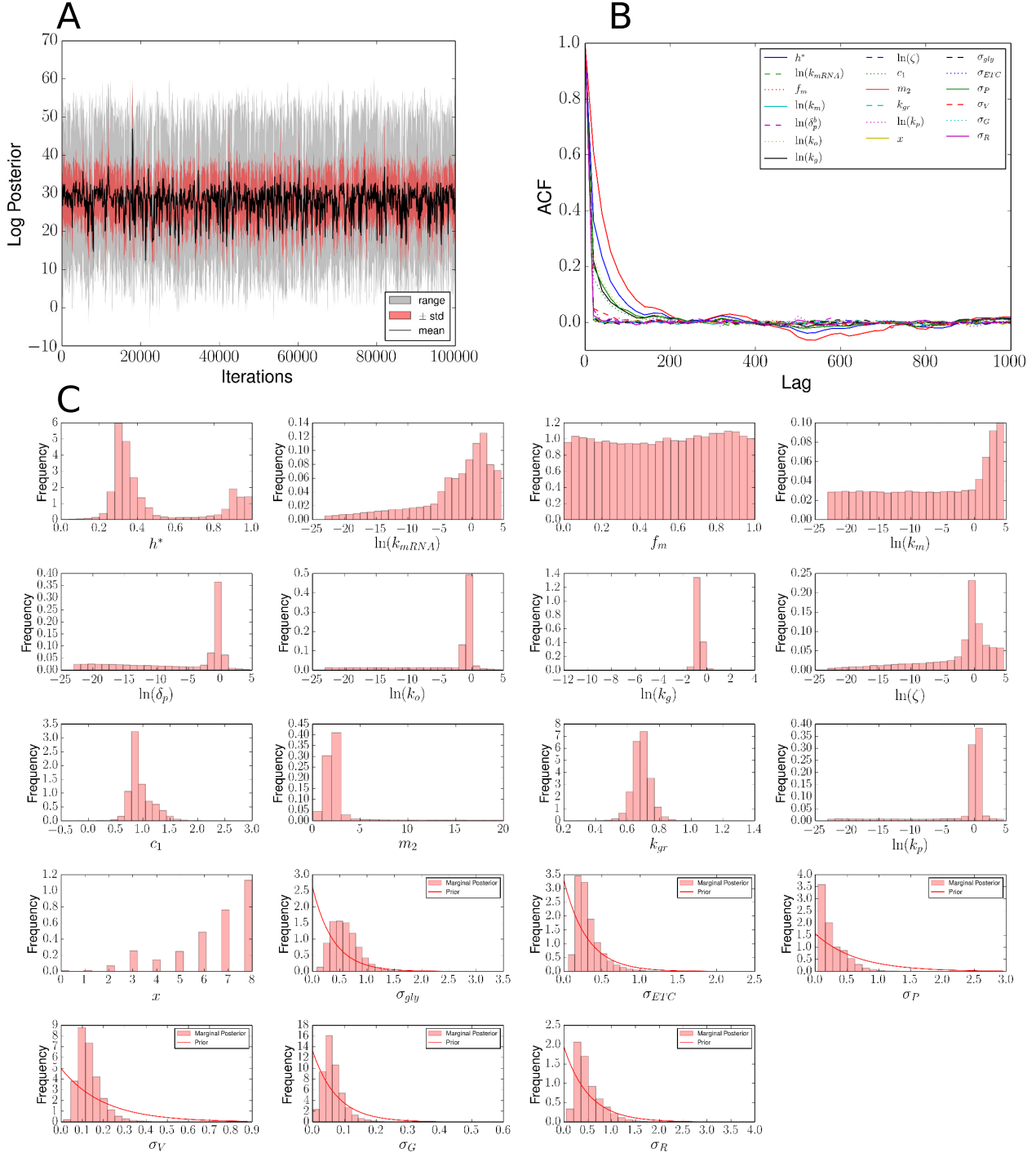


Figure B.9: Metropolis-Hastings posterior samples for tRNA misincorporation model 10^{10} iterations of Metropolis-Hastings were performed, which were thinned to 10^5 samples. The hyperparameter for model uncertainty, Ω , was chosen as $\Omega = 2$ (see Section B.2). The prior for the additional parameter x was chosen to be a uniform discrete distribution between 0 and 8. **(A)** Trajectory of the unnormalised log posterior after thinning. **(B)** Autocorrelation function for each parameter, on thinned samples. **(C)** Approximate marginal posterior distributions for each parameter in the model. Exponential priors for model uncertainties are plotted, all other priors are uniform (see Section B.2). Note the marginal distribution of x peaks at 8, which we expect to be incorrect.

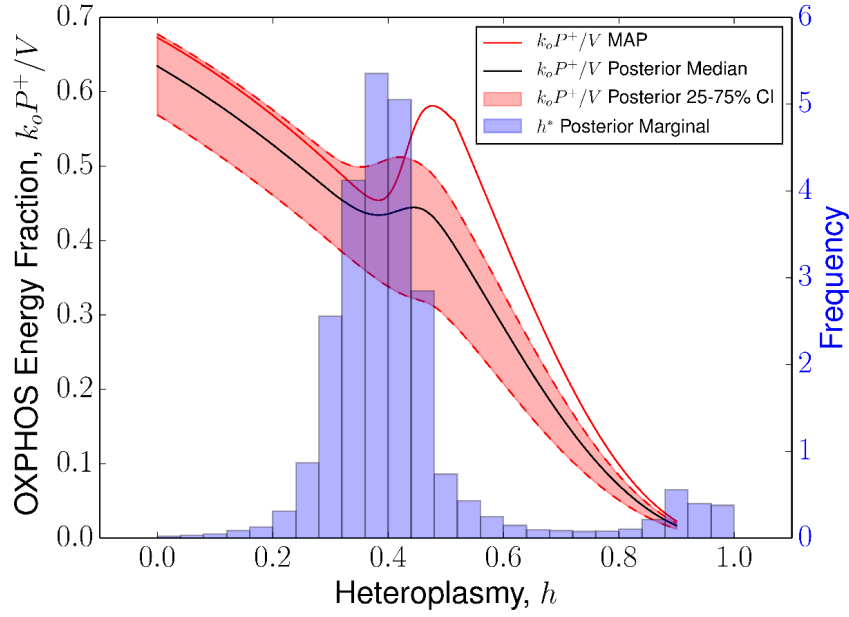


Figure B.10: Relative contribution of OXPHOS to total power supply across heteroplasmy Posterior statistics for the ratio of OXPHOS power supply ($k_o P^+$) to total power supply $k_o P^+ + k_g M_{\text{gly}} = V$. The contribution of ETC power production reduces until the critical heteroplasmy h^* , where a compensatory response stabilises OXPHOS contributions. As $\delta_m \rightarrow 0$ at $h \approx 0.5$ (see Fig. 3.8), OXPHOS power contributions continue to diminish.

B.6 RELATIVE OXPHOS CONTRIBUTION TO POWER SUPPLY

It is interesting to observe the relative contributions of oxidative phosphorylation and glycolysis to power supply. Since Eq.(2.9) states that power supply = demand, where demand corresponds to cell volume, the ratio $f_o = k_o P^+ / V$ determines the relative contribution of OXPHOS to power supply, see Fig. B.10.

For $h < h^*$, we see that OXPHOS has decreasing contributions to power supply. At h^* , OXPHOS contributions stabilise with $0.28 < f_o(h = 0.52 \approx h_{\text{MAP}}^*) < 0.44$ (25-75% CI). The heteroplasmy at which OXPHOS contributions are stabilised corresponds to the hypothesised demand/supply toggle, where the cell attempts to increase power supply as opposed to reducing power demand.

The value of f_o where OXPHOS contributions become stabilised ($f_o(h^*)$) may have wider significance. Mitochondrial metabolism, and especially mitochondrial membrane potential, is connected to a variety of biosynthetic pathways [351] and crucial for maintaining cellular proliferation [250]. $f_o(h^*)$ may represent a minimum ETC flux, relative to power demand, for mitochondria to support their mitochondrial membrane potential without the aid of glycolytic ATP. Below $f_o(h^*)$, we might predict that cells run ATP synthase in reverse, hydrolysing

glycolytic ATP to maintain membrane potential.

C

SUPPLEMENTAL INFORMATION COMPLEMENTING CHAPTER 4

C.1 CORRELATION BETWEEN MITOCHONDRIAL MASS AND CELL VOLUME IS NOT STRONGLY CONSTRAINED

The metabolic scaling argument presented in Section 4.4.1 relies on mitochondrial content scaling exactly as v^1 . At the level of the mean, Miettinen & Björklund [50] found a remarkably strong linear correlation between JC-1 monomer fluorescence and total cellular protein ($R^2 = 0.99$, Ref. [50] Fig. 1B), which can be loosely considered to reflect mitochondrial mass and cell volume respectively. However, due to the complexities raised in Section 4.3, quantitative use of this probe may have issues with reliability. To probe this relationship directly, we attempted to fit models of the form $n \propto v^\alpha$ for unknown α using the MitoTracker data from Miettinen & Björklund (Ref. [50] Fig. 1D). We found that these data did not strongly constrain this more general relationship (see Fig. C.1 below) due to a low signal to noise ratio, so a linear correlation between mitochondrial mass and cell volume remains a potentially parsimonious model.

In this case, it is unclear why mitochondrial mass does not compensate for power demand variation with changing cell size, instead of $r\Delta\Psi$. Correlation between n and v has been observed repeatedly [24, 44, 45, 49] and in Chapter 3 we suggested that the inability to maintain wild-type mtDNA density homeostasis may account for the cellular response to mutation in

mitochondrial DNA. We highlight investigation of the constraints on mitochondrial density homeostasis, and refinements of the precise scaling relationship between n and v , as areas of further research.

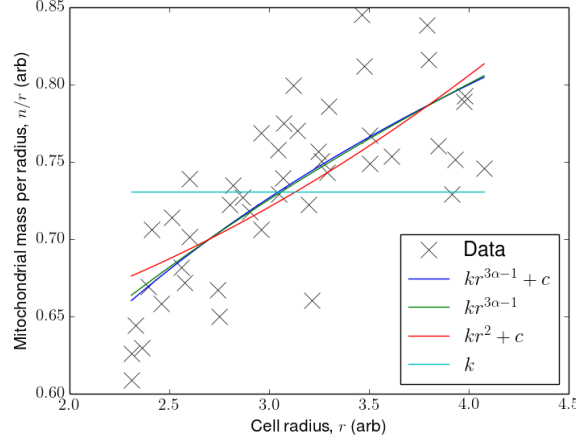


Figure C.1: Data does not strongly constrain the scaling relationship between n and v due to a small dynamic range of volume relative to noise. Median mitochondrial mass (n) per cell radius (r) is given as a function of forward scatter, which correlates with r . Assuming $n \propto v^\alpha$ and $v \propto r^3$ yields $n/r \propto r^{3\alpha-1}$. Several variations of this model are fitted to data from Miettinen & Björklund (Ref. [50] Fig. 1D) using nonlinear least squares. The models perform similarly, with the exception of $n = kr$ which has the worst performance.

C.2 SIMULATING METABOLIC SCALING AND CELL DEATH

To simulate the metabolic scaling relationship $r\Delta\Psi = \min\{r\Delta\Psi_s, r\Delta\Psi_\mu\}$ (see Eq. (4.3) and Eq. (4.5)), we simulated cells growing exponentially according to $v = v_0 e^{kt}$, obeying adder dynamics. The death rate in Eq.(4.4) defines an inhomogeneous Poisson process with rate $\mu(t)$. Defining $P_0(s)$ as the probability of zero death events in the time interval $[t, t + s]$, and $m(t) = \int_0^t \mu(s) ds$, then $P_0(s) = \exp(-[m(t + s) - m(t)])$ for some fixed time t [324]. By setting $t = 0$ at the birth of each cell, it follows from Eq.(4.4) that

$$P_0(s) = \exp\left(-\left[c_\mu s - \frac{m_\mu v_0}{k}[\exp(ks) - 1]\right]\right) \quad (\text{C.1})$$

where m_μ and c_μ are defined in Eq.(4.4). This equation holds until $v > c_\mu/m_\mu$, from which point the death rate is zero for the remaining lifetime of the cell until division (Eq. (4.5)). If $v_0 < c_\mu/m_\mu$, cell survival is determined by evaluating $P_0(s)$, where the interval s is either the time to division or the first time where $\mu(t) = 0$, whichever occurs first. At division, cells divide

their volume according to

$$v_1 = \mathcal{N}\left(\frac{v_f}{2}, \sigma_v\right) \quad (\text{C.2})$$

$$v_2 = v_f - v_1 \quad (\text{C.3})$$

where $\mathcal{N}(m, \sigma)$ is a Gaussian distribution with mean m and standard deviation σ ; v_f is the volume of a cell just before division; and v_1, v_2 are the volumes of the daughter cells. The volumes are resampled until conservation of mass is obeyed, which effectively samples from a truncated normal distribution. Finally, cells are simulated up to a threshold measurement time, and cells (along with their daughters) are updated iteratively. Finally, each cell is assigned a value of $r\Delta\Psi$ at the end of the simulation according to $r\Delta\Psi = \min\{r\Delta\Psi_s, r\Delta\Psi_\mu\}$, with Gaussian noise.

A table of parameters is shown in Table C.1. The parameters for Eq. 4.1 were loosely based on the study of Buttgereit *et al.* [293], who found that protein synthesis, plasma membrane transporters and RNA/DNA synthesis contribute to 20%, 20% and 15% of total cellular respiration rate respectively for quiescent thymocytes. We approximate these as scaling with cell volume, cell surface area, and constant per cell respectively, and assume that these values are comparable in dividing cells. Neglecting other contributions to total power demand (mainly proton leak, 40%, whose scaling with cell size is unclear), implies that these contribute to 36%, 36% and 27% of total cellular power demand which is accounted for in this model. We chose k_v such that $K_v k_v / (K_v k_v + K_v^{2/3} k_s + k_c) \approx 36\%$, which is the fraction of cellular power demand attributable to cellular volume, for a typical cell of volume $v = K_v$. Similarly, $K_v^{2/3} k_s / (K_v k_v + K_v^{2/3} k_s + k_c) \approx 36\%$ (power demand of cell surface area) and $k_c / (K_v k_v + K_v^{2/3} k_s + k_c) \approx 27\%$ (constant power demands). Although we have neglected proton leak, it is unlikely that the addition of this term will change these proportional power demands by orders of magnitude, each showing contributions in the range of tens of percent. We note that the specific choices for these parameters do not have strong implications for our results, as the model has a sufficient number of unconstrained degrees of freedom to replicate the results of [50], through parameters such as $m_\psi, c_\psi, \sigma_\psi$ and so on, whilst also maintaining biophysical plausibility. It is an area of future study to infer such parameters in a more rigorous fashion, given the data.

Parameter	Value	Description	Unit
-	20000	Initial number of cells	-
-	unif(0,0.04)	Initial volume distribution	arb
-	3.47	Total simulation time	T
k	1.2	growth rate	T^{-1}
K_v	0.02	Cell volume increase over cell cycle	arb
c_μ	1.0	Intercept for cell death vs volume	T^{-1}
m_μ	66.6	Gradient for cell death vs volume	T^{-1}
σ_v	0.005	Partitioning error at division	arb
k_v	1.375	Volume parameter in power demand	arb
k_s	0.373	Surface area parameter in power demand	arb
k_c	0.021	Constant parameter in power demand	arb
m_ψ	4.3	Gradient for $r\Delta\Psi$ vs cell death	T
c_ψ	4.3	Intercept for $r\Delta\Psi$ vs cell death	arb
σ_ψ	1.0	Gaussian measurement SD for $r\Delta\Psi$	arb

Table C.1: Parameters for metabolic scaling simulation. Units are T: simulation time, arb: arbitrary.

C.3 COEFFICIENT OF VARIATION OF RELATIVE MITOCHONDRIAL MEMBRANE POTENTIAL AND CELL SIZE

We investigated the hypothesis that the coefficient of variation ($CV_{r\Delta\Psi}(v)$ or just CV) sets the characteristic cell size (as postulated by Miettinen & Björklund [50]), using the metabolic scaling argument presented in Section 4.4.1. Cells were binned *in silico* according to their volume, and the coefficient of variation for $r\Delta\Psi$ was calculated for each bin (Fig. 4.3D).

We use a simple model of variability in $r\Delta\Psi(v)$, which is Gaussian noise with constant SD (accounting for biological and technical sources of noise), in addition to mean behaviour determined by a combination of metabolic scaling and cell death. Since the CV is simply the ratio of the standard deviation to the mean, the CV for $r\Delta\Psi(v)$ (at constant v) is therefore $CV_{r\Delta\Psi}(v) \propto \langle r\Delta\Psi(v) \rangle^{-1}$, where $\langle \dots \rangle$ denotes a mean. Consequently, $CV_{r\Delta\Psi}(v)$ has a minimum which coincides with the maximum $\langle r\Delta\Psi(v) \rangle$, see Fig. 4.3C&D. The data of Miettinen & Björklund shows that the minimum of $CV_{r\Delta\Psi}(v)$ does not coincide with the maximum median $r\Delta\Psi$ with respect to v , see Fig. C.2A. Assuming that the turning point in the median $r\Delta\Psi$ corresponds to that of the mean suggests that the model poses too strong a constraint on the position of the minimum CV.

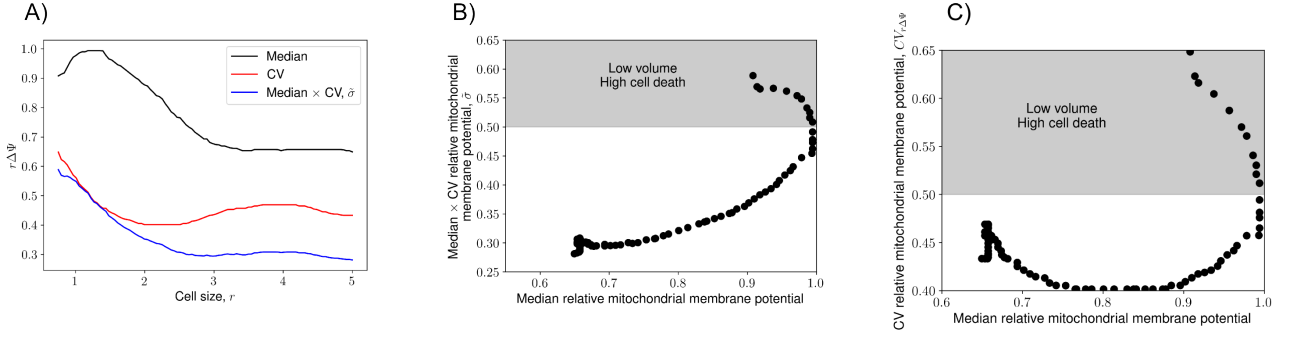


Figure C.2: CV in mitochondrial functionality displays a minimum with respect to the mean, whereas an approximation of the standard deviation does not. Relationship between the variability in $r\Delta\Psi$ and its median, using loess fits of data from Ref. [50] Fig. 1C, 3A. The approximate standard deviation ($\tilde{\sigma}$) is derived by taking the product of the median and the CV for each value of cell radius. **(A)** Direct comparison of the median, CV and $\tilde{\sigma}$ of $r\Delta\Psi$ over cell size. **(B)** Using the data shown in (A), $\tilde{\sigma}$ is plotted against the median $r\Delta\Psi$. The shaded region corresponds to low cell volume and high apoptosis rate. $\tilde{\sigma}$ simply increases monotonically with larger mean values over the region where cell death is negligible. **(C)** The CV of $r\Delta\Psi$ is plotted against the median. The coefficient of variation displays a minimum with respect to the mean. (B) and (C) raise the question of whether the standard deviation or the coefficient of variation is more physiologically relevant in this system, with (C) displaying a minimum whereas (B) does not, and therefore whether variability in mitochondrial functionality sets the “optimal” cell size.

Indeed, it is straightforward to show, using the quotient rule of ordinary calculus, that in order for the turning point of the CV to not coincide with the mean, the standard deviation must have some pure dependence on v . We therefore sought to compare the relationship between the SD, CV and cell radius.

Since the mean relationship between $r\Delta\Psi$ and cell radius (r) was not available, in order to make a preliminary exploration, we approximated the SD by taking the product of the CV and the median – this approximation is referred to as $\tilde{\sigma}$. $\tilde{\sigma}$ will be proportional to the SD if the mean is proportional to the median over all r (although it is quite possible that this is not the case).

Given this approximation, Fig. C.2A shows that $\tilde{\sigma}$ shows pure variation with cell radius (and therefore volume), as expected from the observed lack of alignment between the maximum median $r\Delta\Psi$ and the minimum $CV_{r\Delta\Psi}(v)$. The relationship between the median of $r\Delta\Psi$ and $\tilde{\sigma}$ takes a form where a larger median $r\Delta\Psi$ has a larger $\tilde{\sigma}$ over the region where cell death is small (Fig. C.2B). *This relationship is intuitive, since quantities which have a larger mean often have larger variability.*

In contrast, the relationship between $CV_{r\Delta\Psi}(v)$ and the median $r\Delta\Psi$ does not take a monotonic form (Fig. C.2C). In order to account for this observation (which Miettinen & Björklund speculate may determine the “optimal” cell size [50]), we hypothesize that an analogous, monotonically increasing, relationship (as shown in Fig. C.2B) between the true SD and the mean

$r\Delta\Psi$ holds. If this were the case, then we would be able to account for the seemingly non-intuitive minimum CV (Fig. C.2C) with an intuitive relationship between the mean and the standard deviation of $r\Delta\Psi(v)$ (Fig. C.2B). It therefore remains open as to whether the coefficient of variation or standard deviation in $r\Delta\Psi$ is the more physiologically relevant measure of variability in mitochondrial functionality, and whether either have direct consequences for determining the “optimal” cell size.

C.4 A PARABOLIC FUNCTION OF CELL DEATH AGAINST CELL SIZE AS AN ALTERNATIVE ACCOUNT OF MAXIMAL MITOCHONDRIAL FUNCTIONALITY

Although a combination of cell death and power demand scaling is able to account for a large amount of the data in the study of Miettinen & Björklund [50], we also investigated the hypothesis that cell death alone may be the underlying causal variable for these observations. We considered a simple model where the cell death rate μ varies with cell volume v according to

$$\mu(t) = k_s(v(t) - v_{opt}(t))^2 + k_m \quad (\text{C.4})$$

where cells grow exponentially in time t according to $v = v_0 e^{kt}$ and the optimal volume trajectory is $v_{opt}(t) = v^* e^{kt}$ (i.e. following this trajectory minimises the probability of cell death). k , k_s , k_m are constants and v^* is the optimal initial volume. v_0 is the volume of a cell at the beginning of the cell cycle. Deviations in growth from an optimal trajectory specified by $v_{opt}(t)$ incur a higher probability of cell death. Cells were simulated to undergo “timer” dynamics, whereby cells grow for a fixed amount of time and then divide (see Section C.4.1 for details). Such dynamics generally result in unbounded diffusion of cell volume, however Eq.(C.4) confines cell volumes through cell death. We adopted these dynamics for computational convenience to probe cell cycle effects, because if cells are initially synchronised under these dynamics then they remain so for all time (although our qualitative observations are expected to hold under alternative growth dynamics, see below). We then assumed that $r\Delta\Psi$ is determined through a noisy linear relationship with μ , $r\Delta\Psi = -m_\psi\mu + c_\psi + \mathcal{N}(0, \sigma_\psi)$ (Fig. C.3A).

This model was able to show qualitatively similar behaviour in mean $r\Delta\Psi$ across different cell cycle stages (Fig. C.3C&D) to the data of Miettinen & Björklund (Ref. [50] Fig. 2D), as well as a similar joint distribution in cell volume and $r\Delta\Psi$ (Fig. C.3B, compare with Ref. [50] Fig. 1B). *This model suggests a compatible interpretation of [50] is that $r\Delta\Psi$ is a passive readout of cell death rate, and that the observed biological nonlinearity lies in the relationship between cell*

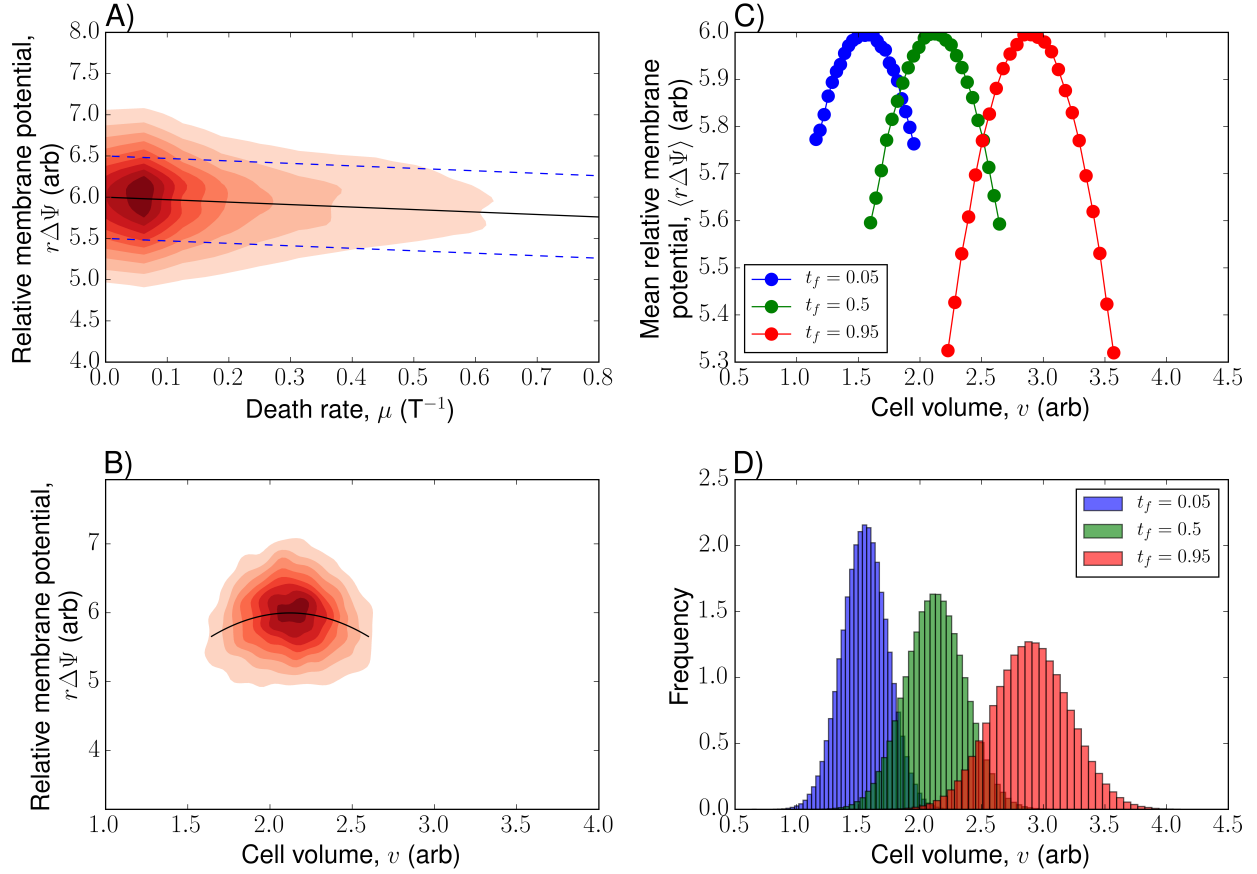


Figure C.3: Optimal cell size, in terms of cell death rate, is able to account for optimal membrane potential Parabolic model of cell death versus cell volume, see Eq.(C.4), simulated using timer cell cycle dynamics. (A) Noisy linear relationship between $r\Delta\Psi$ and μ , $r\Delta\Psi = -m_\psi\mu + c_\psi + \mathcal{N}(0, \sigma_\psi)$, with mean (black) $\pm$ standard deviation (blue dotted). (B) Joint distribution of v and $r\Delta\Psi$ is qualitatively similar to experimental findings (Ref. [50] Fig. 1B). Mean relationship between $r\Delta\Psi$ and v is shown (black curve). All cells are synchronized in their cell cycle stage, and measured 50% through their final cell cycle. (C) During the final generation of growth, cells are measured at different points in the cell cycle, where t_f denotes cell cycle progression as a fraction between 0 and 1. The parabolic relationship between $r\Delta\Psi$ and v shifts to higher values of v with later cell cycle stages, as seen in the data of Miettinen & Björklund (Ref. [50] Fig. 2D). (D) The distribution of cell volumes also shifts to higher v , and becomes broader with the cell cycle, as can be seen in Ref. [50]. This is due to the exponential nature of cell growth amplifying variation in cell size present at birth.

death and cell volume. Although the parabolic cell death model in Eq.(C.4) shows larger cell death for large cells than is suggested by the data of [50], Eq.(C.4) is still a feasible model if cell dynamics are assumed which have an intrinsic optimal volume, such as adder or sizer dynamics. In this case, the steepness parameter of Eq.(C.4) could be reduced, with qualitatively similar implications for $r\Delta\Psi$ due to the absence of direct coupling between $r\Delta\Psi$ and growth dynamics.

It is important to stress that one may invert the causality in the above model, by writing $r\Delta\Psi(t) = k_s(v(t) - v_{opt}(t))^2 + k_m$ and $\mu = -a(r\Delta\Psi) + b$ for constants a and b . If we assume Gaussian noise in $r\Delta\Psi$, then this is equivalent to the model discussed above. It is therefore important to rule out cell death as a potential confounding variable in future confirmatory studies.

C.4.1 SIMULATING PARABOLIC FUNCTION OF CELL DEATH AGAINST CELL SIZE

This simulation proceeds similarly to Section C.2. In this case, μ follows Eq.(C.4). This may be written as $\mu(t) = k_s\tilde{v}\exp(2kt) + k_m$ where $\tilde{v} = (v_0 - v^*)^2$ and

$$P_0(s) = \exp\left(-\left[\frac{k_s\tilde{v}}{2k}(\exp(2ks) - 1) + k_ms\right]\right). \quad (\text{C.5})$$

Cells are initialised with uniformly distributed volumes. All cells are simulated for a fixed number of generations and are removed according to Eq.(C.5). Cell volumes are partitioned between daughters in the same way as Section C.2. Cells are then assigned a value of $r\Delta\Psi$ according to a linear relationship with Gaussian noise. A table of parameters is shown in Table C.2. All simulations and analysis were performed in `Python` and source code is available from the authors upon request.

Parameter	Value	Description	Unit
-	200000	Initial number of cells	-
-	unif(0, 4)	Initial volume distribution	arb
-	6	Total number of generations	-
k	2.0	Growth rate	T^{-1}
k_s	5.0	Death rate steepness	T^{-1}
k_m	0.01	Minimum death rate	T^{-1}
v^*	1.5	Optimal initial volume	arb
σ_v	0.05	Partitioning error at division	arb
m_ψ	0.3	Gradient for $r\Delta\Psi$ vs cell death	T
c_ψ	6.0	Intercept for $r\Delta\Psi$ vs cell death	arb
σ_ψ	0.5	Gaussian measurement SD for $r\Delta\Psi$	arb

Table C.2: Parameters for simulation with parabolic function of cell death. Units are T: simulation time, arb: arbitrary.

D

SUPPLEMENTAL INFORMATION COMPLEMENTING CHAPTER 5

D.1 SUMMARY STATISTICS FOR THE PROPORTION OF INDIVIDUALS WITH A PATHOLOGICALLY MUTATED REGION OF A FIXED SIZE

We may determine the probability of observing one or more regions seeded at generation i^* , given our posterior estimate for λ_m (see Fig. 5.3E). We define Q_{i^*} as the proportion of N individuals with one or more mutated regions from generation i^* . Relabelling $\lambda_m \equiv \lambda$, and indexing each individual by j , corresponding to a single draw from the approximate posterior $\lambda_j \sim P(\lambda_j | \rho(\mathbf{S}, \mathcal{D}))$, we are able to simulate as many individuals (N) as we have available samples from the approximate posterior ($N \approx 10^5$). We may therefore write

$$Q_{i^*} = \frac{1}{N} \sum_{j=1}^N I(X_{i^*,j}), \quad (\text{D.1})$$

which is the approximate posterior mean of the indicator function (defined in Eq.(5.11)). We may then determine the standard deviation of Q_{i^*} . Observing that

$$\mathbb{E}(I(X_{i^*,j})) = 0 \cdot P(X_{i^*,j} = 0) + 1 \cdot P(X_{i^*,j} = 1) + 1 \cdot P(X_{i^*,j} = 2) \dots \quad (\text{D.2})$$

$$= 1 - P(X_{i^*,j} = 0) \quad (\text{D.3})$$

(where $\mathbb{E}()$ denotes an expectation) it follows that, since all $X_{i^*,j}$ are independent,

$$\mathbb{E}(Q_{i^*}) = \frac{1}{N} \sum_{j=0}^N (1 - P(X_{i^*,j} = 0)) \quad (\text{D.4})$$

$$\mathbb{V}(Q_{i^*})^{1/2} = \left[\frac{1}{N^2} \sum_{j=0}^N \{ (1 - P(X_{i^*,j} = 0) - [1 - P(X_{i^*,j} = 0)]^2 \} \right]^{1/2} \quad (\text{D.5})$$

where $\mathbb{V}()$ denotes the variance and

$$P(X_{i^*,j} = 0) = (1 - \lambda_j)^{4N_b 2^{i^*}}. \quad (\text{D.6})$$

Note that the standard deviation of Q_{i^*} scales as $1/N$ due to uncertainty arising from sampling noise in λ_j which diminishes as more samples from the approximate posterior distribution are obtained. Also note that the number of bases used in Eq.(D.6) is $N_b = 298$, as this is the number of pathological bases in the set of bases considered. An empirical estimate of the standard error may be obtained by treating each individual as a Bernoulli trial with probability $p = P(X_{i^*} \geq 1)$, and SEM $\sqrt{p(1-p)/N}$. These are compatible with our uncertainty in the approximate posterior mean of the indicator function, shown in Eq.(D.5).

D.2 DETAILS FOR THE VISUALISATION OF MUTATED REGIONS

The visual representation in Fig. 5.4 was created by taking the mean of the approximate posterior samples for the mutation rate $\mathbb{E}(\lambda_m|\mathcal{D})$ as the mutation rate of the individual. The number of mutated neuronal regions for $20 \leq i^* \leq 27$ were drawn using Eq.(5.1), where $X_{i^*=20}$ was redrawn until $X_{i^*=20} \geq 1$. We think this is fair for illustrative purposes since Fig. 5.3E shows that $\sim 79\%$ of all simulated individuals possess at least one region seeded at generation $i^* = 20$. Each circular mutated region has an area proportional to the number of pathologically mutated neurons in the corresponding region: $k\pi r^2 = 2^{36} \cdot 2^{-(i^*+2)}$ where k is chosen arbitrarily. The radius and angle of each region relative to the centre of the whole brain was chosen via a uniform random distribution. Although the relative sizes of each region are to scale, the area of the whole brain is certainly not to scale with the mutated regions.

D.3 MULTIMODALITY IN THE NUMBER OF PATHOLOGICAL NEURONAL MUTATIONS IS DRIVEN BY THE LARGEST PATHOLOGICAL REGION

In Fig. 5.3D, the distribution for the total number of mutations per individual shows multimodality. In Fig. D.1 we show that individuals with the largest numbers of mutated neurons in total, also had the largest single mutated regions. Let us modify our notation for the fraction of the brain affected by a mutation (f , see Eq.(5.2)), to be indexed by both the mutant generation number i , and the individual j (f_{ij}). Then overlaying a scatter of $\max_i(f_{ij})$ against the total number of mutations per individual j , on top of a histogram of the total number of mutations, reveals that the modes are driven by individuals with the largest value of $\max_i f_{ij}$ – see Fig. D.1.

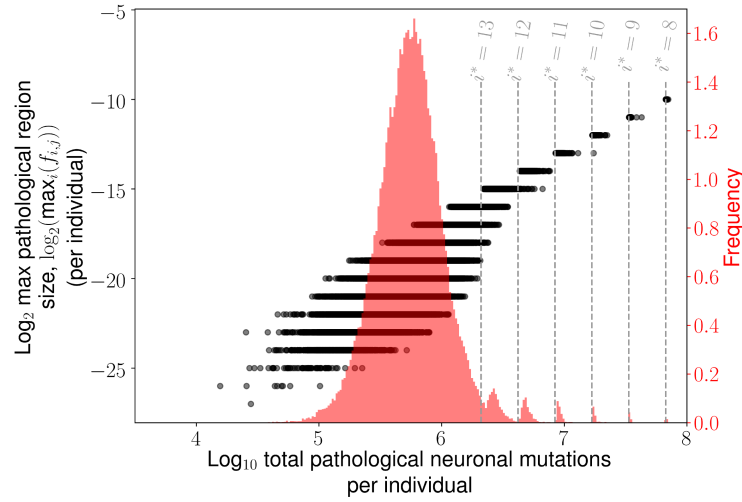


Figure D.1: Multimodality in the distribution of total pathological mutations per individual is driven by the largest pathological region per individual. Correlation between the largest pathological region over all mutations i , for individuals j , and the total number of pathological mutations per individual (grey point for each individual), overlaid with the distribution of total pathological mutations per individual. The number of cells corresponding to regions seeded at generation $i^* = 8, 9, \dots, 13$ are shown in grey. At fixed i^* , e.g. $i^* = 11$, for individuals whose largest mutant region occurred at $i^* = 12$ (corresponding to $\log_2(f) = 13$), the width of the corresponding mode of the distribution of total pathological mutations is small relative to the number of mutant cells associated with e.g. mutations seeded at $i^* = 12$. Hence, individuals with a mutation seeded at $i^* = 12$ have sufficiently more pathologically mutated cells to induce a separate mode.

D.4 CONSTANT REGION SIZE MODEL

D.4.1 SUMMARY

In this section, we consider an alternative model where the brain is modelled as consisting of independent regions of some characteristic size f , which are homogeneously mutated at the

mean variant allele fraction (VAF) experimentally observed in the case genes (132617 bp) in patients demonstrating neurodegenerative disease ($n = 40$). $f = 1$ implies that if the brain is mutated at a particular base pair then this mutation is present across the entire brain. However, if $f < 1$ then there exists $1/f$ regions in the brain, each consisting of Nf cells, where N is the number of cells in the brain. We assume that each region has a probability p_m of being homogeneously mutated at a particular base pair in the case genes (Fig. D.2A). This model does not make a commitment to a tree-like development instead assuming that mutations occupy characteristic volumes of tissue. This makes the model, in some sense, less realistic but it has the virtue of being a different perspective that allows us to probe the robustness of our observations. This model also lends itself more naturally to the diffusive nature of neurodevelopment (when $f < 1$), unlike the branching process model.

The parameters f and p_m were considered as shared amongst all patients, for parsimony. We investigated the effect of varying f by orders of magnitude through approximate Bayesian computation (see below). We again restricted our attention to patients with neurodegenerative disease to avoid sampling bias in our extrapolations. We found that, under this model, the relative size of independent regions was not strongly constrained by the data (Fig. D.2C), finding only that large region sizes $\gtrsim 10\%$ are unlikely. (The tree-like structure considered in Chapter 5 puts a much stronger constraint on the distribution of sizes of patches of mutated cells.)

The data did constrain the mean number of mutated regions per individual such that the lower bound was approximately 1 region per individual ($0.942 \pm 6.6 \times 10^{-3}$ mutated regions per individual, 5th percentile of posterior distribution, see Fig. D.2D). Consequently, we found that 63% of simulated individuals carried non-zero levels of pathologically mutated cells in case genes, carrying between $0.33\text{--}5.0 \times 10^5$ pathologically mutated cells per affected individual (95% Bayesian credible interval, see Fig. D.2E). This number of cells with pathological mutations is in broad agreement with Fig. 5.3D. Thus, even this remarkably crude model, that makes no commitments to a developmental hierarchy, suggests that zones of pathological mutations occur in the brains of most people. That we observe a subset of simulated individuals lacking mutated regions (and therefore pathologically mutated cells) is due to our broad priors in f and p_m allowing for independent brain regions to be large enough that it is unlikely for any pathological regions to arise in an individual (this limit is $f \gtrsim \tilde{p}p_m$ where $\tilde{p} = 298/132617$ is the probability of a region being pathologically mutated, given that it is mutated, see below – Fig. D.2G).

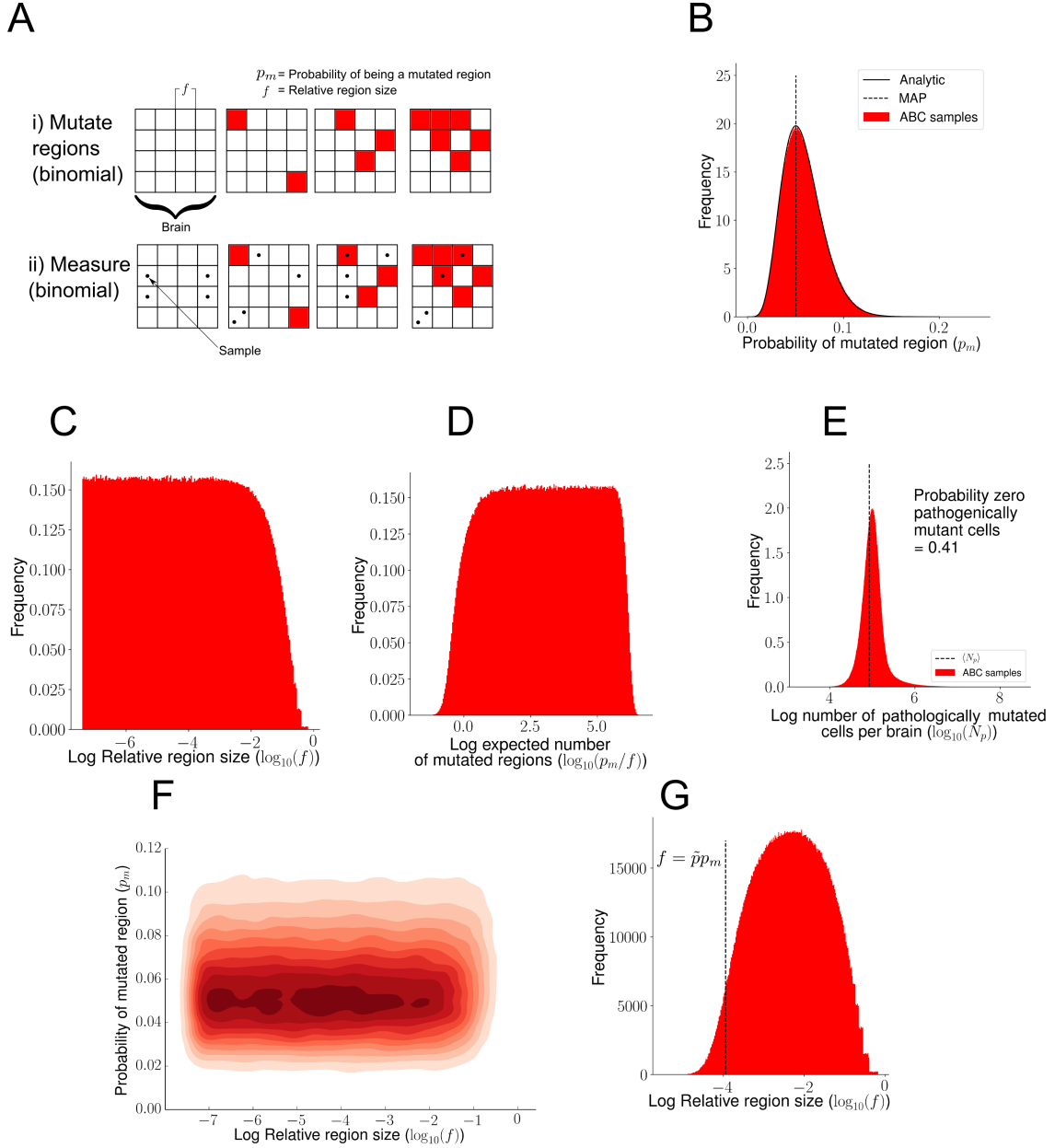


Figure D.2: Constant region size model of neurodevelopment. (A) We use a simplified model of mutagenesis in the brain, where the brain consists of independent ‘regions’ of relative size f , where f is an unknown constant. Each region (small square) has an independent probability of being homogeneously mutated in any of the case genes (p_m , red square), with the number of mutated regions being binomially distributed. We model the measurement process each sample (black circle) being an independent trial, with the number of samples being positive for a mutation as binomially distributed. (B) The marginal posterior for p_m from ABC overlaps with the analytic posterior distribution if p_m is treated independently of f . $(p_m)_{\text{MAP}} = 6119$. (C) The marginal posterior distribution of f is not strongly constrained by the data, although support reduces markedly as $f \rightarrow 1$. (D) The mean number of mutated regions per individual is constrained strongly enough such that the lower bound is approximately 1. (E) Using the probability that a region is pathologically mutated ($\tilde{p} = 298/132617$) 37% of simulated individuals had zero pathologically mutated cells but those that did tended to have $\sim 10^5$ mutated neurons. (F) Joint distribution of p_m and $\log_{10}(f)$ shows that p_m and f are well-approximated as being independent. (G) For individuals with 0 pathologically mutated cells (see E), we may plot the posterior distribution of relative region sizes. The support decays when $f \ll \tilde{p}p_m$. This is the limit that the mean number of pathologically mutated regions per individual is much less than 1, see Eq.(D.13).

D.4.2 APPROXIMATE BAYESIAN COMPUTATION FOR FIXED REGION SIZE MODEL

We consider the brain as consisting of ‘regions’, each having a relative mass (f) where $0 < f \leq 1$, implying that each brain contains $1/f$ such regions. For the subset of the data we considered (the case genes in the case cohort), all samples measuring positive for somatic mutations corresponded to unique genes in unique patients. Consequently, we used the parsimonious model that positive samples were homogeneously mutated in a particular case gene at the average VAF (v) determined experimentally ($v = 0.87\%$). We modelled each region as having an independent probability p_m of being homogeneously mutated at a particular base of the case genes.

Defining Y_i as the number of mutated regions in the brain of a single individual (i), then Y_i follows a binomial distribution

$$Y_i \sim \text{Binom}\left(\frac{1}{f}, p_m\right) \quad (\text{D.7})$$

where $\text{Binom}(n, p)$ is a binomial trial with n attempts each with probability p of success. Note that $1/f$ is constrained to be an integer, and may be obtained by rounding $1/f$ for a given f . We then model the measurement process for each individual as consisting of n_i samples. Since there are Y_i patches, each of which occupies a fraction f of the brain, then fY_i gives the total fraction of the brain which is mutated, where $0 \leq fY_i \leq 1$. This results in M_i samples positive for mutation for individual i , with

$$M_i \sim \text{Binom}(n_i, fY_i). \quad (\text{D.8})$$

Note that we have modelled all individuals as sharing the only two free parameters in the model: f and p_m .

To infer these parameters, we used an approximate Bayesian computation (ABC) rejection algorithm. We used the uninformative priors

$$\log(f) \sim \text{unif}\left(\log\left(\frac{3415}{86 \times 10^9}\right), 0\right) \quad (\text{D.9})$$

$$p_m \sim \text{unif}(0, 1) \quad (\text{D.10})$$

where we chose to use a prior for f which was uniform over log-space, since we were uncertain of its value over orders of magnitude. The lower limit of the prior for f was chosen such that f could not be smaller than the approximate relative size occupied by a single sample i.e. 3415 cells out of the total number of cells in the brain. We have idealised the brain as consisting

purely of neurons, of which there are approximately 86×10^9 per brain [323].

We simulated each patient by drawing (f, p_m) from the prior, and determining M_i with Eq.(D.8). We then used the following summary statistics for the simulated model ($\mathbf{S}$)

$$\mathbf{S} = \left(\sum_{i=1}^{40} M_i, \sum_{i=1}^{40} I(M_i) \right) \quad (\text{D.11})$$

where I is the indicator function defined in Eq.(5.11), and the upper summation limit corresponds to the number of case patients. The first element counts the total number of positive samples across all patients, whereas the second element counts the total number of affected patients. The summary statistic for the data is $\hat{\mathbf{S}} = (6, 6)$. We used the distance metric

$$\rho(\mathbf{S}, \hat{\mathbf{S}}) = \sum |\mathbf{S} - \hat{\mathbf{S}}| \leq \epsilon \quad (\text{D.12})$$

and required the tolerance $\epsilon = 0$ to accept (f, p_m) as a draw from the approximate posterior $P(f, p_m | \rho(\mathbf{S}, \hat{\mathbf{S}}(\mathcal{D})) = 0) \approx P(f, p_m | \mathcal{D})$ for data $\mathcal{D}$. This therefore enforces the ABC rejection algorithm to only accept samples where every mutation occurs once in each affected patient (as observed in the data), and that the correct number of mutant samples and patients are observed overall.

Given posterior samples of (f, p_m) , we may also infer the distribution of the number of pathologically mutated neurons per brain. Dropping the patient subscript i , and using Eq.(D.7), the number of pathologically mutated regions per patient ($\tilde{Y}$) is

$$\tilde{Y} \sim \text{Binom}(Y, \tilde{p}) \quad (\text{D.13})$$

where $\tilde{p}$ is the probability of a mutated region being pathological, which we take as $298/132617$, being the proportion of base pairs in the case genes being associated with pathology. From this, we approximate the number of pathologically mutated cells (N_p) as

$$N_p \approx \tilde{Y} \cdot (Nf) \cdot v \quad (\text{D.14})$$

where N = number of neurons per brain and v = is the proportion of mutant cells in a positive sample, which we take as the average VAF $v = 0.87 \times 10^{-2}$. The term (Nf) is the number of neurons per region.

It is instructive to observe that the mean number of mutant cells, $\langle N_p \rangle$, for fixed f and p_m

is $\langle N_p \rangle = Nfv\langle \tilde{Y} \rangle = Nfv\tilde{p}\langle Y \rangle$. Using Eq.(D.7), $\langle N_p \rangle = Nfv\tilde{p}p_m/f = Nv\tilde{p}p_m$ which is independent of f . Note that if $f \ll \tilde{p}p_m$ then $\langle \tilde{Y} \rangle \gg 1$, in which case an individual is likely to harbour at least one pathological mutation (see Fig. D.2G). Substituting $N = 86 \times 10^9$ neurons, $v = 0.87 \times 10^{-2}$, and the maximum *a posteriori* estimate of $(p_m)_{MAP} = 6/119$, yields $\langle N_p \rangle = 1.70 \times 10^5$ pathologically mutated neurons per brain, which approximately corresponds to the peak shown in Fig. D.2E.

REFERENCES

- [1] Berg, J., Tymoczko, J., and Stryer, L. *Biochemistry, Fifth Edition*. W. H. Freeman, (2002).
- [2] Schon, E. A., DiMauro, S., and Hirano, M. Human mitochondrial DNA: roles of inherited and somatic mutations. *Nat. Rev. Genet.* **13**(12), 878–890 (2012).
- [3] Lin, M. T. and Beal, M. F. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* **443**(7113), 787–795 (2006).
- [4] Keogh, M. J. and Chinnery, P. F. Mitochondrial DNA mutations in neurodegeneration. *Biochim. Biophys. Acta.* **1847**(11), 1401–1411 (2015).
- [5] Wallace, D. C. Mitochondria and cancer. *Nature Rev. Cancer* **12**(10), 685–698 (2012).
- [6] Yuan, Y., Ju, Y. S., Kim, Y., Li, J., Wang, Y., Yang, Y., Martincorena, I., Creighton, C., Weinstein, J. N., Xu, Y., et al. Comprehensive molecular characterization of mitochondrial genomes in human cancers. *bioRxiv* , 161356 (2017).
- [7] Kauppila, T. E., Kauppila, J. H., and Larsson, N.-G. Mammalian mitochondria and aging: an update. *Cell Metab.* **25**(1), 57–71 (2017).
- [8] López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., and Kroemer, G. The hallmarks of aging. *Cell* **153**(6), 1194–1217 (2013).
- [9] Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., et al. Sequence and organization of the human mitochondrial genome. *Nature* **290**(5806), 457 (1981).
- [10] Hoitzing, H., Johnston, I. G., and Jones, N. S. What is the function of mitochondrial networks? A theoretical assessment of hypotheses and proposal for future research. *BioEssays* **37**(6), 687–700 (2015).
- [11] Vassar, R., Ngai, J., and Axel, R. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* **74**(2), 309–318 (1993).
- [12] Chang, H. H., Hemberg, M., Barahona, M., Ingber, D. E., and Huang, S. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**(7194), 544–547 (2008).
- [13] Snijder, B., Sacher, R., Rämö, P., Damm, E.-M., Liberali, P., and Pelkmans, L. Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* **461**(7263), 520 (2009).

- [14] Spencer, S. L., Gaudet, S., Albeck, J. G., Burke, J. M., and Sorger, P. K. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459**(7245), 428 (2009).
- [15] Márquez-Jurado, S., Díaz-Colunga, J., Neves, R. P., Martinez-Lorente, A., Almazán, F., Guantes, R., and Iborra, F. J. Mitochondrial levels determine variability in cell death by modulating apoptotic gene expression. *Nat. Commun.* **9**(1), 389 (2018).
- [16] Bahar, R., Hartmann, C. H., Rodriguez, K. A., Denny, A. D., Busuttil, R. A., Dollé, M. E., Calder, R. B., Chisholm, G. B., Pollock, B. H., Klein, C. A., et al. Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441**(7096), 1011 (2006).
- [17] Swain, P. S., Elowitz, M. B., and Siggia, E. D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA* **99**(20), 12795–12800 (2002).
- [18] Elowitz, M. B., Levine, A. J., Siggia, E. D., and Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**(5584), 1183–1186 (2002).
- [19] Newman, J. R., Ghaemmaghami, S., Ihmels, J., Breslow, D. K., Noble, M., DeRisi, J. L., and Weissman, J. S. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**(7095), 840–846 (2006).
- [20] Huh, D. and Paulsson, J. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.* **43**(2), 95 (2011).
- [21] Huh, D. and Paulsson, J. Random partitioning of molecules at cell division. *Proc. Natl. Acad. Sci. USA* , 201013171 (2011).
- [22] Johnston, I. G. and Jones, N. S. Closed-form stochastic solutions for non-equilibrium dynamics and inheritance of cellular components over many cell divisions. In *Proc. R. Soc. A*, volume 471, 20150050. The Royal Society, (2015).
- [23] Das Neves, R. P., Jones, N. S., Andreu, L., Gupta, R., Enver, T., and Iborra, F. J. Connecting variability in global transcription rate to mitochondrial variability. *PLoS Biol.* **8**(12), e1000560 (2010).
- [24] Johnston, I. G., Gaal, B., Das Neves, R. P., Enver, T., Iborra, F. J., and Jones, N. S. Mitochondrial variability as a source of extrinsic cellular noise. *PLoS Comput. Biol.* **8**(3), 1–14 (2012).
- [25] Guantes, R., Rastrojo, A., das Neves, R. P., Lima, A., Aguado, B., and Iborra, F. J. Global variability in gene expression and alternative splicing is modulated by mitochondrial content. *Genome Res.* **25**(5), 633–644 (2015).
- [26] Raj, A. and van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**(2), 216–226 (2008).
- [27] Gaál, B. *Mitochondrial and Transcription Rate Heterogeneity of Mouse Embryonic Stem Cells*. PhD thesis, University of Oxford, (2014).

- [28] Sukumar, M., Liu, J., Mehta, G. U., Patel, S. J., Roychoudhuri, R., Crompton, J. G., Klebanoff, C. A., Ji, Y., Li, P., Yu, Z., et al. Mitochondrial membrane potential identifies cells with enhanced stemness for cellular therapy. *Cell Metab.* **23**(1), 63–76 (2016).
- [29] Ansó, E., Weinberg, S. E., Diebold, L. P., Thompson, B. J., Malinge, S., Schumacker, P. T., Liu, X., Zhang, Y., Shao, Z., Steadman, M., et al. The mitochondrial respiratory chain is essential for haematopoietic stem cell function. *Nat. Cell Biol.* **19**(6), 614 (2017).
- [30] Mizutani, S., Miyato, Y., Shidara, Y., Asoh, S., Tokunaga, A., Tajiri, T., and Ohta, S. Mutations in the mitochondrial genome confer resistance of cancer cells to anticancer drugs. *Cancer Sci.* **100**(9), 1680–1687 (2009).
- [31] Yao, Z., Jones, A., Fassone, E., Sweeney, M., Lebiedzinska, M., Suski, J., Wieckowski, M., Tajeddine, N., Hargreaves, I., Yasukawa, T., et al. PGC-1 β mediates adaptive chemoresistance associated with mitochondrial DNA mutations. *Oncogene* **32**(20), 2592 (2013).
- [32] Dhar, R., Missarova, A., Lehner, B., and Carey, L. B. Single cell functional genomics reveals the importance of mitochondria in cell-to-cell variation in proliferation, drug resistance and mutation outcome. *bioRxiv*, 346361 (2018).
- [33] Kiviet, D. J., Nghe, P., Walker, N., Boulineau, S., Sunderlikova, V., and Tans, S. J. Stochasticity of metabolism and growth at the single-cell level. *Nature* **514**(7522), 376–379 (2014).
- [34] Takhaveev, V. and Heinemann, M. Metabolic heterogeneity in clonal microbial populations. *Curr. Opin. Microbiol.* **45**, 30–38 (2018).
- [35] Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J., and Letellier, T. Mitochondrial threshold effects. *Biochem. J.* **370**(Pt 3), 751–762 (2003).
- [36] Rocher, C., Taanman, J.-W., Pierron, D., Faustin, B., Benard, G., Rossignol, R., Malgat, M., Pedespan, L., and Letellier, T. Influence of mitochondrial DNA level on cellular energy metabolism: implications for mitochondrial diseases. *J. Bioenerg. Biomembr.* **40**, 59–67 (2008).
- [37] Picard, M., Zhang, J., Hancock, S., Derbeneva, O., Golhar, R., Golik, P., O’Hearn, S., Levy, S., Potluri, P., Lvova, M., et al. Progressive increase in mtDNA 3243A>G heteroplasmy causes abrupt transcriptional reprogramming. *Proc. Natl. Acad. Sci. USA* **111**(38), E4033–E4042 (2014).
- [38] Palikaras, K. and Tavernarakis, N. Mitochondrial homeostasis: the interplay between mitophagy and mitochondrial biogenesis. *Exp. Gerontol.* **56**, 182–188 (2014).
- [39] Birky, C. W., Maruyama, T., and Fuerst, P. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* **103**(3), 513–527 (1983).
- [40] Chinnery, P. F. and Samuels, D. C. Relaxed replication of mtDNA: a model with implications for the expression of disease. *Am. J. Hum. Genet.* **64**, 1158 (1999).

- [41] Capps, G. J., Samuels, D. C., and Chinnery, P. F. A model of the nuclear control of mitochondrial DNA replication. *J. Theor. Biol.* **221**(4), 565–583 (2003).
- [42] Wonnapijit, P., Chinnery, P., and Samuels, D. The distribution of mitochondrial DNA heteroplasmy due to random genetic drift. *Am. J. Hum. Genet.* **83**, 582 (2008).
- [43] Johnston, I. G. and Jones, N. S. Evolution of cell-to-cell variability in stochastic, controlled, heteroplasmic mtDNA populations. *Am. J. Hum. Genet.* **99**(5), 1150–1162 (2016).
- [44] Posakony, J. W., England, J. M., and Attardi, G. Mitochondrial growth and division during the cell cycle in HeLa cells. *J. Cell Biol.* **74**(2), 468–491 (1977).
- [45] Rafelski, S. M., Viana, M. P., Zhang, Y., Chan, Y.-H. M., Thorn, K. S., Yam, P., Fung, J. C., Li, H., Costa, L. d. F., and Marshall, W. F. Mitochondrial network size scaling in budding yeast. *Science* **338**(6108), 822–824 (2012).
- [46] Iborra, F. J., Kimura, H., and Cook, P. R. The functional organization of mitochondrial genomes in human cells. *BMC Biol.* **2**(1), 9 (2004).
- [47] Tauber, J., Dlasková, A., Šantorová, J., Smolková, K., Alán, L., Špaček, T., Plecítá-Hlavatá, L., Jaburek, M., and Jezek, P. Distribution of mitochondrial nucleoids upon mitochondrial network fragmentation and network reintegration in HEPG2 cells. *Int. J. Biochem. Cell Biol.* **45**(3), 593–603 (2013).
- [48] Osman, C., Noriega, T. R., Okreglak, V., Fung, J. C., and Walter, P. Integrity of the yeast mitochondrial genome, but not its distribution and inheritance, relies on mitochondrial fission and fusion. *Proc. Natl. Acad. Sci. USA* **112**(9), E947–E956 (2015).
- [49] Jajoo, R., Jung, Y., Huh, D., Viana, M. P., Rafelski, S. M., Springer, M., and Paulsson, J. Accurate concentration control of mitochondria and nucleoids. *Science* **351**(6269), 169–172 (2016).
- [50] Miettinen, T. P. and Björklund, M. Cellular allometry of mitochondrial functionality establishes the optimal cell size. *Dev. Cell* **39**(3), 370–382 (2016).
- [51] El-Hattab, A. W. and Scaglia, F. Mitochondrial DNA depletion syndromes: review and updates of genetic basis, manifestations, and therapeutic options. *Neurotherapeutics* **10**(2), 186–198 (2013).
- [52] Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J. N., Rovio, A. T., Bruder, C. E., Bohlooly-Y, M., Gidlöf, S., Oldfors, A., Wibom, R., et al. Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**(6990), 417 (2004).
- [53] Jiang, M., Kauppila, T. E. S., Motori, E., Li, X., Atanassov, I., Folz-Donahue, K., Bonekamp, N. A., Albarran-Gutierrez, S., Stewart, J. B., and Larsson, N.-G. Increased total mtDNA copy number cures male infertility despite unaltered mtDNA mutation load. *Cell Metab.* **26**(2), 429–436 (2017).

- [54] Williams, S. L., Mash, D. C., Züchner, S., and Moraes, C. T. Somatic mtDNA mutation spectra in the aging human putamen. *PLoS Genet.* **9**(12), e1003990 (2013).
- [55] Kennedy, S. R., Salk, J. J., Schmitt, M. W., and Loeb, L. A. Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage. *PLoS Genet.* **9**(9), e1003794 (2013).
- [56] Stewart, J. B. and Larsson, N.-G. Keeping mtDNA in shape between generations. *PLoS Genet.* **10**(10), e1004670 (2014).
- [57] Kauppi, J. H. and Stewart, J. B. Mitochondrial DNA: radically free of free-radical driven mutations. *Biochim. Biophys. Acta.* **1847**(11), 1354–1361 (2015).
- [58] Kauppi, J. H., Bonekamp, N. A., Mourier, A., Isokallio, M. A., Just, A., Kauppi, T. E., Stewart, J. B., and Larsson, N.-G. Base-excision repair deficiency alone or combined with increased oxidative stress does not increase mtDNA point mutations in mice. *Nucleic Acids Res.* **46**, 6642–6669 (2018).
- [59] Phillips, A. F., Millet, A. R., Tigano, M., Dubois, S. M., Crimmins, H., Babin, L., Charpentier, M., Piganeau, M., Brunet, E., and Sfeir, A. Single-molecule analysis of mtDNA replication uncovers the basis of the common deletion. *Mol. Cell* **65**(3), 527–538 (2017).
- [60] Eigen, M. and Schuster, P. A principle of natural self-organization. *Naturwissenschaften* **64**(11), 541–565 (1977).
- [61] Nowak, M. A. *Evolutionary dynamics*. Harvard University Press, (2006).
- [62] Jayaprakash, A. D., Benson, E. K., Gone, S., Liang, R., Shim, J., Lambertini, L., Toloue, M. M., Wigler, M., Aaronson, S. A., and Sachidanandam, R. Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. *Nucleic Acids Res.* **43**(4), 2177–2187 (2015).
- [63] Morris, J., Na, Y.-J., Zhu, H., Lee, J.-H., Giang, H., Ulyanova, A. V., Baltuch, G. H., Brem, S., Chen, H. I., Kung, D. K., et al. Pervasive within-mitochondrion single-nucleotide variant heteroplasmy as revealed by single-mitochondrion sequencing. *Cell Rep.* **21**(10), 2706–2713 (2017).
- [64] Satoh, M. and Kuroiwa, T. Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. *Exp. Cell Res.* **196**(1), 137–140 (1991).
- [65] Burgstaller, J. P., Johnston, I. G., Jones, N. S., Albrechtová, J., Kolbe, T., Vogl, C., Futschik, A., Mayrhofer, C., Klein, D., Sabitzer, S., et al. MtDNA segregation in heteroplasmic tissues is common in vivo and modulated by haplotype differences and developmental stage. *Cell Rep.* **7**(6), 2031–2041 (2014).
- [66] Acton, B., Lai, I., Shang, X., Jurisicova, A., and Casper, R. Neutral mitochondrial heteroplasmy alters physiological function in mice. *Biol. Reprod.* **77**(3), 569–576 (2007).

- [67] Sharpley, M. S., Marciniak, C., Eckel-Mahan, K., McManus, M., Crimi, M., Waymire, K., Lin, C. S., Masubuchi, S., Friend, N., Koike, M., et al. Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. *Cell* **151**(2), 333–343 (2012).
- [68] Hoitzing, H., Gammage, P. A., Minczuk, M., Johnston, I. G., and Jones, N. S. Energetic costs of cellular and therapeutic control of stochastic mtDNA populations. *bioRxiv*, 145292 (2017).
- [69] Bagwan, N., Bonzon-Kulichenko, E., Calvo, E., Lechuga-Vieco, A. V., Michalakopoulos, S., Trevisan-Herraz, M., Ezkurdia, I., Rodríguez, J. M., Magni, R., Latorre-Pellicer, A., et al. Comprehensive quantification of the modified proteome reveals oxidative heart damage in mitochondrial heteroplasmy. *Cell Rep.* **23**(12), 3685–3697 (2018).
- [70] Ewens, W. J. *Mathematical population genetics: I. theoretical introduction*, volume 27. Springer Science & Business Media, (2004).
- [71] Poovathingal, S. K., Gruber, J., Halliwell, B., and Gunawan, R. Stochastic drift in mitochondrial DNA point mutations: a novel perspective ex silico. *PLoS Comput. Biol.* **5**(11), e1000572 (2009).
- [72] Burgstaller, J. P., Kolbe, T., Havlicek, V., Hembach, S., Poulton, J., Piálek, J., Steinborn, R., Rülcke, T., Brem, G., Jones, N. S., and Johnston, I. G. Large-scale genetic analysis reveals mammalian mtDNA heteroplasmy dynamics and variance increase through lifetimes and generations. *Nat. Commun.* **9**(1), 2488 (2018).
- [73] Johnston, I. G., Burgstaller, J. P., Havlicek, V., Kolbe, T., Rülcke, T., Brem, G., Poulton, J., and Jones, N. S. Stochastic modelling, bayesian inference, and new *in vivo* measurements elucidate the debated mtDNA bottleneck mechanism. *eLife* **4**, e07464 (2015).
- [74] Burr, S. P., Pezet, M., and Chinnery, P. F. Mitochondrial DNA heteroplasmy and purifying selection in the mammalian female germ line. *Dev. Growth Differ.* **60**(1), 21–32 (2018).
- [75] Floros, V. I., Pyle, A., Dietmann, S., Wei, W., Tang, W. W., Irie, N., Payne, B., Capalbo, A., Noli, L., Coxhead, J., et al. Segregation of mitochondrial DNA heteroplasmy through a developmental genetic bottleneck in human embryos. *Nat. Cell Biol.* **20**(2), 144 (2018).
- [76] Jenuth, J. P., Peterson, A. C., Fu, K., and Shoubridge, E. A. Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. *Nat. Genet.* **14**(2), 146 (1996).
- [77] Cao, L., Shitara, H., Horii, T., Nagao, Y., Imai, H., Abe, K., Hara, T., Hayashi, J.-I., and Yonekawa, H. The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. *Nat. Genet.* **39**(3), 386 (2007).
- [78] Cree, L. M., Samuels, D. C., de Sousa Lopes, S. C., Rajasimha, H. K., Wonnapijit, P., Mann, J. R., Dahl, H.-H. M., and Chinnery, P. F. A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat. Genet.* **40**(2), 249 (2008).
- [79] Wai, T., Teoli, D., and Shoubridge, E. A. The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nat. Genet.* **40**(12), 1484 (2008).

- [80] Hoitzing, H., Johnston, I. G., and Jones, N. S. Stochastic models for evolving cellular populations of mitochondria: Disease, development, and ageing. In *Stochastic Processes, Multiscale Modeling, and Numerical Methods for Computational Cellular Biology*, 287–314. Springer (2017).
- [81] Ross, J. M., Stewart, J. B., Hagström, E., Brené, S., Mourier, A., Coppotelli, G., Freyer, C., Lagouge, M., Hoffer, B. J., Olson, L., et al. Germline mitochondrial DNA mutations aggravate ageing and can impair brain development. *Nature* **501**(7467), 412 (2013).
- [82] Khrapko, K., Bodyak, N., Thilly, W. G., Van Orsouw, N. J., Zhang, X., Coller, H. A., Perls, T. T., Upton, M., Vijg, J., and Wei, J. Y. Cell-by-cell scanning of whole mitochondrial genomes in aged human heart reveals a significant fraction of myocytes with clonally expanded deletions. *Nucleic Acids Res.* **27**(11), 2434–2441 (1999).
- [83] Taylor, R. W., Barron, M. J., Borthwick, G. M., Gospel, A., Chinnery, P. F., Samuels, D. C., Taylor, G. A., Plusa, S. M., Needham, S. J., Greaves, L. C., et al. Mitochondrial DNA mutations in human colonic crypt stem cells. *J. Clin. Invest.* **112**(9), 1351–1360 (2003).
- [84] Kraytsberg, Y., Kudryavtseva, E., McKee, A. C., Geula, C., Kowall, N. W., and Khrapko, K. Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat. Genet.* **38**(5), 518 (2006).
- [85] Bender, A., Krishnan, K. J., Morris, C. M., Taylor, G. A., Reeve, A. K., Perry, R. H., Jaros, E., Hersheson, J. S., Betts, J., Klopstock, T., et al. High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and parkinson disease. *Nat. Genet.* **38**(5), 515 (2006).
- [86] Li, M., Schröder, R., Ni, S., Madea, B., and Stoneking, M. Extensive tissue-related and allele-related mtDNA heteroplasmy suggests positive selection for somatic mutations. *Proc. Natl. Acad. Sci. USA* **112**(8), 2491–2496 (2015).
- [87] Aiken, J., Bua, E., Cao, Z., Lopez, M., Wanagat, J., McKenzie, D., and McKiernan, S. Mitochondrial DNA deletion mutations and sarcopenia. *Ann. N. Y. Acad. Sci.* **959**(1), 412–423 (2002).
- [88] Bua, E., Johnson, J., Herbst, A., Delong, B., McKenzie, D., Salamat, S., and Aiken, J. M. Mitochondrial DNA-deletion mutations accumulate intracellularly to detrimental levels in aged human skeletal muscle fibers. *Am. J. Hum. Genet.* **79**(3), 469–480 (2006).
- [89] Vincent, A. E., Rosa, H. S., Pabis, K., Lawless, C., Chen, C., Grünwald, A., Rygiel, K. A., Rocha, M. C., Reeve, A. K., Falkous, G., et al. Sub-cellular origin of mtDNA deletions in human skeletal muscle. *Ann. Neurol.* (2018).
- [90] Wanagat, J., Cao, Z., Pathare, P., and Aiken, J. M. Mitochondrial DNA deletion mutations colocalize with segmental electron transport system abnormalities, muscle fiber atrophy, fiber splitting, and oxidative damage in sarcopenia. *FASEB J.* **15**(2), 322–332 (2001).
- [91] Khrapko, K. and Vijg, J. Mitochondrial DNA mutations and aging: devils in the details? *Trends Genet.* **25**(2), 91–98 (2009).

- [92] Vermulst, M., Bielas, J. H., Kujoth, G. C., Ladiges, W. C., Rabinovitch, P. S., Prolla, T. A., and Loeb, L. A. Mitochondrial point mutations do not limit the natural lifespan of mice. *Nat. Genet.* **39**(4), 540 (2007).
- [93] Lakshmanan, L. N., Yee, Z., Ng, L. F., Gunawan, R., Halliwell, B., and Gruber, J. Clonal expansion of mitochondrial DNA deletions is a private mechanism of ageing in long-lived animals. *Aging Cell* **17**(5), e12814 (2018).
- [94] Zhao, Q., Wang, J., Levichkin, I. V., Stasinopoulos, S., Ryan, M. T., and Hoogenraad, N. J. A mitochondrial specific stress response in mammalian cells. *EMBO J.* **21**(17), 4411–4419 (2002).
- [95] Durieux, J., Wolff, S., and Dillin, A. The cell-non-autonomous nature of electron transport chain-mediated longevity. *Cell* **144**(1), 79–91 (2011).
- [96] Zhang, Q., Wu, X., Chen, P., Liu, L., Xin, N., Tian, Y., and Dillin, A. The mitochondrial unfolded protein response is mediated cell-non-autonomously by retromer-dependent Wnt signaling. *Cell* **174**(4), 870–883 (2018).
- [97] Tynismaa, H., Carroll, C. J., Raimundo, N., Ahola-Erkkilä, S., Wenz, T., Ruhanen, H., Guse, K., Hemminki, A., Peltola-Mjøsund, K. E., Tulkki, V., et al. Mitochondrial myopathy induces a starvation-like response. *Hum. Mol. Genet.* **19**(20), 3948–3958 (2010).
- [98] Khan, N. A., Nikkanen, J., Yatsuga, S., Jackson, C., Wang, L., Pradhan, S., Kivelä, R., Pessia, A., Velagapudi, V., and Suomalainen, A. mTORC1 regulates mitochondrial integrated stress response and mitochondrial myopathy progression. *Cell Metab.* **26**(2), 419–428 (2017).
- [99] Simard, M.-L., Mourier, A., Greaves, L. C., Taylor, R. W., and Stewart, J. B. A novel histochemistry assay to assess and quantify focal cytochrome *c* oxidase deficiency. *J. Pathol.* (2018).
- [100] Vincent, A. E. and Picard, M. Multi-level heterogeneity of mitochondrial respiratory chain deficiency. *J. Pathol.* (2018).
- [101] Spees, J. L., Olson, S. D., Whitney, M. J., and Prockop, D. J. Mitochondrial transfer between cells can rescue aerobic respiration. *Proc. Natl. Acad. Sci. USA* **103**(5), 1283–1288 (2006).
- [102] Torralba, D., Baixauli, F., and Sánchez-Madrid, F. Mitochondria know no boundaries: mechanisms and functions of intercellular mitochondrial transfer. *Front. Cell Dev. Biol.* **4**, 107 (2016).
- [103] Rustom, A., Saffrich, R., Markovic, I., Walther, P., and Gerdes, H.-H. Nanotubular highways for intercellular organelle transport. *Science* **303**(5660), 1007–1010 (2004).
- [104] Berridge, M. V., McConnell, M. J., Grasso, C., Bajzikova, M., Kovarova, J., and Neuzil, J. Horizontal transfer of mitochondria between mammalian cells: beyond co-culture approaches. *Curr. Opin. Genetics Dev.* **38**, 75–82 (2016).
- [105] Islam, M. N., Das, S. R., Emin, M. T., Wei, M., Sun, L., Westphalen, K., Rowlands, D. J., Quadri, S. K., Bhattacharya, S., and Bhattacharya, J. Mitochondrial transfer from bone-marrow-derived stromal cells to pulmonary alveoli protects against acute lung injury. *Nat. Med.* **18**(5), 759 (2012).

- [106] Ahmad, T., Mukherjee, S., Pattnaik, B., Kumar, M., Singh, S., Rehman, R., Tiwari, B. K., Jha, K. A., Barhanpurkar, A. P., Wani, M. R., Roy, S. S., Mabalirajan, U., Ghosh, B., and Agrawal, A. Miro1 regulates intercellular mitochondrial transport & enhances mesenchymal stem cell rescue efficacy. *EMBO J.* **33**(9), 994–1010 (2014).
- [107] Hayakawa, K., Esposito, E., Wang, X., Terasaki, Y., Liu, Y., Xing, C., Ji, X., and Lo, E. H. Transfer of mitochondria from astrocytes to neurons after stroke. *Nature* **535**(7613), 551 (2016).
- [108] Berridge, M. V., Schneider, R. T., and McConnell, M. J. Mitochondrial transfer from astrocytes to neurons following ischemic insult: guilt by association? *Cell Metab.* **24**(3), 376–378 (2016).
- [109] Bacman, S. R., Williams, S. L., Pinto, M., Peralta, S., and Moraes, C. T. Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat. Med.* **19**(9), 1111–1113 (2013).
- [110] Gammage, P. A., Rorbach, J., Vincent, A. I., Rebar, E. J., and Minczuk, M. Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. *EMBO Mol. Med.* **6**(4), 458–466 (2014).
- [111] Reddy, P., Ocampo, A., Suzuki, K., Luo, J., Bacman, S. R., Williams, S. L., Sugawara, A., Okamura, D., Tsunekawa, Y., Wu, J., et al. Selective elimination of mitochondrial mutations in the germline by genome editing. *Cell* **161**(3), 459–469 (2015).
- [112] Pereira, C. V., Bacman, S. R., Arguello, T., Zekonyte, U., Williams, S. L., Edgell, D. R., and Moraes, C. T. mitoTev-TALE: a monomeric DNA editing enzyme to reduce mutant mitochondrial DNA levels. *EMBO Mol. Med.* **10**(9), e8084 (2018).
- [113] Peeva, V., Blei, D., Trombly, G., Corsi, S., Szukszto, M. J., Rebelo-Guiomar, P., Gammage, P. A., Kudin, A. P., Becker, C., Altmüller, J., et al. Linear mitochondrial DNA is rapidly degraded by components of the replication machinery. *Nat. Commun.* **9**(1), 1727 (2018).
- [114] Palmer, J. W., Tandler, B., and Hoppel, C. L. Biochemical properties of subsarcolemmal and interfibrillar mitochondria isolated from rat cardiac muscle. *J. Biol. Chem.* **252**(23), 8731–8739 (1977).
- [115] McKenna, M. C., Stevenson, J. H., Huang, X., and Hopkins, I. B. Differential distribution of the enzymes glutamate dehydrogenase and aspartate aminotransferase in cortical synaptic mitochondria contributes to metabolic compartmentation in cortical synaptic terminals. *Neurochem. Int.* **37**(2-3), 229–241 (2000).
- [116] Benador, I. Y., Veliova, M., Mahdavian, K., Petcherski, A., Wikstrom, J. D., Assali, E. A., Acín-Pérez, R., Shum, M., Oliveira, M. F., Cinti, S., et al. Mitochondria bound to lipid droplets have unique bioenergetics, composition, and dynamics that support lipid droplet expansion. *Cell Metab.* **27**(4), 869–885 (2018).
- [117] Wikstrom, J. D., Twig, G., and Shirihai, O. S. What can mitochondrial heterogeneity tell us about mitochondrial dynamics and autophagy? *Int. J. Biochem. Cell Biol.* **41**(10), 1914–1927 (2009).

- [118] Brown, T. A., Tkachuk, A. N., Shtengel, G., Kopek, B. G., Bogenhagen, D. F., Hess, H. F., and Clayton, D. A. Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction. *Mol. Cell Biol.* **31**(24), 4994–5010 (2011).
- [119] Busch, K. B., Kowald, A., and Spelbrink, J. N. Quality matters: how does mitochondrial network dynamics and quality control impact on mtDNA integrity? *Phil. Trans. R. Soc. B* **369**(1646), 20130442 (2014).
- [120] Allen, J. F. Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. *J. Theor. Biol.* **165**(4), 609–631 (1993).
- [121] Allen, J. F. The function of genomes in bioenergetic organelles. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358**(1429), 19–38 (2003).
- [122] Lane, N. Mitonuclear match: optimizing fitness and fertility over generations drives ageing within generations. *Bioessays* **33**(11), 860–869 (2011).
- [123] Schagger, H. and Pfeiffer, K. Supercomplexes in the respiratory chains of yeast and mammalian mitochondria. *EMBO J.* **19**(8), 1777–1783 (2000).
- [124] Enríquez, J. A. Supramolecular organization of respiratory complexes. *Annu. Rev. Physiol.* **78**, 533–561 (2016).
- [125] Schäfer, E., Seelert, H., Reifschneider, N. H., Krause, F., Dencher, N. A., and Vonck, J. Architecture of active mammalian respiratory chain supercomplexes. *J. Biol. Chem.* **281**(22), 15370–15375 (2006).
- [126] Acín-Pérez, R., Fernández-Silva, P., Peleato, M. L., Pérez-Martos, A., and Enriquez, J. A. Respiratory active mitochondrial supercomplexes. *Mol. Cell* **32**(4), 529–539 (2008).
- [127] Ikeda, K., Shiba, S., Horie-Inoue, K., Shimokata, K., and Inoue, S. A stabilizing factor for mitochondrial respiratory supercomplex assembly regulates energy metabolism in muscle. *Nat. Commun.* **4**, 2147 (2013).
- [128] Mannella, C. A. The relevance of mitochondrial membrane topology to mitochondrial function. *Biochim. Biophys. Acta, Mol. Basis Dis.* **1762**(2), 140–147 (2006).
- [129] Strauss, M., Hofhaus, G., Schröder, R. R., and Kühlbrandt, W. Dimer ribbons of ATP synthase shape the inner mitochondrial membrane. *EMBO J.* **27**(7), 1154–1160 (2008).
- [130] Davies, K. M., Anselmi, C., Wittig, I., Faraldo-Gómez, J. D., and Kühlbrandt, W. Structure of the yeast F_1F_o -ATP synthase dimer and its role in shaping the mitochondrial cristae. *Proc. Natl. Acad. Sci. USA* **109**(34), 13602–13607 (2012).
- [131] Quintana-Cabrera, R., Quirin, C., Glytsou, C., Corrado, M., Urbani, A., Pellattiero, A., Calvo, E., Vázquez, J., Enríquez, J. A., Gerle, C., et al. The cristae modulator Optic atrophy 1 requires mitochondrial ATP synthase oligomers to safeguard mitochondrial function. *Nat. Commun.* **9**(1), 3399 (2018).

- [132] Gilkerson, R. W., Selker, J. M., and Capaldi, R. A. The cristal membrane of mitochondria is the principal site of oxidative phosphorylation. *FEBS Lett.* **546**(2-3), 355–358 (2003).
- [133] Vogel, F., Bornhövd, C., Neupert, W., and Reichert, A. S. Dynamic subcompartmentalization of the mitochondrial inner membrane. *J. Cell Biol.* **175**(2), 237–247 (2006).
- [134] Rieger, B., Junge, W., and Busch, K. B. Lateral pH gradient between OXPHOS complex IV and F₁F₀-ATP-synthase in folded mitochondrial membranes. *Nat. Commun.* **5**, 3103 (2014).
- [135] Frey, T. G. and Mannella, C. A. The internal structure of mitochondria. *Trends Biochem. Sci.* **25**(7), 319–324 (2000).
- [136] Eisner, V., Picard, M., and Hajnóczky, G. Mitochondrial dynamics in adaptive and maladaptive cellular stress responses. *Nat. Cell Biol.* **20**, 755–765 (2018).
- [137] Patten, D. A., Wong, J., Khacho, M., Soubannier, V., Mailloux, R. J., Pilon-Larose, K., MacLaurin, J. G., Park, D. S., McBride, H. M., Trinkle-Mulcahy, L., et al. OPA1-dependent cristae modulation is essential for cellular adaptation to metabolic demand. *EMBO J.* **33**(22), 2676–2691 (2014).
- [138] Mannella, C. A., Pfeiffer, D. R., Bradshaw, P. C., Moraru, I. I., Slepchenko, B., Loew, L. M., Hsieh, C.-e., Buttle, K., and Marko, M. Topology of the mitochondrial inner membrane: dynamics and bioenergetic implications. *IUBMB Life* **52**(3-5), 93–100 (2001).
- [139] Scorrano, L., Ashiya, M., Buttle, K., Weiler, S., Oakes, S. A., Mannella, C. A., and Korsmeyer, S. J. A distinct pathway remodels mitochondrial cristae and mobilizes cytochrome *c* during apoptosis. *Dev. Cell* **2**(1), 55–67 (2002).
- [140] Yamaguchi, R., Lartigue, L., Perkins, G., Scott, R. T., Dixit, A., Kushnareva, Y., Kuwana, T., Ellisman, M. H., and Newmeyer, D. D. Opa1-mediated cristae opening is Bax/Bak and BH3 dependent, required for apoptosis, and independent of Bak oligomerization. *Mol. Cell* **31**(4), 557–569 (2008).
- [141] Cogliati, S., Enriquez, J. A., and Scorrano, L. Mitochondrial cristae: where beauty meets functionality. *Trends Biochem. Sci.* **41**(3), 261–273 (2016).
- [142] Taylor, R. C., Cullen, S. P., and Martin, S. J. Apoptosis: controlled demolition at the cellular level. *Nature Rev. Mol. Cell Biol.* **9**(3), 231–241 (2008).
- [143] Tam, Z. Y., Cai, Y. H., and Gunawan, R. Elucidating cytochrome *c* release from mitochondria: Insights from an *in silico* three-dimensional model. *Biophys. J.* **99**(10), 3155–3163 (2010).
- [144] Khalifat, N., Puff, N., Bonneau, S., Fournier, J.-B., and Angelova, M. I. Membrane deformation under local pH gradient: mimicking mitochondrial cristae dynamics. *Biophys. J.* **95**(10), 4924–4933 (2008).
- [145] Khalifat, N., Fournier, J.-B., Angelova, M. I., and Puff, N. Lipid packing variations induced by pH in cardiolipin-containing bilayers: the driving force for the cristae-like shape instability. *Biochim. Biophys. Acta.* **1808**(11), 2724–2733 (2011).

- [146] Acehan, D., Malhotra, A., Xu, Y., Ren, M., Stokes, D. L., and Schlame, M. Cardiolipin affects the supramolecular organization of ATP synthase in mitochondria. *Biophys. J.* **100**(9), 2184–2192 (2011).
- [147] Zhang, M., Mileykovskaya, E., and Dowhan, W. Cardiolipin is required for supercomplex formation in the inner mitochondrial membrane. *J. Biol. Chem.* **277**(46), 43553–43556 (2002).
- [148] Pfeiffer, K., Gohil, V., Stuart, R. A., Hunte, C., Brandt, U., Greenberg, M. L., and Schagger, H. Cardiolipin stabilizes respiratory chain supercomplexes. *J. Biol. Chem.* **278**(52), 52873–52880 (2003).
- [149] McKenzie, M., Lazarou, M., Thorburn, D. R., and Ryan, M. T. Mitochondrial respiratory chain supercomplexes are destabilized in Barth syndrome patients. *J. Mol. Biol.* **361**(3), 462–469 (2006).
- [150] Paradies, G., Petrosillo, G., Pistolese, M., and Ruggiero, F. M. Reactive oxygen species affect mitochondrial electron transport complex I activity through oxidative cardiolipin damage. *Gene* **286**(1), 135–141 (2002).
- [151] Fouret, G., Tolika, E., Lecomte, J., Bonafos, B., Aoun, M., Murphy, M. P., Ferreri, C., Chatgililoglu, C., Dubreucq, E., Coudray, C., and Feillet-Coudraya, C. The mitochondrial-targeted antioxidant, MitoQ, increases liver mitochondrial cardiolipin content in obesogenic diet-fed rats. *Biochim. Biophys. Acta* **1847**(10), 1025–1035 (2015).
- [152] Nicholls, D. G. Mitochondrial membrane potential and aging. *Aging Cell* **3**(1), 35–40 (2004).
- [153] Perry, S. W., Norman, J. P., Barbieri, J., Brown, E. B., and Gelbard, H. A. Mitochondrial membrane potential probes and the proton gradient: a practical usage guide. *Biotechniques* **50**(2), 98 (2011).
- [154] Duchen, M. R., Leyssens, A., and Crompton, M. Transient mitochondrial depolarizations reflect focal sarcoplasmic reticular calcium release in single rat cardiomyocytes. *J. Cell Biol.* **142**(4), 975–988 (1998).
- [155] OReilly, C. M., Fogarty, K. E., Drummond, R. M., Tuft, R. A., and Walsh, J. V. Quantitative analysis of spontaneous mitochondrial depolarizations. *Biophys. J.* **85**(5), 3350–3357 (2003).
- [156] Schwarzlander, M., Logan, D. C., Johnston, I. G., Jones, N. S., Meyer, A. J., Fricker, M. D., and Sweetlove, L. J. Pulsing of membrane potential in individual mitochondria: a stress-induced mechanism to regulate respiratory bioenergetics in arabidopsis. *Plant Cell* **24**(3), 1188–1201 (2012).
- [157] Grimm, S. and Brdiczka, D. The permeability transition pore in cell death. *Apoptosis* **12**(5), 841–855 (2007).
- [158] Denton, R., Randle, P., and Martin, B. Stimulation by calcium ions of pyruvate dehydrogenase phosphate phosphatase. *Biochem. J.* **128**(1), 161 (1972).

- [159] Hüser, J. and Blatter, L. A. Fluctuations in mitochondrial membrane potential caused by repetitive gating of the permeability transition pore. *Biochem. J.* **343**(2), 311–317 (1999).
- [160] Wang, W., Fang, H., Groom, L., Cheng, A., Zhang, W., Liu, J., Wang, X., Li, K., Han, P., Zheng, M., et al. Superoxide flashes in single mitochondria. *Cell* **134**(2), 279–290 (2008).
- [161] Schwarzländer, M., Logan, D. C., Fricker, M. D., and Sweetlove, L. J. The circularly permuted yellow fluorescent protein cpYFP that has been used as a superoxide probe is highly responsive to pH but not superoxide in mitochondria: implications for the existence of superoxide ‘flashes’. *Biochem. J.* **437**(3), 381–387 (2011).
- [162] Schwarzländer, M., Wagner, S., Ermakova, Y. G., Belousov, V. V., Radi, R., Beckman, J. S., Buettner, G. R., Demareux, N., Duchon, M. R., Forman, H. J., et al. The ‘mitoflash’ probe cpYFP does not respond to superoxide. *Nature* **514**(7523), E12 (2014).
- [163] Agarwal, A., Wu, P.-H., Hughes, E. G., Fukaya, M., Tischfield, M. A., Langseth, A. J., Wirtz, D., and Bergles, D. E. Transient opening of the mitochondrial permeability transition pore induces microdomain calcium transients in astrocyte processes. *Neuron* **93**(3), 587–605 (2017).
- [164] Twig, G., Elorza, A., Molina, A. J., Mohamed, H., Wikstrom, J. D., Walzer, G., Stiles, L., Haigh, S. E., Katz, S., Las, G., et al. Fission and selective fusion govern mitochondrial segregation and elimination by autophagy. *EMBO J.* **27**(2), 433–446 (2008).
- [165] Lemasters, J. J. Variants of mitochondrial autophagy: Types 1 and 2 mitophagy and micromitophagy (type 3). *Redox Biol.* **2**, 749–754 (2014).
- [166] Ye, K., Lu, J., Ma, F., Keinan, A., and Gu, Z. Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. *Proc. Natl. Acad. Sci. USA* **111**(29), 10654–10659 (2014).
- [167] Yoneda, M., Miyatake, T., and Attardi, G. Complementation of mutant and wild-type human mitochondrial DNAs coexisting since the mutation event and lack of complementation of DNAs introduced separately into a cell within distinct organelles. *Mol. Cell. Biol.* **14**(4), 2699–2712 (1994).
- [168] Hayashi, J.-I., Takemitsu, M., Goto, Y.-i., and Nonaka, I. Human mitochondria and mitochondrial genome function as a single dynamic cellular unit. *J. Cell Biol.* **125**(1), 43–50 (1994).
- [169] Enriquez, J. A., Cabezas-Herrera, J., Bayona-Bafaluy, M. P., and Attardi, G. Very rare complementation between mitochondria carrying different mitochondrial DNA mutations points to intrinsic genetic autonomy of the organelles in cultured human cells. *J. Biol. Chem.* **275**(15), 11207–11215 (2000).
- [170] Ono, T., Isobe, K., Nakada, K., and Hayashi, J.-I. Human cells are protected from mitochondrial dysfunction by complementation of DNA products in fused mitochondria. *Nat. Genet.* **28**(3), 272–275 (2001).

- [171] Wilkens, V., Kohl, W., and Busch, K. Restricted diffusion of OXPHOS complexes in dynamic mitochondria delays their exchange between cristae and engenders a transitory mosaic distribution. *J. Cell Sci.* **126**(1), 103–116 (2013).
- [172] Ferree, A. W., Trudeau, K., Zik, E., Benador, I. Y., Twig, G., Gottlieb, R. A., and Shirihai, O. S. Mitotimer probe reveals the impact of autophagy, fusion, and motility on subcellular distribution of young and old mitochondrial protein and on relative mitochondrial protein age. *Autophagy* **9**(11), 1887–1896 (2013).
- [173] Stewart, J. B. and Chinnery, P. F. The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. *Nature Rev. Genet.* **16**(9), 530–542 (2015).
- [174] Chomyn, A., Martinuzzi, A., Yoneda, M., Daga, A., Hurko, O., Johns, D., Lai, S., Nonaka, I., Angelini, C., and Attardi, G. MELAS mutation in mtDNA binding site for transcription termination factor causes defects in protein synthesis and in respiration but no change in levels of upstream and downstream mature transcripts. *Proc. Natl. Acad. Sci. USA* **89**(10), 4221–4225 (1992).
- [175] Miyabayashi, S., Hanamizu, H., Nakamura, R., Endo, H., and Tada, K. Defects of mitochondrial respiratory enzymes in cloned cells from MELAS fibroblasts. *J. Inherit. Metab. Dis.* **15**(5), 797–802 (1992).
- [176] Gilkerson, R. W., Schon, E. A., Hernandez, E., and Davidson, M. M. Mitochondrial nucleoids maintain genetic autonomy but allow for functional complementation. *J. Cell Biol.* **181**(7), 1117–1128 (2008).
- [177] Yang, L., Long, Q., Liu, J., Tang, H., Li, Y., Bao, F., Qin, D., Pei, D., and Liu, X. Mitochondrial fusion provides an ‘initial metabolic complementation’ controlled by mtDNA. *Cell Mol. Life Sci.* **72**(13), 2585–2598 (2015).
- [178] Ono, T., Kasahara, Y., Nakada, K., and Hayashi, J.-I. Presence of interaction but not complementation between human mtDNAs carrying different mutations within a tRNA^{Leu} (UUR) gene. *Biochem. Biophys. Res. Commun.* **314**(4), 1107–1112 (2004).
- [179] Jacobs, H. T., Lehtinen, S. K., and Spelbrink, J. N. No sex please, we’re mitochondria: a hypothesis on the somatic unit of inheritance of mammalian mtDNA. *BioEssays* **22**(6), 564–572 (2000).
- [180] Kukat, C., Wurm, C. A., Spåhr, H., Falkenberg, M., Larsson, N.-G., and Jakobs, S. Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc. Natl. Acad. Sci. USA* **108**(33), 13534–13539 (2011).
- [181] Kukat, C., Davies, K. M., Wurm, C. A., Spåhr, H., Bonekamp, N. A., Köhl, I., Joos, F., Polosa, P. L., Park, C. B., Posse, V., Falkenberg, M., Jakobs, S., Köhlbrandt, W., and Larsson, N.-G. Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc. Natl. Acad. Sci. USA* **112**(36), 11288–11293 (2015).

- [182] McFaline-Figueroa, J. R., Vevea, J., Swayne, T. C., Zhou, C., Liu, C., Leung, G., Boldogh, I. R., and Pon, L. A. Mitochondrial quality control during inheritance is associated with lifespan and mother–daughter age asymmetry in budding yeast. *Aging Cell* **10**(5), 885–895 (2011).
- [183] Katajisto, P., Döhla, J., Chaffer, C. L., Pentinmikko, N., Marjanovic, N., Iqbal, S., Zoncu, R., Chen, W., Weinberg, R. A., and Sabatini, D. M. Asymmetric apportioning of aged mitochondria between daughter cells is required for stemness. *Science* **348**(6232), 340–343 (2015).
- [184] Adams, W. C., Chen, Y.-H., Kratchmarov, R., Yen, B., Nish, S. A., Lin, W.-H. W., Rothman, N. J., Luchsinger, L. L., Klein, U., Busslinger, M., et al. Anabolism-associated mitochondrial stasis driving lymphocyte differentiation over self-renewal. *Cell Rep.* **17**(12), 3142–3152 (2016).
- [185] Wilkens, V., Kohl, W., and Busch, K. Restricted diffusion of oxphos complexes in dynamic mitochondria delays their exchange between cristae and engenders a transitory mosaic distribution. *J. Cell Sci.*, jcs-108852 (2012).
- [186] Ozawa, T., Natori, Y., Sato, M., and Umezawa, Y. Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nat. Methods* **4**(5), 413–419 (2007).
- [187] van der Wijst, M. G. and Rots, M. G. Mitochondrial epigenetics: an overlooked layer of regulation? *Trends Genet.* **31**(7), 353–356 (2015).
- [188] van der Wijst, M. G., van Tilburg, A. Y., Ruiters, M. H., and Rots, M. G. Experimental mitochondria-targeted DNA methylation identifies GpC methylation, not CpG methylation, as potential regulator of mitochondrial gene expression. *Sci. Rep.* **7**(1), 177 (2017).
- [189] Nakanishi, N., Fukuoh, A., Kang, D., Iwai, S., and Kuraoka, I. Effects of DNA lesions on the transcription reaction of mitochondrial RNA polymerase: implications for bypass RNA synthesis on oxidative DNA lesions. *Mutagenesis* **28**(1), 117–123 (2012).
- [190] Safdar, A., Bourgeois, J. M., Ogborn, D. I., Little, J. P., Hettinga, B. P., Akhtar, M., Thompson, J. E., Melov, S., Mocellin, N. J., Kujoth, G. C., et al. Endurance exercise rescues progeroid aging and induces systemic mitochondrial rejuvenation in mtDNA mutator mice. *Proc. Natl. Acad. Sci. USA* **108**(10), 4135–4140 (2011).
- [191] Safdar, A., Annis, S., Kraytsberg, Y., Laverack, C., Saleem, A., Popadin, K., Woods, D. C., Tilly, J. L., and Khrapko, K. Amelioration of premature aging in mtDNA mutator mouse by exercise: the interplay of oxidative stress, PGC-1 α , p53, and DNA damage. a hypothesis. *Curr. Opin. Genet. Dev.* **38**, 127–132 (2016).
- [192] Suhm, T., Kaimal, J. M., Dawitz, H., Peselj, C., Masser, A. E., Hanzén, S., Ambrožič, M., Smialowska, A., Björck, M. L., Brzezinski, P., et al. Mitochondrial translation efficiency controls cytoplasmic protein homeostasis. *Cell Metab.* **27**(6), 1309–1322 (2018).
- [193] Itsara, L. S., Kennedy, S. R., Fox, E. J., Yu, S., Hewitt, J. J., Sanchez-Contreras, M., Cardozo-Pelaez, F., and Pallanck, L. J. Oxidative stress is not a major contributor to somatic mitochondrial DNA mutations. *PLoS genetics* **10**(2), e1003974 (2014).

- [194] Brandt, T., Mourier, A., Tain, L. S., Partridge, L., Larsson, N.-G., and Kühlbrandt, W. Changes of mitochondrial ultrastructure and function during ageing in mice and drosophila. *Elife* **6**, e24662 (2017).
- [195] Aryaman, J., Johnston, I. G., and Jones, N. S. Mitochondrial heterogeneity. *ArXiv e-prints* (2018).
- [196] Aryaman, J., Bowles, C., Jones, N. S., and Johnston, I. G. Mitochondrial network fragmentation modulates mutant mtDNA accumulation independently of absolute fission-fusion rates. *bioRxiv*, 409128 (2018).
- [197] Aryaman, J., Johnston, I. G., and Jones, N. S. Mitochondrial DNA density homeostasis accounts for a threshold effect in a cybrid model of a human mitochondrial disease. *Biochem. J.* **474**(23), 4019–4034 (2017).
- [198] Aryaman, J., Hoitzing, H., Burgstaller, J. P., Johnston, I. G., and Jones, N. S. Mitochondrial heterogeneity, metabolic scaling and cell death. *BioEssays* **39**(7), 1700001 (2017).
- [199] Keogh, M. J., Wei, W., Aryaman, J., Walker, L., Ameele, J. v. d., Coxhead, J., Wilson, I., Bashton, M., Beck, J., West, J., Chen, R., Haudenschield, C., Barthä, G., Luo, S., Morris, C. M., Jones, N. S., Attemsand, J., and Chinnery, P. F. High prevalence of focal and multi-focal somatic genetic variants in the human brain. *Nat. Commun.* **9**(1), 4257 (2018).
- [200] Wei, W., Keogh, M. J., Aryaman, J., Golder, Z., Kullar, P. J., Wilson, I., Talbot, K., Turner, M. R., McKenzie, C.-A., Troakes, C., Attems, J., Smith, C., Sarraj, S. A., Morris, C., Ansorge, O., Jones, N. S., Ironside, J. W., and Chinnery, P. F. Frequency and signature of somatic variants in 1461 human brain exomes. *Genet. Med.* (2018).
- [201] Keogh, M. J., Wei, W., Aryaman, J., Wilson, I., Talbot, K., Turner, M. R., McKenzie, C.-A., Troakes, C., Attems, J., Smith, C., et al. Oligogenic genetic variation of neurodegenerative disease genes in 980 postmortem human brains. *J. Neurol. Neurosurg. Psychiatry* **89**(8), 813–816 (2018).
- [202] Payne, B. A., Wilson, I. J., Yu-Wai-Man, P., Coxhead, J., Deehan, D., Horvath, R., Taylor, R. W., Samuels, D. C., Santibanez-Koref, M., and Chinnery, P. F. Universal heteroplasmy of human mitochondrial DNA. *Hum. Mol. Genet.* **22**(2), 384–390 (2012).
- [203] Kowald, A. and Klipp, E. Mathematical models of mitochondrial aging and dynamics. In *Prog. Mol. Biol. Transl. Sci.*, volume 127, 63–92. Elsevier (2014).
- [204] Mouli, P. K., Twig, G., and Shirihai, O. S. Frequency and selectivity of mitochondrial fusion are key to its quality maintenance function. *Biophys. J.* **96**(9), 3509–3518 (2009).
- [205] Patel, P. K., Shirihai, O., and Huang, K. C. Optimal dynamics for quality control in spatially distributed mitochondrial networks. *PLoS Comp. Biol.* **9**(7), e1003108 (2013).
- [206] Tam, Z. Y., Gruber, J., Halliwell, B., and Gunawan, R. Mathematical modeling of the role of mitochondrial fusion and fission in mitochondrial DNA maintenance. *PloS ONE* **8**(10), e76230 (2013).

- [207] Tam, Z. Y., Gruber, J., Halliwell, B., and Gunawan, R. Context-dependent role of mitochondrial fusion-fission in clonal expansion of mtDNA mutations. *PLoS Comp. Biol.* **11**(5), e1004183 (2015).
- [208] Figge, M. T., Reichert, A. S., Meyer-Hermann, M., and Osiewacz, H. D. Deceleration of fusion-fission cycles improves mitochondrial quality control during aging. *PLoS Comp. Biol.* **8**(6), e1002576 (2012).
- [209] Hoitzing, H. *Controlling mitochondrial dynamics: population genetics and networks*. PhD thesis, Imperial College London, (2017).
- [210] Durham, S. E., Samuels, D. C., Cree, L. M., and Chinnery, P. F. Normal levels of wild-type mitochondrial DNA maintain cytochrome *c* oxidase activity for two pathogenic mitochondrial DNA mutations but not for m.3243A→G. *Am. J. Hum. Genet.* **81**(1), 189–195 (2007).
- [211] Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp. Phys.* **22**(4), 403–434 (1976).
- [212] Grima, R., Thomas, P., and Straube, A. V. How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations? *J. Chem. Phys.* **135**(8), 084103 (2011).
- [213] Gardiner, C. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics. Springer Berlin Heidelberg, (1985).
- [214] Jacobs, K. *Stochastic processes for physicists: understanding noisy systems*. Cambridge University Press, (2010).
- [215] Constable, G. W., Rogers, T., McKane, A. J., and Tarnita, C. E. Demographic noise can reverse the direction of deterministic selection. *Proc. Natl. Acad. Sci. USA*, 201603693 (2016).
- [216] Cagalinec, M., Safiulina, D., Liiv, M., Liiv, J., Choubey, V., Wareski, P., Veksler, V., and Kaasik, A. Principles of the mitochondrial fusion and fission cycle in neurons. *J. Cell Sci.* **126**(10), 2187–2197 (2013).
- [217] Wilkinson, D. J. *Stochastic modelling for systems biology*. CRC press, (2011).
- [218] Van Kampen, N. G. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, (1992).
- [219] Gillespie, D. T. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **58**, 35–55 (2007).
- [220] Grima, R. An effective rate equation approach to reaction kinetics in small volumes: Theory and application to biochemical reactions in nonequilibrium steady-state conditions. *J. Chem. Phys.* **133**(3), 07B604 (2010).
- [221] Parsons, T. L. and Rogers, T. Dimension reduction for stochastic dynamical systems forced onto a manifold by large drift: a constructive approach with examples from theoretical biology. *J. Phys. A* **50**(41), 415601 (2017).

- [222] Kimura, M. The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. *Genetics* **61**(4), 893–903 (1969).
- [223] Zheng, W., Khrapko, K., Collier, H. A., Thilly, W. G., and Copeland, W. C. Origins of human mitochondrial point mutations as DNA polymerase γ -mediated errors. *Mutat. Res.* **599**(1), 11–20 (2006).
- [224] Chen, H., Vermulst, M., Wang, Y. E., Chomyn, A., Prolla, T. A., McCaffery, J. M., and Chan, D. C. Mitochondrial fusion is required for mtDNA stability in skeletal muscle and tolerance of mtDNA mutations. *Cell* **141**(2), 280–289 (2010).
- [225] Safdar, A., Bourgeois, J. M., Ogborn, D. I., Little, J. P., Hettinga, B. P., Akhtar, M., Thompson, J. E., Melov, S., Mocellin, N. J., Kujoth, G. C., et al. Endurance exercise rescues progeroid aging and induces systemic mitochondrial rejuvenation in mtDNA mutator mice. *Proc. Natl. Acad. Sci. USA* **108**(10), 4135–4140 (2011).
- [226] Yan, Z., Lira, V. A., and Greene, N. P. Exercise training-induced regulation of mitochondrial quality. *Exerc. Sport Sci. Rev.* **40**(3), 159 (2012).
- [227] Suen, D.-F., Narendra, D. P., Tanaka, A., Manfredi, G., and Youle, R. J. Parkin overexpression selects against a deleterious mtDNA mutation in heteroplasmic cybrid cells. *Proc. Natl. Acad. Sci. USA* **107**(26), 11835–11840 (2010).
- [228] Kandul, N. P., Zhang, T., Hay, B. A., and Guo, M. Selective removal of deletion-bearing mitochondrial DNA in heteroplasmic drosophila. *Nat. Commun.* **7**, 13100 (2016).
- [229] Sukhorukov, V. M., Dikov, D., Reichert, A. S., and Meyer-Hermann, M. Emergence of the mitochondrial reticulum from fission and fusion dynamics. *PLoS Comput. Biol.* **8**(10), e1002745 (2012).
- [230] Zamponi, N., Zamponi, E., Cannas, S. A., Billoni, O. V., Helguera, P. R., and Chialvo, D. R. Mitochondrial network complexity emerges from fission/fusion dynamics. *Sci. Rep.* **8**(1), 363 (2018).
- [231] Boudoures, A. L., Saben, J., Drury, A., Scheaffer, S., Modi, Z., Zhang, W., and Moley, K. H. Obesity-exposed oocytes accumulate and transmit damaged mitochondria due to an inability to activate mitophagy. *Dev. Biol.* **426**(1), 126–138 (2017).
- [232] de Almeida, M. J., Luchsinger, L. L., Corrigan, D. J., Williams, L. J., and Snoeck, H.-W. Dye-independent methods reveal elevated mitochondrial mass in hematopoietic stem cells. *Cell Stem Cell* **21**(6), 725–729 (2017).
- [233] Wachsmuth, M., Huebner, A., Li, M., Madea, B., and Stoneking, M. Age-related and heteroplasmy-related variation in human mtDNA copy number. *PLoS Genet.* **12**(3), e1005939 (2016).
- [234] Dölle, C., Flønes, I., Nido, G. S., Miletic, H., Osuagwu, N., Kristoffersen, S., Lilleng, P. K., Larsen, J. P., Tysnes, O.-B., Haugarvoll, K., et al. Defective mitochondrial DNA homeostasis in the substantia nigra in Parkinson disease. *Nat. Commun.* **7**, 13548 (2016).

- [235] Perier, C., Bender, A., García-Arumí, E., Melia, M. J., Bove, J., Laub, C., Klopstock, T., Elstner, M., Mounsey, R. B., Teismann, P., et al. Accumulation of mitochondrial DNA deletions within dopaminergic neurons triggers neuroprotective mechanisms. *Brain* **136**(8), 2369–2378 (2013).
- [236] Cree, L., Patel, S., Pyle, A., Lynn, S., Turnbull, D., Chinnery, P., and Walker, M. Age-related decline in mitochondrial DNA copy number in isolated human pancreatic islets. *Diabetologia* **51**(8), 1440–1443 (2008).
- [237] Mengel-From, J., Thinggaard, M., Dalgård, C., Kyvik, K. O., Christensen, K., and Christiansen, L. Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Hum. Genet.* **133**(9), 1149–1159 (2014).
- [238] Narendra, D., Tanaka, A., Suen, D.-F., and Youle, R. J. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *J. Cell Biol.* **183**(5), 795–803 (2008).
- [239] Kim, I., Rodriguez-Enriquez, S., and Lemasters, J. J. Selective degradation of mitochondria by mitophagy. *Arch. Biochem. Biophys.* **462**(2), 245–253 (2007).
- [240] Dai, Y., Zheng, K., Clark, J., Swerdlow, R. H., Pulst, S. M., Sutton, J. P., Shinobu, L. A., and Simon, D. K. Rapamycin drives selection against a pathogenic heteroplasmic mitochondrial DNA mutation. *Hum. Mol. Genet.* **23**(3), 637–647 (2013).
- [241] McWilliams, T. G., Prescott, A. R., Allen, G. F., Tamjar, J., Munson, M. J., Thomson, C., Muqit, M. M., and Ganley, I. G. mito-QC illuminates mitophagy and mitochondrial architecture in vivo. *J. Cell Biol.* , jcb-201603039 (2016).
- [242] McWilliams, T. G., Prescott, A. R., Montava-Garriga, L., Ball, G., Singh, F., Barini, E., Muqit, M. M., Brooks, S. P., and Ganley, I. G. Basal mitophagy occurs independently of PINK1 in mouse tissues of high metabolic demand. *Cell Metab.* **27**(2), 439–449 (2018).
- [243] Woods, D. C., Khrapko, K., and Tilly, J. L. Influence of maternal aging on mitochondrial heterogeneity, inheritance, and function in oocytes and preimplantation embryos. *Genes* **9**(5), 265 (2018).
- [244] Salk, J. J., Schmitt, M. W., and Loeb, L. A. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nature Rev. Genet.* (2018).
- [245] Elliott, H. R., Samuels, D. C., Eden, J. A., Relton, C. L., and Chinnery, P. F. Pathogenic mitochondrial DNA mutations are common in the general population. *Am. J. Hum. Genet.* **83**(2), 254–260 (2008).
- [246] Newkirk, J., Taylor, R., Howell, N., Bindoff, L., Chinnery, P., Alberti, K., Turnbull, D., and Walker, M. Maternally inherited diabetes and deafness: prevalence in a hospital diabetic population. *Diabetic Med.* **14**(6), 457–460 (1997).

- [247] Majamaa, K., Moilanen, J. S., Uimonen, S., Remes, A. M., Salmela, P. I., Kärppä, M., Majamaa-Voltti, K. A., Rusanen, H., Sorri, M., Peuhkurinen, K. J., et al. Epidemiology of a3243g, the mutation for mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes: prevalence of the mutation in an adult population. *Am. J. Hum. Genet.* **63**(2), 447–454 (1998).
- [248] Manwaring, N., Jones, M. M., Wang, J. J., Rochtchina, E., Howard, C., Mitchell, P., and Sue, C. M. Population prevalence of the MELAS A3243G mutation. *Mitochondrion* **7**(3), 230–233 (2007).
- [249] Sasarman, F., Antonicka, H., and Shoubridge, E. A. The A3243G tRNA^{Leu} (UUR) MELAS mutation causes amino acid misincorporation and a combined respiratory chain assembly defect partially suppressed by overexpression of EFTu and EFG2. *Hum. Mol. Genet.* **17**(23), 3697–3707 (2008).
- [250] Martínez-Reyes, I., Diebold, L. P., Kong, H., Schieber, M., Huang, H., Hensley, C. T., Mehta, M. M., Wang, T., Santos, J. H., Woychik, R., et al. TCA cycle and mitochondrial membrane potential are necessary for diverse biological functions. *Mol. Cell* **61**(2), 199–209 (2015).
- [251] Appleby, R. D., Porteous, W. K., Hughes, G., James, A. M., Shannon, D., Wei, Y.-H., and Murphy, M. P. Quantitation and origin of the mitochondrial membrane potential in human cells lacking mitochondrial DNA. *Eur. J. Biochem.* **262**(1), 108–116 (1999).
- [252] Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O’Shea, E., Pilpel, Y., and Barkai, N. Noise in protein expression scales with natural protein abundance. *Nat. Genet.* **38**(6), 636–643 (2006).
- [253] Bentlage, H. A. and Attardi, G. Relationship of genotype to phenotype in fibroblast-derived transmittochondrial cell lines carrying the 3243 mutation associated with the MELAS encephalomyopathy: shift towards mutant genotype and role of mtDNA copy number. *Hum. Mol. Gen.* **5**(2), 197–205 (1996).
- [254] Buttgereit, F., Brand, M., and Müller, M. Cona induced changes in energy metabolism of rat thymocytes. *Biosci. Rep.* **12**(2), 109–114 (1992).
- [255] Brand, M. D., Couture, P., Else, P. L., Withers, K. W., and Hulbert, A. Evolution of energy metabolism. Proton permeability of the inner membrane of liver mitochondria is greater in a mammal than in a reptile. *Biochem. J.* **275**(1), 81–86 (1991).
- [256] Rolfe, D. and Brown, G. C. Cellular energy utilization and molecular origin of standard metabolic rate in mammals. *Physiol. Rev.* **77**(3), 731–758 (1997).
- [257] Ojala, D., Montoya, J., and Attardi, G. tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**(5806), 470–474 (1981).
- [258] Chujo, T., Ohira, T., Sakaguchi, Y., Goshima, N., Nomura, N., Nagao, A., and Suzuki, T. LRPPRC/SLIRP suppresses PNPase-mediated mRNA decay and promotes polyadenylation in human mitochondria. *Nucleic Acids Res.* **40**(16), 8033–8047 (2012).

- [259] Lane, N. and Martin, W. The energetics of genome complexity. *Nature* **467**(7318), 929–934 (2010).
- [260] Lane, N. Energetics and genetics across the prokaryote-eukaryote divide. *Biol. Direct* **6**, 35 (2011).
- [261] Johnston, I. G. and Williams, B. P. Evolutionary inference across eukaryotes identifies specific pressures favoring mitochondrial gene retention. *Cell Syst.* **2**(2), 101–111 (2016).
- [262] Kowald, A. and Kirkwood, T. B. Evolution of the mitochondrial fusion–fission cycle and its role in aging. *Proc. Natl. Acad. Sci. USA* **108**(25), 10237–10242 (2011).
- [263] Dieteren, C. E., Gielen, S. C., Nijtmans, L. G., Smeitink, J. A., Swarts, H. G., Brock, R., Willems, P. H., and Koopman, W. J. Solute diffusion is hindered in the mitochondrial matrix. *Proc. Natl. Acad. Sci. USA* **108**(21), 8657–8662 (2011).
- [264] Davis, A. F. and Clayton, D. A. In situ localization of mitochondrial DNA replication in intact mammalian cells. *J. Cell Biol.* **135**(4), 883–893 (1996).
- [265] Imamura, H., Nhat, K. P. H., Togawa, H., Saito, K., Iino, R., Kato-Yamada, Y., Nagai, T., and Noji, H. Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators. *Proc. Natl. Acad. Sci. USA* **106**(37), 15651–15656 (2009).
- [266] Whitfield, M. L., Sherlock, G., Saldanha, A. J., Murray, J. I., Ball, C. A., Alexander, K. E., Matese, J. C., Perou, C. M., Hurt, M. M., Brown, P. O., et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* **13**(6), 1977–2000 (2002).
- [267] Murphy, M. P. How mitochondria produce reactive oxygen species. *Biochem. J.* **417**(1), 1–13 (2009).
- [268] Brodie, A. E. and Reed, D. J. Cellular recovery of glyceraldehyde-3-phosphate dehydrogenase activity and thiol status after exposure to hydroperoxides. *Arch. Biochem. Biophys.* **276**(1), 212–218 (1990).
- [269] Knight, R. J., Kofoed, K. F., Schelbert, H. R., and Buxton, D. B. Inhibition of glyceraldehyde-3-phosphate dehydrogenase in post-ischaemic myocardium. *Cardiovasc. Res.* **32**(6), 1016–1023 (1996).
- [270] James, A. M., Yau-Huei, W., Cheng-Yoong, P., and Murphy, M. P. Altered mitochondrial function in fibroblasts containing MELAS or MERRF mitochondrial DNA mutations. *Biochem. J.* **318**(2), 401–407 (1996).
- [271] Danielson, S. R., Carelli, V., Tan, G., Martinuzzi, A., Schapira, A. H., Savontaus, M.-L., and Cortopassi, G. A. Isolation of transcriptomal changes attributable to LHON mutations and the cybridization process. *Brain* **128**(5), 1026–1037 (2005).

- [272] Hao, H., Morrison, L. E., and Moraes, C. T. Suppression of a mitochondrial tRNA gene mutation phenotype associated with changes in the nuclear background. *Hum. Mol. Gen.* **8**(6), 1117–1124 (1999).
- [273] Wilkins, H. M., Carl, S. M., and Swerdlow, R. H. Cytoplasmic hybrid (cybrid) cell lines as a practical model for mitochondriopathies. *Redox Biol.* **2**, 619–631 (2014).
- [274] Otten, A. B., Theunissen, T. E., Derhaag, J. G., Lambrichs, E. H., Boesten, I. B., Winandy, M., van Montfoort, A. P., Tarbashevich, K., Raz, E., Gerards, M., et al. Differences in strength and timing of the mtDNA bottleneck between zebrafish germline and non-germline cells. *Cell Rep.* **16**(3), 622–630 (2016).
- [275] Goto, Y.-i., Nonaka, I., and Horai, S. A mutation in the tRNA^{Leu} (UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* **348**(6302), 651–653 (1990).
- [276] Viscomi, C., Bottani, E., Civiletto, G., Cerutti, R., Moggio, M., Fagiolari, G., Schon, E. A., Lamperti, C., and Zeviani, M. In vivo correction of cox deficiency by activation of the ampk/pgc-1 α axis. *Cell Metab.* **14**(1), 80–90 (2011).
- [277] Jain, I. H., Zazzeron, L., Goli, R., Alexa, K., Schatzman-Bone, S., Dhillon, H., Goldberger, O., Peng, J., Shalem, O., Sanjana, N. E., et al. Hypoxia as a therapy for mitochondrial disease. *Science* **352**(6281), 54–61 (2016).
- [278] Bough, K. J., Wetherington, J., Hassel, B., Pare, J. F., Gawryluk, J. W., Greene, J. G., Shaw, R., Smith, Y., Geiger, J. D., and Dingledine, R. J. Mitochondrial biogenesis in the anticonvulsant mechanism of the ketogenic diet. *Ann. Neurol.* **60**(2), 223–235 (2006).
- [279] Ahola-Erkkilä, S., Carroll, C., Peltola-Mjösund, K., Tulkki, V., Mattila, I., Seppänen-Laakso, T., Orešič, M., Tynjismaa, H., and Suomalainen, A. Ketogenic diet slows down mitochondrial myopathy progression in mice. *Hum. Mol. Gen.* **19**(10), 1974–1984 (2010).
- [280] Steriade, C., Andrade, D. M., Faghfoury, H., Tarnopolsky, M. A., and Tai, P. Mitochondrial encephalopathy with lactic acidosis and stroke-like episodes (MELAS) may respond to adjunctive ketogenic diet. *Pediatr. Neurol.* **50**(5), 498–502 (2014).
- [281] Nicholls, D. G. The effective proton conductance of the inner membrane of mitochondria from brown adipose tissue. *Eur. J. Biochem.* **77**(2), 349–356 (1977).
- [282] Smiley, S. T., Reers, M., Mottola-Hartshorn, C., Lin, M., Chen, A., Smith, T. W., Steele, G., and Chen, L. B. Intracellular heterogeneity in mitochondrial membrane potentials revealed by a J-aggregate-forming lipophilic cation JC-1. *Proc. Natl. Acad. Sci. USA* **88**(9), 3671–3675 (1991).
- [283] Mathur, A., Hong, Y., Kemp, B. K., Barrientos, A. A., and Erusalimsky, J. D. Evaluation of fluorescent dyes for the detection of mitochondrial membrane potential changes in cultured cardiomyocytes. *Cardiovasc. Res.* **46**(1), 126–138 (2000).

- [284] Tzur, A., Moore, J. K., Jorgensen, P., Shapiro, H. M., and Kirschner, M. W. Optimizing optical flow cytometry for cell volume-based sorting and analysis. *PLOS ONE* **6**(1), e16053 (2011).
- [285] Zhou, W., Kang, H. C., O’Grady, M., Chambers, K. M., Dubbels, B., Melquist, P., and Gee, K. R. CellTraceTM Far Red & CellTrackerTM Deep Red—long term live cell tracking for flow cytometry and fluorescence microscopy. *J. Biol. Methods* **3**(1), e38 (2016).
- [286] Miettinen, T. P. and Björklund, M. Mitochondrial function and cell size: An allometric relationship. *Trends Cell Biol.* (2017).
- [287] West, G. B., Brown, J. H., and Enquist, B. J. A general model for the origin of allometric scaling laws in biology. *Science* **276**(5309), 122–126 (1997).
- [288] West, G. B., Woodruff, W. H., and Brown, J. H. Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. *Proc. Natl. Acad. Sci. USA* **99**(suppl 1), 2473–2478 (2002).
- [289] Aon, M. A., Cortassa, S., and O’Rourke, B. Percolation and criticality in a mitochondrial network. *Proc. Natl. Acad. Sci. USA* **101**(13), 4447–4452 (2004).
- [290] Darveau, C.-A., Suarez, R. K., Andrews, R. D., and Hochachka, P. W. Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* **417**(6885), 166–170 (2002).
- [291] Savage, V. M., Deeds, E. J., and Fontana, W. Sizing up allometric scaling theory. *PLoS Comput. Biol.* **4**(9), e1000171 (2008).
- [292] Kolokotronis, T., Savage, V., Deeds, E. J., and Fontana, W. Curvature in metabolic scaling. *Nature* **464**(7289), 753–756 (2010).
- [293] Buttgerit, F. and Brand, M. D. A hierarchy of ATP-consuming processes in mammalian cells. *Biochem. J.* **312**(1), 163–167 (1995).
- [294] Gottlieb, E., Armour, S., Harris, M., and Thompson, C. Mitochondrial membrane potential regulates matrix configuration and cytochrome *c* release during apoptosis. *Cell Death Differ.* **10**(6), 709–717 (2003).
- [295] Vargas-Garcia, C. A., Soltani, M., and Singh, A. Conditions for cell size homeostasis: A stochastic hybrid systems approach. *IEEE Life Sci. Lett.* (2016).
- [296] Sompayrac, L. and Maaløe, O. Autorepressor model for control of DNA replication. *Nature* **241**(109), 133–135 (1973).
- [297] Amir, A. Cell size regulation in bacteria. *Phys. Rev. Lett.* **112**(20), 208102 (2014).
- [298] Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., and Jun, S. Cell-size control and homeostasis in bacteria. *Curr. Biol.* **25**(3), 385–391 (2015).
- [299] Kerr, J. F., Wyllie, A. H., and Currie, A. R. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br. J. Cancer* **26**(4), 239 (1972).

- [300] Modi, S. K., Vargas-Garcia, C. A., Ghusinga, K. R., and Singh, A. Analysis of noise mechanisms in cell size control. *bioRxiv*, 080465 (2016).
- [301] Otera, H., Miyata, N., Kuge, O., and Mihara, K. Drp1-dependent mitochondrial fission via mid49/51 is essential for apoptotic cristae remodeling. *J. Cell Biol.* **212**(5), 531–544 (2016).
- [302] Van Dijk, D., Dhar, R., Missarova, A. M., Espinar, L., Blevins, W. R., Lehner, B., and Carey, L. B. Slow-growing cells within isogenic populations have increased RNA polymerase error rates and DNA damage. *Nat. Commun.* **6** (2015).
- [303] Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M. C., Tassani, S., Piva, F., et al. An estimation of the number of cells in the human body. *Ann. Hum. Biol.* **40**(6), 463–471 (2013).
- [304] Frank, S. A. and Nowak, M. A. Cell biology: developmental predisposition to cancer. *Nature* **422**(6931), 494 (2003).
- [305] Frank, S. A. Somatic evolutionary genomics: mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. *Proc. Natl. Acad. Sci. U.S.A.* **107**(suppl 1), 1725–1730 (2010).
- [306] Tomasetti, C., Vogelstein, B., and Parmigiani, G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1999–2004 (2013).
- [307] Frank, S. A. Somatic mosaicism and cancer: inference based on a conditional Luria–Delbrück distribution. *J. Theor. Biol.* **223**(4), 405–412 (2003).
- [308] Youssoufian, H. and Pyeritz, R. E. Human genetics and disease: Mechanisms and consequences of somatic mosaicism in humans. *Nat. Rev. Genet.* **3**(10), 748 (2002).
- [309] Singleton, A. and Hardy, J. The evolution of genetics: Alzheimer’s and Parkinson’s diseases. *Neuron* **90**(6), 1154–1163 (2016).
- [310] Guerreiro, R., Brás, J., and Hardy, J. SnapShot: Genetics of ALS and FTD. *Cell* **160**(4), 798–798 (2015).
- [311] Luria, S. E. and Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**(6), 491 (1943).
- [312] Lea, D. E. and Coulson, C. A. The distribution of the numbers of mutants in bacterial populations. *J. Genet.* **49**(3), 264 (1949).
- [313] Zheng, Q. Progress of a half century in the study of the Luria–Delbrück distribution. *Math. Biosci.* **162**(1-2), 1–32 (1999).
- [314] Zheng, Q. Statistical and algorithmic methods for fluctuation analysis with SALVADOR as an implementation. *Math. Biosci.* **176**(2), 237–252 (2002).

- [315] Durrett, R. *Branching Process Models of Cancer*. Mathematical Biosciences Institute Lecture Series. Springer International Publishing, (2015).
- [316] Evrony, G. D., Lee, E., Mehta, B. K., Benjamini, Y., Johnson, R. M., Cai, X., Yang, L., Haseley, P., Lehmann, H. S., Park, P. J., et al. Cell lineage analysis in human brain using endogenous retroelements. *Neuron* **85**(1), 49–59 (2015).
- [317] Lodato, M. A., Woodworth, M. B., Lee, S., Evrony, G. D., Mehta, B. K., Karger, A., Lee, S., Chittenden, T. W., DGama, A. M., Cai, X., et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, 94–98 (2015).
- [318] Bae, T., Tomasini, L., Mariani, J., Zhou, B., Roychowdhury, T., Franjic, D., Pletikos, M., Pattni, R., Chen, B.-J., Venturini, E., et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science* , eaan8690 (2017).
- [319] Roach, J. C., Glusman, G., Smit, A. F., Huff, C. D., Hubley, R., Shannon, P. T., Rowen, L., Pant, K. P., Goodman, N., Bamshad, M., et al. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* **328**, 636–639 (2010).
- [320] Tomasetti, C. and Vogelstein, B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* **347**(6217), 78–81 (2015).
- [321] Tomasetti, C., Li, L., and Vogelstein, B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science* **355**(6331), 1330–1334 (2017).
- [322] Frost, B., Ollesch, J., Wille, H., and Diamond, M. I. Conformational diversity of wild-type Tau fibrils specified by templated conformation change. *J. Biol. Chem.* **284**(6), 3546–3551 (2009).
- [323] Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., Lent, R., Herculano-Houzel, S., et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* **513**, 532–541 (2009).
- [324] Ross, S. M. et al. *Stochastic processes*, volume 2. John Wiley & Sons New York, (1996).
- [325] Csilléry, K., Blum, M. G., Gaggiotti, O. E., and François, O. Approximate Bayesian computation (ABC) in practice. *Trends Ecol. Evol.* **25**(7), 410–418 (2010).
- [326] Cooper, D. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum. Genet.* **133**, 19 (2014).
- [327] Yamaguchi, Y. and Miura, M. Programmed cell death in neurodevelopment. *Dev. Cell* **32**, 478–490 (2015).
- [328] Oppenheim, R. W. Cell death during development of the nervous system. *Annu. Rev. Neurosci.* **14**, 453–501 (1991).
- [329] Vogel, F. and Motulsky, A. G. *Vogel and Motulsky’s Human Genetics: Problems and Approaches*. Springer Science & Business Media, (2013).

- [330] Florio, M. and Huttner, W. B. Neural progenitors, neurogenesis and the evolution of the neo-cortex. *Development* **141**(11), 2182–2194 (2014).
- [331] Rakic, P. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. *Trends Neurosci.* **18**(9), 383–388 (1995).
- [332] Marín, O. and Rubenstein, J. L. Cell migration in the forebrain. *Annu. Rev. Neurosci.* **26**, 441–483 (2003).
- [333] Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S., and Kriegstein, A. R. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* **409**, 714–720 (2001).
- [334] Horton, J. C. and Adams, D. L. The cortical column: a structure without a function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **360**, 837–862 (2005).
- [335] Kauppila, J. H., Baines, H. L., Bratic, A., Simard, M.-L., Freyer, C., Mourier, A., Stamp, C., Filograna, R., Larsson, N.-G., Greaves, L. C., et al. A phenotype-driven approach to generate mouse models with pathogenic mtDNA mutations causing mitochondrial disease. *Cell Rep.* **16**(11), 2980–2990 (2016).
- [336] Brandon, M. C., Lott, M. T., Nguyen, K. C., Spolim, S., Navathe, S. B., Baldi, P., and Wallace, D. C. MITOMAP: a human mitochondrial genome database 2004 update. *Nucleic Acids Res.* **33**(suppl_1), D611–D613 (2005).
- [337] Smith, A. C. and Robinson, A. J. Mitominer v3. 1, an update on the mitochondrial proteomics database. *Nucleic Acids Res.* **44**(D1), D1258–D1261 (2015).
- [338] Lewis, S. C., Uchiyama, L. F., and Nunnari, J. ER-mitochondria contacts couple mtDNA synthesis with mitochondrial division in human cells. *Science* **353**(6296), aaf5549 (2016).
- [339] Medeiros, T. C., Thomas, R. L., Ghillebert, R., and Graef, M. Autophagy balances mtDNA synthesis and degradation by DNA polymerase POLG during starvation. *J. Cell Biol.*, jcb-201801168 (2018).
- [340] Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**(1), 97–109 (1970).
- [341] Haario, H., Saksman, E., and Tamminen, J. An adaptive metropolis algorithm. *Bernoulli* **7**(2), 223–242 (2001).
- [342] Alberts, B., Johnson, A., Lewis, J., Roberts, K., Raff, M., and Walter, P. *Molecular Biology of the Cell*. Garland Science, (2002).
- [343] Wasserman, L. *All of Statistics: A Concise Course in Statistical Inference*. Springer Texts in Statistics. Springer New York, (2013).
- [344] Tzur, A., Kafri, R., LeBleu, V. S., Lahav, G., and Kirschner, M. W. Cell growth and size homeostasis in proliferating animal cells. *Science* **325**(5937), 167–171 (2009).

- [345] Neumann, F. R. and Nurse, P. Nuclear size control in fission yeast. *J. Cell Biol.* **179**(4), 593–600 (2007).
- [346] Schuler, M., Muehlbauer, P., Guzzie, P., and Eastmond, D. Noscapine hydrochloride disrupts the mitotic spindle in mammalian cells and induces aneuploidy as well as polyploidy in cultured human lymphocytes. *Mutagenesis* **14**(1), 51–56 (1999).
- [347] Bereiter-Hahn, J. and Vöth, M. Distribution and dynamics of mitochondrial nucleoids in animal cells in culture. In *Experimental Biology Online*, 58–77. Springer (1998).
- [348] Alán, L., Zelenka, J., Ježek, J., Dlasková, A., and Ježek, P. Fluorescent in situ hybridization of mitochondrial DNA and RNA. *Acta Biochim. Pol.* **57**(4), 403 (2010).
- [349] Chatre, L. and Ricchetti, M. Large heterogeneity of mitochondrial DNA transcription and initiation of replication exposed by single-cell imaging. *J. Cell Sci.* **126**(4), 914–926 (2013).
- [350] Sakaue-Sawano, A., Kurokawa, H., Morimura, T., Hanyu, A., Hama, H., Osawa, H., Kashiwagi, S., Fukami, K., Miyata, T., Miyoshi, H., et al. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**(3), 487–498 (2008).
- [351] Ahn, C. S. and Metallo, C. M. Mitochondria as biosynthetic factories for cancer proliferation. *Cancer Metab.* **3**(1), 1 (2015).