

SYSTEMS BIOLOGY OF AGING: MODELING & ANALYSIS OF MITOCHONDRIAL GENOME INTEGRITY

SURESH KUMAR POOVATHINGAL

(M.S., Technische Universität Hamburg, Germany)

(B. Engg., Bangalore University, India)



A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMICAL & BIOMOLECULAR ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE

2011

Dedicated to;

My Family
For all their love & support



देवानां च ऋषीणां च गुरुं काञ्चनसंनिभम् ।
बुद्धिभूतं त्रिलोकेशं तं नमामि बृहस्पतिम् ॥५॥

Acknowledgements

Research progress and activities during my PhD work has been an incredible journey and a monumental milestone in my academic life. It has been enjoyable, valuable and inspiring from the very beginning. I could not have embarked on this expedition and traveled this far without the passionate and continued support of many people. I am deeply grateful to them and I want to thank some of them in particular.

First and foremost, my heart felt gratitude to my advisors, Dr. Rudiyanto Gunawan and Dr. Jan Gruber. I am greatly indebted to Dr. Gunawan for his immense support, timely inputs, patience and precise guidance during the past years of my PhD work. I am always amazed by your ability to identify important problems in a field. It has been a great honor to work with you and to get an excellent opportunity to learn from you.

I would also like to express my deep admiration to Dr. Gruber, for being a great and bright scientist, always open to discuss new ideas and find time for brainstorming. Several of these discussions have been instrumental in crystallizing the framework of this dissertation. In him, I have realized a great mentor, good motivator and a good friend for life. Your dedication, encouragement, and support constantly inspired me to better myself and aim higher. I am at a loss of words.

I would like to admire and thank Professor Barry Halliwell for his continued support and constructive feedback on different aspects of my research work and to Dr. Saif Abdul Kadir Khan for all his support in completing this dissertation. I would also like to thank Ng Li Fang from Department of Biochemistry, for her help in providing me with the experimental data. My heart felt appreciation to Dr. Lakshminarayanan Samavedham and Dr. Thilo Hagen for their acceptance of being in my oral qualifying examination committee. Their suggestions and critiques have been very useful in planning my research. I also would like to thank the final thesis reviewers for spending their time in evaluating the work of this dissertation.

I highly appreciate my fellow lab mates for all their help and support during my PhD work. Specifically, I am grateful to Lakshminarayanan Lakshmanan, Sridharan Srinath and Thanneer Malai Perumal and other collaborators Dr. Ce Belle Chan and Dr Sebastian Shaffer, who took their time away from their own over-extended lives to provide feedback on this thesis and other publications related to the present dissertation. Your suggestions have been definitely valuable and have made the contents related to my thesis more presentable. Unrestrained comments and constructive feedbacks from unknown reviewers of our publications have definitely helped us to bring out a good research work.

Thanks also to scientific staff of ChBE, NUS, especially, Mr Boey Kok Hong and Ms Samantha Fam for their technical and administrative support. Sincere thanks also goes to Mr Anand Vaidya of VSA services, for his useful suggestions and relentless technical support related to Linux programming. The exciting journey during my PhD work would have been certainly less exciting, if not for the engaging and heartwarming time spent with friends in NUS and Singapore.

I profoundly appreciate Chemical and Biomolecular Engineering department at Na-

tional University of Singapore for providing me an opportunity to carry out research and for providing scholarship for my PhD work, ministry of education for providing grants to perform research projects. I would also like to express my sincere gratitude to the professors of ChBE and NUS for the valuable lectures and seminars hosted in the department. As I mention of courses, I must certainly mention of Dr. Lakshminarayanan Samavedham, whose course on advanced mathematical modeling has been really instrumental for this dissertation. His knowledge, organization and enthusiasm are some things that can never be appreciated enough.

Finally and most importantly, I thank my family members for all their support and perseverance during my PhD. I am grateful to my family: my parents, Mr. Sukumaran and Ms. Ramani and my younger brother, Rajesh. Without your love, moral support and guidance through the years of my life, I would certainly not have reached this far. I thank my wife Madhavi, for her love, patience and support during the course of this work and in aiding me for successfully completing this journey. A special thanks to my son, Anirudh, who has always been able to make me smile during the past few months, no matter how daunting the graduation seemed to be! Last but not the least, I wish to thank my in-laws, Mr. Haridas and Ms. Padmaja, for supporting my wife during the crucial stages of my PhD work.

Suresh Poovathingal
NUS
March 2011

Contents

Acknowledgements	i
Summary	v
List of Figures	xi
List of Tables	xiii
Keywords	xiv
Abbreviations & Nomenclature	xvi
1 Introduction	1
1.1 Theories of aging	1
1.2 Molecular mechanisms of aging	3
1.3 Mitochondria	6
1.3.1 Oxidative phosphorylation process	7
1.3.2 Generation of Reactive Oxygen Species	9
1.3.3 Mitochondrial genetics	16
1.3.4 Mitochondrial DNA mutagenesis	18
1.4 Research motivation - accumulation of mtDNA mutations in aging	20
1.4.1 Mitochondrial Free Radical Theory of Aging	20
1.4.2 Mouse models - human aging phenotypes	23
1.4.3 Mutation accumulation dynamics - significance of stochasticity	24
1.5 Thesis scope & organization	29
2 Systems biology of aging — an overview	32
2.1 Mathematical modeling in biology	32
2.1.1 Historical perspective	32
2.1.2 Characteristics of modeling	34
2.1.3 Advantages of modeling	35
2.1.4 Different kinds of models	37
2.2 Origin of cellular stochasticity	38
2.2.1 Stochasticity in biology	38
2.2.2 Stochasticity in aging process	40
2.3 Systems Biology & aging research	41
2.4 Models of cellular aging - an overview	43
2.4.1 Models of intra-cellular mechanisms	44
2.4.2 Role of mitochondria in aging	45

3	Stochastic drift in mitochondrial DNA point mutations: a novel perspective <i>ex silico</i>	58
3.1	Introduction	58
3.2	Methods	60
3.2.1	Chemical Master Equation	60
3.2.2	Gillespie's algorithm	61
3.2.3	<i>In silico</i> mouse model	63
3.2.4	Simulations of mouse development	68
3.2.5	Simulation of postnatal stage	69
3.2.6	Simulation of POLG mutator mouse models	70
3.2.7	Model parameters	70
3.2.8	Computation of mtDNA mutation frequency	78
3.3	Results & Discussion	78
3.3.1	Statistical features of the RMC assay.	78
3.3.2	Transgenic mouse studies.	83
3.3.3	Rate of mitochondrial turnover	84
3.3.4	Significance of animal development.	89
3.4	Conclusions	92
4	Elucidating mechanisms of age-dependent clonal accumulation of mitochondrial DNA mutations — an <i>in silico</i> approach	96
4.1	Introduction	96
4.2	Methods	97
4.2.1	Cell-level modeling details	98
4.2.2	Tissue-level modeling details	100
4.2.3	Model parameters	101
4.2.4	Modeling details of different hypotheses	105
4.3	Results & Discussion	110
4.3.1	Clonal expansion mechanism	110
4.3.2	Influence of mtDNA replicative advantage	115
4.3.3	Cluster size distribution of clonally expanded cells	119
4.4	Conclusions	123
5	Sarcopenia <i>in silico</i> — a multiscale stochastic-deterministic modeling approach	126
5.1	Introduction	126
5.2	Methods	128
5.2.1	Stochastic model	129
5.2.2	Reaction-diffusion model	131
5.3	Results & Discussion	134
5.3.1	Simulation & validation Results	134
5.3.2	Parametric perturbation analysis	139
5.3.3	Perturbation analysis related to Caloric Restriction	145
5.4	Conclusions	152
6	Maximizing signal-to-noise ratio in the Random Mutation Capture assay	156
6.1	Introduction	156
6.2	Methods	157
6.2.1	Statistical optimization of the RMC assay	157
6.2.2	Validation of RMC optimization	161
6.3	Results & Discussion	162

6.4	Conclusions	167
7	Global parameter estimation methods for stochastic biochemical systems	170
7.1	Introduction	170
7.2	Methods	172
7.2.1	Estimation of state density function	172
7.2.2	Parameter estimation methods	173
7.2.3	Global optimization algorithm	176
7.3	Results	177
7.3.1	Case Study 1: RNA dynamics in <i>E. coli</i>	177
7.3.2	Case Study 2: Galactose uptake model in <i>S. cerevisiae</i>	180
7.3.3	Case Study 3: Stochastic model of a synthetic toggle switch	186
7.3.4	Case Study 4: The Schlögl Model	189
7.4	Discussion	190
7.5	Conclusions	194
8	Conclusions & future directions	195
8.1	Conclusions	195
8.2	Recommendations	200
	References	207
A	Citations of Random Mutation Capture Assay	229
A.1	Literature citations: RMC assay	229
B	COMSOL report — sarcopenic model	233
B.1	COMSOL report	233

Summary

The mitochondrial free radical theory of aging (mFRTA) implicates reactive oxygen species (ROS) as the causal agent of degeneration of mitochondrial genome integrity and subsequent cellular respiratory dysfunctions, leading to tissue degeneration and aging. While several premises of mitochondrial theory of aging are intensely debated, there exist overwhelming evidence supporting the importance of mtDNA mutations in aging. Significant contests in the premises of mFRTA are associated with high uncertainty in the reported age-dependent mitochondrial data such as the mtDNA mutation burden in cells. The source of such variability may be multifaceted, including both intrinsic cellular stochasticity and variability due to measurement protocols. In addition, while there exists ample amount of experimental evidence hinting for the direct role of mitochondria and mtDNA in cellular physiology and aging, there is still a large knowledge gap in understanding the mechanisms of origin and accumulation dynamics of mtDNA mutations during aging. Consequently, to better understand mtDNA mutation dynamics during aging, the focus of the present thesis revolves around the following aspects:

- ★ to develop parsimonious stochastic models of mtDNA mutation dynamics that encompass only the most relevant biological processes, in order to elucidate the origin, mechanism and consequences of mutation accumulation,
- ★ to design experiments that minimize variability associated with mtDNA mutation measurements, i.e. maximize the signal-to-noise ratio, and

-
- ★ to develop practical kinetic parameter estimation methods for stochastic models based on either single cell or cell population level data.

In order to capture the stochasticity associated with the inherent mtDNA turnover dynamics, a stochastic model of mouse heart tissue was developed based on chemical master equation (CME). The first model captured the random drift of mtDNA mutations in heart tissue during the embryonic development and post-natal life of mouse population. The model specifically simulated the mtDNA point mutation accumulation in both wild-type and mutator mice. The model was formulated as a Markov based framework of CME and solved using a Monte Carlo approach known as the Stochastic Simulation Algorithm (SSA). The simulation outcome signified the coupling between experimental variability and intrinsic stochasticity of mtDNA mutation within the measurement data. The contribution of measurement variability to the overall data uncertainty, arising due to the statistically suboptimal sampling procedure, was observed to be substantial and thereby causing significant uncertainty in determining the trend of mtDNA mutation accumulation dynamics from the experimental data. Model simulations further indicated that *de novo* mutations during embryonic development play a critical role in post-birth mutation accumulation in postmitotic tissues like heart. This effect was found to be much more prominent in the case of mutator mice, where much of the high mutation burden in the heart tissue was already acquired at birth, as seen in the experiments as well.

One paradox associated with the role of mtDNA mutations in cellular aging is the extent of tissue level mutation burden found in homogenate studies. Although mutation load have been found to increase with age, the overall mutation burden was found to be below any functionally relevant level, considering cells contain several hundred copies of mtDNA. Single cell studies have greatly contributed in resolving this mystery by indicating focal accumulation of mtDNA mutations in some cells of a tissue. Such focal accumulation resulted from a process know as clonal expansion, which has been found to

cause degeneration of post mitotic tissue like heart and brain. Since the exact mechanism of clonal expansion is not known completely, a stochastic modeling framework of mtDNA mutation dynamics was also developed to gain insights into the origin and dynamics of this process. The first model described above was extended to include nuclear retrograde signaling in rat cardiomyocytes. The model was further modified to test different hypothesis commonly proffered to explain the clonal expansion process. Simulation results indicated that random segregation of mtDNA copies during developmental cell lineage with replicative advantage of mutant mtDNA in the presence of nuclear retrograde response was required to best validate the experimental data. Similar to the results obtained from the point mutation model, the expansion of *de novo* mtDNA mutations during embryogenesis was found to play an important role in mosaicity of mtDNA mutation burden in cells of postmitotic tissue.

Further, the stochastic mtDNA mutation models developed above were integrated into a multi-scale hybrid modeling (stochastic and deterministic) framework to gain understanding of the etiology and progression of *sarcopenia*, an important pathology associated with aging. Turnover dynamics of mtDNA in muscle progenitor cells during embryogenesis was simulated using stochastic model, while the progression of mtDNA mutations in muscle fibers after birth was modeled using a reaction-diffusion partial differential equation, describing the active transport of mtDNA and clonal expansion. Key attributes of rat skeletal muscle, such as the extent of clonal expansion in muscle fiber at different ages were validated against the experimental data. Simulation results further indicated that mtDNA diffusivity parameter, which has been observed to be modulated by the mitochondrial fusion-fission process, plays a very important role in the progression of sarcopenia and the subsequent tissue degeneration. Particularly, parametric sensitivity analysis of the hybrid model provided insights on the development and progression of sarcopenia, thereby providing important insights into the possible interventions to retard/prolong the severity

of sarcopenia. Specifically, using the hybrid model, different hypothesized mechanisms of the beneficial effects of caloric restriction in retarding the progression of sarcopenia was analyzed. Simulation results indicated that the ability of the post-mitotic cells to sustain enhanced stress, in terms of withstanding higher cellular mtDNA mutation burden, might be a potential mechanism to explain the intervention of CR in reducing the progression of sarcopenia.

Model simulations pointed to a large uncertainty in the mutation estimation data, which arise from the sampling protocol of single molecules into real-time PCR wells in the Random Mutation Capture (RMC) assay. In this thesis, an alternative protocol was proposed based on a statistical modeling of the protocol. The statistical design of experiment was developed using linearized variance propagation analysis and Monte Carlo methods. The design of experiment revealed an optimal dilution factor of about 1.6 mutant DNA template per well quantifying *unamplified* wells, compared to 0.1 template per well of the conventional RMC assay quantifying *amplified* wells. Experimental validation and simulation results indicated that the optimized protocol produced a substantial improvement in measurement accuracy (reduction in the measurement variability) and 10-times improvement in the information offered per PCR well, i.e. the optimal protocol can achieve the same coefficient of variation using one-tenth the number of wells in the original assay.

Finally, the determination of "kinetic" rate constants of stochastic models, like those mentioned above, is non-trivial and is one of the most critical challenges faced in the development of such models. While there exist methods for deterministic models (e.g. for differential equations), the identification of kinetic parameters for stochastic models is not yet routine. In this thesis, several parameter estimation methods for CME models have been proposed that utilizes data from single cell measurements or cell population study (e.g. flow cytometry). The methods were developed based on maximum likelihood and probability density distance objective functions. Application to examples demonstrated

that these methods can provide accurate and robust estimates of the kinetic rate constants.

List of Figures

1.1	Mitochondrion: Energy producing organelle of Eukaryotic cells.	7
1.2	Oxidative Phosphorylation process and the Electron Transport Chain in the inner mitochondrial membrane.	8
1.3	Free radical chemistry and cellular anti-oxidant defense system.	14
1.4	The Mammalian mitochondrial genome.	16
1.5	Mitochondrial ROS Vicious Cycle.	22
1.6	Technological challenges in quantifying somatic mtDNA point mutations.	26
1.7	Stochastic origin and clonal expansion dynamics of mtDNA mutation.	28
2.1	Synergism between the natural system and formal system.	34
2.2	Taxonomy of biological modeling in systems biology.	38
3.1	Stochastic mouse mtDNA turnover model.	64
3.2	Results of choice of mtDNA replication model.	76
3.3	Stochastic determinants of age-dependent dynamics in the observed mtDNA point mutation frequency.	79
3.4	Point mutation distribution in cells of heart tissue from different <i>in silico</i> mice.	80
3.5	Stochastic evolution of mtDNA states.	81
3.6	Average mtDNA point mutation frequencies in WT and POLG mutator mice.	84
3.7	Stochastic determinants of age-dependent point mutation dynamics in mutator mice.	85
3.8	Comparison of mutation frequency (experimental and simulation data) for different mtDNA turnover rates.	88
3.9	Mitochondrial DNA point mutation accumulation during mouse development.	90
3.10	Simulations with exclusion of developmental phase.	91
3.11	Average age-dependent mtDNA point mutation accumulation in wild type and POLG mutator mouse models.	93
3.12	Mitochondrial DNA point mutation burden under an elevated oxidative burden assumption.	93
4.1	Stochastic model of mtDNA turnover dynamics in a mouse cardiomyocytes.	100
4.2	Simulation results of the fold increase at different replicative advantage factors and at different initial load.	106
4.3	Parametric perturbation analysis on mouse heart ($n = 3$).	107
4.4	Comparison of the results from different hypothesis testing on the mechanisms of mtDNA mutant clonal expansion process.	110
4.5	Parametric perturbation analysis on mouse heart ($n = 3$), considering model used for hypothesis D.	111

4.6	Mitochondrial DNA deletion accumulation in mouse heart tissue ($n = 3$). Influence of animal development in the mtDNA mutation heterogeneity observed in the mouse heart tissue. Simulations were performed using the model and parameters of hypothesis D.	112
4.7	Comparison of simulation outcome of using different clonal expansion hypothesis for capturing the dynamics of mtDNA mutation accumulation in mouse heart ($n=3$).	114
4.8	Comparison of simulation outcome on the state space distribution of mtDNA, obtained using different clonal expansion hypothesis for capturing the dynamics of mtDNA mutation accumulation in mouse heart ($n=3$).	114
4.9	Parametric perturbation analysis on mouse heart ($n = 3$). State space probability density function of Total mtDNA count ($W + M$) and mutation burden (F_M) in mouse cardiomyocytes of the complete heart tissue simulations performed using the model and parameters of hypothesis D.	116
4.10	Influence of mutant mtDNA replicative advantage on mtDNA mutation clonal expansion dynamics. Simulations performed using the model and parameters of hypothesis D.	118
4.11	<i>in silico</i> perturbation analysis to understand the influence of different parametric variation on the origin and dynamics of mtDNA mutation clusters in the murine tissue. Simulations were performed using the model and parameters of hypothesis D.	120
4.12	Results on the origin and accumulation of <i>de novo</i> mutation in clonally expanded mutant cells during mouse life span. Simulations were performed using the model and parameters of hypothesis D.	122
4.13	Parametric perturbation analysis on mouse heart ($n = 3$). Simulations were performed using the model and parameters of hypothesis D.	123
5.1	Details of the hybrid (Stochastic-Deterministic) model of mtDNA turnover process in mouse skeletal muscle tissue.	130
5.2	Stochastic model of mtDNA turnover dynamics in a mouse skeletal muscle progenitor cells (myoblast cells).	131
5.3	Progression of the COX deficiency in the skeletal muscle fiber with age, from the onset to the complete dysfunction of muscle fiber.	134
5.4	Validation of the simulation results with the data from the histochemical studies on rat skeletal muscle tissue (simulation statistics is based on five mouse tissue($n = 5$)).	137
5.5	Simulation results on the number of fibers lost as a function of age.	137
5.6	Effect of changes in effective mitochondrial diffusivity on the progression of COX deficient regions in muscle fiber (simulation statistics is based on five mouse tissue($n = 5$)).	140
5.7	Progression of the COX deficiency in the skeletal muscle fiber with age — Significance of animal development.	141
5.8	Results of different parametric perturbation analysis on COX deficient attributes of skeletal muscle (simulation statistics is based on five mouse tissue($n = 5$)).	144
5.9	Results of different parametric analysis related to beneficial effects of caloric restriction (CR) (simulation statistics is based on five mouse tissue($n = 5$)).	148
5.10	Simulation and validation results of muscle fiber subjected to caloric restriction (simulation statistics is based on five mouse tissue($n = 5$)).	151

5.11	Results of mutation burden in tissue arising from different short-term dietary regimen transfer (simulation statistics is based on five mouse tissue($n = 5$)).	153
6.1	Statistical aspects associated with the RMC assay.	158
6.2	Results of Optimized RMC assay considering the approximation of sampling process to the Poisson process.	165
6.3	Variation of the Coefficient of variation (CV) obtained using Monte Carlo analysis.	165
6.4	Influence of the experimental errors (type - I and II errors) on the performance of both the optimized and original RMC protocol assuming Poisson random sampling.	166
6.5	Variability in the point mutation frequency measurement data in mouse heart tissues.	168
7.1	mRNA Dynamics Model in <i>E. coli</i>	178
7.2	Normalized objective function contours of the ML and DFD methods in the <i>E. coli</i> RNA dynamics model.	181
7.3	Gene Expression Model for the Preferential Galactose Uptake in Yeast Cells.	183
7.4	Validation of the parameter estimates obtained from the optimization of the complete yEGFP model.	185
7.5	Stochastic dynamics of synthetic gene toggle switch engineered in <i>E. coli</i>	188
7.6	Effect of the finite sampling noise on the parameter estimation of <i>E. coli</i> RNA dynamics model.	193

List of Tables

1.1	Comparison of mammalian nuclear and mitochondrial genome.	17
3.1	Model parameters used in the simulations of the <i>in silico</i> wild-type mice. . .	68
3.2	Model parameters used in the simulations of the <i>in silico</i> POLG heterozygous mice's ($POLG^{+/mut}$) mice	71
3.3	Model parameters used in the simulations of the <i>in silico</i> POLG heterozygous mice's ($POLG^{mut/mut}$) mice.	77
3.4	Turnover rates of some mammalian mitochondrial proteins.	87
4.1	Model parameters of the stochastic mtDNA turnover process in the cardiomyocytes of <i>in silico</i> mouse model.	98
4.2	Details of different hypotheses proposed to explain the clonal expansion of pathogenic mtDNA mutations.	110
5.1	Morphometric details of the <i>Vastus lateralis</i> murine skeletal muscle fibers, used for hybrid stochastic-deterministic modeling.	132
6.1	Estimation of the false positives and false negatives frequencies	162
6.2	Pseudo-code for simulating the sampling protocol used in the original RMC assay.	163
6.3	Pseudo-code for simulating the sampling protocol used in the optimized RMC assay, developed in this work.	164
6.4	Mock RMC experimental results using the optimized and original protocols.	166
7.1	SSA formulation of the RNA dynamics model.	179
7.2	Parameter estimation of <i>E.coli</i> RNA dynamics model.	180
7.3	SSA formulation of the <i>S. cerevisiae</i> reduced Galactose uptake model.	182
7.4	SSA formulation of the <i>S. cerevisiae</i> complete Galactose uptake model.	184
7.5	Parameter estimation of <i>S. cerevisiae</i> reduced yEGFP model.	184
7.6	Parameter estimation results of <i>S. cerevisiae</i> full yEGFP model.	185
7.7	Parameter estimation of <i>S. cerevisiae</i> reduced yEGFP model using maximum probability distance measure.	186
7.8	Parameter estimation results of <i>S. cerevisiae</i> full yEGFP model using maximum probability distance measure.	187
7.9	Parameter estimation of synthetic toggle switch in <i>E. coli</i>	188
7.10	Parameter estimation of Schlögl model.	190

Keywords

Computational biology
Mitochondrial DNA
Aging
mtDNA mutations
Stochastic drift
Clonal mtDNA expansion

Abbreviations & Nomenclature

Abbreviations

Abbreviations	Description
8OHdG	8-hydroxydeoxyguanosine
AL	<i>ad libitum</i>
ATP	Adenosine triphosphate
bps	Base pairs
CDF	Cumulative distribution function
CME	Chemical master equation
COX	Cytochrome C oxidase
CR	Calorie restriction
DE	Differential evolution
DFD	Density function distance
DOE	Design of experiments
dpc	Days post coitum
ETC	Electron transport chain
ETS	Electron transfer system
FRTA	Free radical theory of aging
GFP, gfp	Green fluorescent protein
MC	Monte Carlo
mFRTA	Mitochondrial free radical theory of aging

ML	Maximum likelihood
mtDNA	Mitochondrial DNA
nDNA	Nuclear DNA
ODE	Ordinary differential equation
OXPHOS	Oxidative phosphorylation
PCR	Polymerase chain reaction
PDF	Probability density function
POLG	Polymerase- γ enzyme
POLG ^{+/<i>mut</i>}	Heterozygous POLG transgenic mouse
POLG ^{<i>mut/mut</i>}	Homozygous POLG transgenic mouse
RMC	Random mutation capture
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RRF	Ragged red fiber phenotype
SDH	Succinic dehydrogenase enzyme
SOS	Survival of the slowest hypothesis
SOT	Survival of tiny
SSA	Stochastic simulation algorithm
TFAM	Mitochondrial transcription factor A

Nomenclature

Nomenclature	Description
W	Wild type mtDNA molecules
M	Mutant mtDNA molecules
k	Reaction rate constant
k_m	Reaction rate constant for <i>de novo</i> mtDNA mutation
k_d	Reaction rate constant for mtDNA degradation
k_R	mtDNA replication rate

$P(\cdot), f(\cdot)$	Probability density function
ν_j	Change in states stoichiometric vector due to the j^{th} reaction
τ	Time leap to the next reaction; Gillespie's algorithm
j	Next reaction index; Gillespie's algorithm
ν_{max}	Maximum rate of mtDNA replication
K_H	Hill Constant
n_H, n	Hill Coefficient
F_W	Wild-type mtDNA proportion; $W / (W + M)$
F_M	Mutant mtDNA proportion; $M / (W + M)$
$t_{1/2}$	Half life of mtDNA turnover
D_{eff}	Effective diffusivity of mtDNA
p_0	Proportion of unamplified PCR wells in RMC assay
λ	Average mtDNA mutant template count per PCR well
n_{well}	Number of PCR wells per plate used in RMC protocol
CV	Coefficient of Variation
V	Variance
$L(\cdot)$	Likelihood function
$F(\cdot)$	Cumulative density function
s, S	Standard deviation

Superscripts

dev	Pertaining to developmental stage
PN, PM	Pertaining to post natal or post mitotic stage
$\hat{\cdot}$	unbiased estimator of the parameter

Subscripts

<i>ss</i>	Related to the steady state value
<i>WT</i>	Pertaining to wild-type mouse
<i>mut</i>	Pertaining to POLG mutator mouse models
<i>e</i>	Related to the experimental data
0	Pertaining to initial conditions

Other notations used in specific chapters are described *in situ*.

Introduction

1.1 Theories of aging

Biological aging may be described as a process of intrinsic, progressive and generalized physical deterioration that occurs over time, beginning at about the age of reproductive maturity. Aging is therefore characterized by declining ability to respond to stress and by increasing the homeostatic imbalance and incidence of pathology, and death remains the ultimate consequence of aging [1, 2]. Organismal aging cannot be characterized as disease, however, the incidence of several diseases increase exponentially with advancing age. Except in the case of prokaryotes, some protists and some multi-cellular eukaryotes, the rate of aging varies enormously among different species, with some species dying within days and several others surviving for more than a century [3, 4]. To distinguish species that age without undergoing physical deterioration, "senescence" is commonly preferred compared to aging.

The prime motivation of gerontological research, like any other biological research is to understand and appreciate the mechanisms that cause cellular pathology and eventually result in organismal death and further possibly develop interventions that can reverse the process of aging [2]. In recent decades, there has been enormous increase in aging related research [2, 5, 6]. This can be primarily attributed to: (i) extraordinary lengthening of average & maximum human life span, (ii) the consequent increase in the proportion of elderly in the population and, (iii) the related increases in the proportion of national health care expenditures.

Biogerontology is a vast and diversified field, which is littered with large number of aging theories. An attempt by Medvedev to rationalize the multiplicity of aging hypotheses resulted in a listing of more than 300 theories of aging [7]. Advances in experimental techniques to investigate the phenomenological complexities of the aging phenotypes, has helped clear the path towards unraveling the mechanisms of aging process [2]. Also, with recent development of different theoretical approaches, a significant simplification of the theoretical underpinnings of aging research has resulted. Nevertheless, the intrinsic complexity of aging remains a significant challenge in understanding how aging is caused. Several studies have indicated that the search for a single cause of aging, such as a single gene or the decline of the functionality of a key body system may be futile and aging may be considered as a complex multi-factorial process [8, 9]. Several processes may interact simultaneously and may operate at several levels of functional organization [10].

Some of the most widely used classification of aging theories are: *programmed* and *error* theories. According to the programmed theories, aging depends on biological clocks regulating the timetable of lifespan through different stages of organism's life. The regulation depends on genes modulating the signals that control different tissue systems responsible for cellular homeostasis. Greater details on the programmed and evolutionary theories of aging can be found elsewhere [2, 4, 11]. The error theories however, identify environmental insults to cellular systems (like, mitochondrial DNA damage, oxygen radicals accumulation, DNA cross linking, accumulation of other damaged macromolecules like proteins and lipids) to cause cellular morbidity. Since the error theories have recently gained much of the interest in the gerontological research domain [2, 6], the focus of the present dissertation is on aging arising as a consequence of accumulation of cellular macromolecule damage. Subsequent sections below discusses some of the important theories on molecular mechanisms of aging.

1.2 Molecular mechanisms of aging

Several research findings indicate that one of the most important mechanisms associated with aging are related to damage of cellular macromolecules. However, it should be noted that by the virtue of the complexity of aging process, none of the mechanisms discussed below may be sufficient to explain the process, independently. A close interplay between different mechanisms may be manifesting the organism and cellular level aging. In the text below, highlights of some of the major theories attributing the molecular mechanisms to aging process are described:

Somatic Mutation Theory: Age-dependent accumulation of somatic mutations and other types of damage to DNA cause an increase in cellular morbidity [2]. In this case, it has been suggested that capacity of DNA repair and the associated accumulation of the DNA mutations may be an important determinant of the aging rate at the cellular and organismal level. However, it is not yet clear whether the age-related increase in somatic mutations result from an increase in the damaging agents or decreased ability of senescent cells to repair DNA [12].

Telomere Loss Theory: In mitotic tissue, a progressive decline in cellular division capacity with age appears to be linked to the shortening of telomeres, which protects the chromosome ends [13]. This is known as the "end-replication problem" - the inability of normal DNA copy machinery to proceed copying to the end of the DNA strand. This occurs due to the absence of the enzyme *telomerase*, which is normally expressed only in germline cells (in testes and ovaries) and in certain adult stem cells. In the case of somatic tissues, it has been suggested that the absence of telomerase activity acts as an intrinsic cellular division check, perhaps to protect the cells against runaway cell divisions, as in cancers [14]. The telomere shortening in some cells limit its proliferative life span. The telomere loss and the subsequent onset of the replicative senescence have been proposed to contribute to cellular aging and other

age-associated diseases [15]. Cellular stress and inflammatory responses have an accelerated influence on the telomere shortening [16].

A significant challenge to the role of telomere shortening in the context of aging comes from the studies related to yeast and rodent. A mutant strain of yeast, EST-1 exhibits ever increasing telomere shortening. However for this stain, it has been observed that there is no immediate loss of viability [17]. Also, in the case of normal aging in yeast, it has been seen that aging process is not accompanied with shortening of telomeres. In certain rodents, which are not specifically long lived, have been found to have immensely long telomeres [17].

Altered Protein/Waste Accumulation Theory: Protein turnover is essential to preserve the cellular functionality by removing the damaged or redundant proteins from the cell. Accumulation of damaged proteins occur due to age-dependent impairment of protein turnover mechanism, which results in range of age-related disorders, including cataract, Alzheimer's and Parkinson's diseases [18]. Chaperones and Proteasomes are main caretakers in the maintenance of protein integrity. It has been found that their functionality also decreases with aging [19, 20].

Age related accumulation of lipofuscin in secondary lysosomes of postmitotic cells have been implicated to cause cellular aging [21]. Lipofuscin is derived from non-enzymatic glycosylation of long lived proteins [22] and DNA [23]. Inhibition of lysosomal enzymes has been found to accelerate the accumulation of lypofuscin in brain [24]. However, In the context of the relevance of lipofuscin accumulation in cellular aging, what is still unclear is, whether the accumulation is itself damaging to cell or simply the consequence of a more fundamental problem that is directly relevant to cellular aging.

Network Theory of aging: Early proliferation in different aging theories have resulted from the consideration of individual aspects of aging as discrete independent pro-

cesses, and thus can only provide incomplete explanation for the complex aging process. However, due to possibly large scale interactions and dependencies among different competing mechanisms, a reductionist approach may prove to be inefficient. This led to an initiative of considering "network theories" of aging, in which the contribution of various mechanisms are considered together in capturing the interactions and potential synergisms among different aging processes [6, 9]. Development of such model networks can provide significant insight into the difference between "upstream" mechanisms that sets a process in train and "end stage" mechanisms that dominate different cellular phenotypes at the end of its life [8]. For example, a gradual accumulation of mtDNA mutations occurring over several years may result in steady increase in production of ROS and a gradual decline in energy production. Although the buildup of mtDNA mutations initiates the process, what ultimately destroys the cell, is the threshold at which the homeostatic mechanism fails. Understanding these connections are likely to be important in developing effective interventions against age-related cellular deterioration.

Mitochondrial Theory of aging: An important connection between oxidative stress and aging, is the accumulation of mtDNA mutations with age [25]. Age related increase of different types of mutations have been found to be associated with different types of tissues like skeletal muscle [26], brain [27], heart [28] and gut tissue [29]. Cells in which the mtDNA mutations reach a certain critical threshold, are likely to suffer from impaired ATP production resulting in a decline of tissue bioenergenesis [25]. As hypothesized by Harman, free radicals produced in the course of cellular energy metabolism leads to oxidative damage of mitochondrial macromolecules [1, 30]. This is one of the factors that is responsible for aging and degenerative diseases according to Free Radical Theory of Aging (FRTA). Since the mitochondrial theory of aging is the major focus of the present thesis, greater details on mitochondrial components

and the consequence of mtDNA mutations on the mitochondrial respiratory process and the cellular aging/morbidity will be outlined in the subsequent sections.

1.3 Mitochondria

Mitochondria are cellular organelles that are primarily responsible for cellular energy synthesis (Figure 1.1). It is now generally accepted that mitochondria are endosymbionts which originated from purple bacteria around 1.5 billion years ago [31]. Most of the initial details of how the symbiotic process developed into the present day cellular system are largely unknown. The eubacteria is believed to have evolved into mitochondria, which are able to safely replicate within the bounds of its host cell [32]. As a remnant from this time, the symbiont has transferred most of its genetic functionality to the eukaryotic host. The remaining genetic information exists in the form of a circular DNA known as the mitochondrial DNA (mtDNA). The intricate compartmentalization of mitochondria is of great advantage to the cell due to greater energy production resulting from the larger surface area available for ATP production. Although cells retained their ability to produce energy independent of mitochondria (by means of glycolytic pathway), most of the cellular energy requirement is provided by the mitochondrial respiratory functions.

Figure 1.1 shows the internal organization of a mitochondrion. Mitochondria occupy a substantial portion of cellular cytoplasmic volume, as they are a significant contributor for the cellular energy requirements via the Oxidative Phosphorylation process (OXPHOS). Earlier description of mitochondria was stiff, elongated cylindrical organelles, resembling their predecessor bacteria. However, recent research have indicated that mitochondria are remarkably mobile and plastic organelles, constantly changing their shapes and frequently undergoing fusion and fission process with their neighboring members [34].

Mitochondrial morphology comprises of outer membrane, inner membrane and inner mitochondrial matrix (Figure 1.1). The Krebs cycle (or Tri carboxylic acid cycle (TCA))

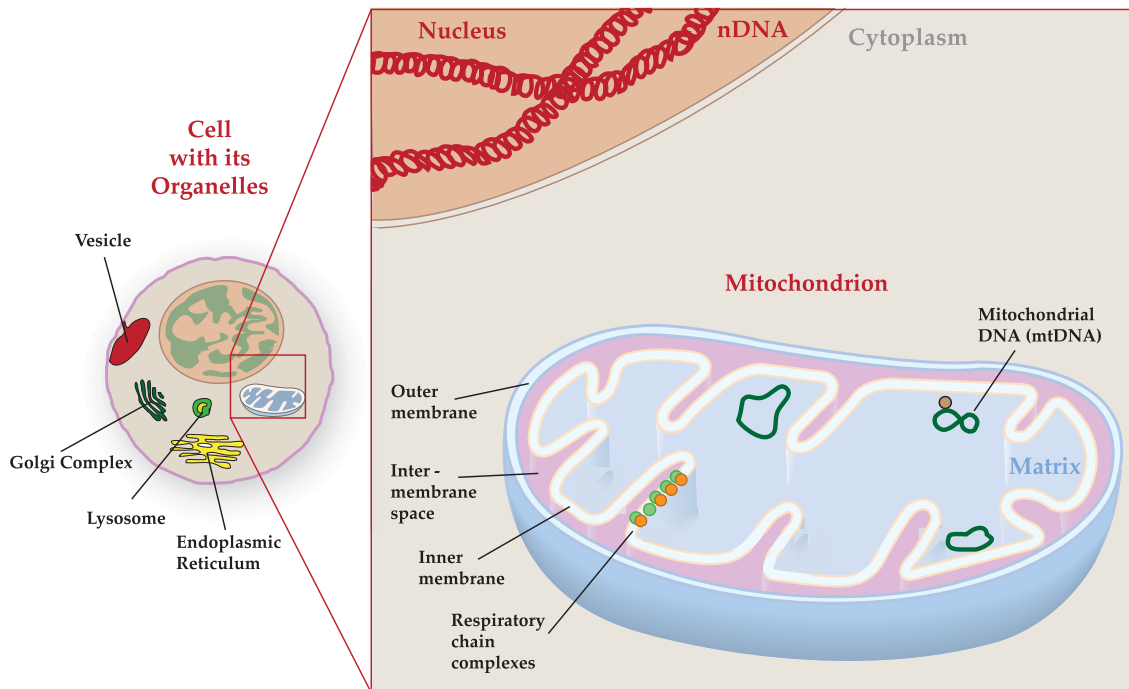


Figure 1.1: **Mitochondrion: Energy producing organelle of Eukaryotic cells.** The structure of a mitochondrion comprises of inner and outer phospholipid membranes, with the inner membrane having a complex convoluted structure known as the *cristae*. The space within the inner mitochondrial membrane is called matrix and the space in between the inner and outer mitochondrial membrane is know as the inter-membrane space. (illustration adapted and modified from [33]).

involved in the Adenosine tri phosphate (ATP) synthesis takes place in the mitochondrial matrix, and the respiratory chain complexes or the electron transport chain (ETC) rest within the inner mitochondrial membrane. The proteins that make up the ETC are synthesized from both the mitochondrial and nuclear DNA. The large surface area of the highly convoluted inner mitochondrial membrane provides a higher energy yield per unit volume of the organelle [32].

1.3.1 Oxidative phosphorylation process

In Eukaryotic cells, mitochondria generate energy by means of the OXPHOS. As discussed earlier, apart from the OXPHOS process, eukaryotic cells utilize anaerobic glycolysis for some of its energy requirements. However, glycolysis process is inefficient and generates only 2 ATP molecules for every glucose molecule. On the other hand, OXPHOS process generates about 36 ATP molecules for every glucose molecule that is consumed

during the cellular respiration [32, 35]. The glycolysis process takes place in the cytosol and yields reducing equivalents like NADH. During this process, pyruvate is produced as the end product and both pyruvate and NAD⁺ are transported into the mitochondrial matrix, where they participate in Krebs cycle. In the Krebs cycle, energy rich electrons are made available for NAD⁺ to form NADH (during Krebs cycle 3 molecules of NADH are produced per cycle). NADH then carry the electrons and passes them to the ETC or the electron transport system (ETS) complexes (Figure 1.2). Electrons are then transferred from donors to downstream acceptors such as oxygen by a series of redox reactions. During the transfer of electrons through the ETC complexes, the protons are pumped across the inner mitochondrial membrane to create membrane potential, thus resulting in the generation of ATP.

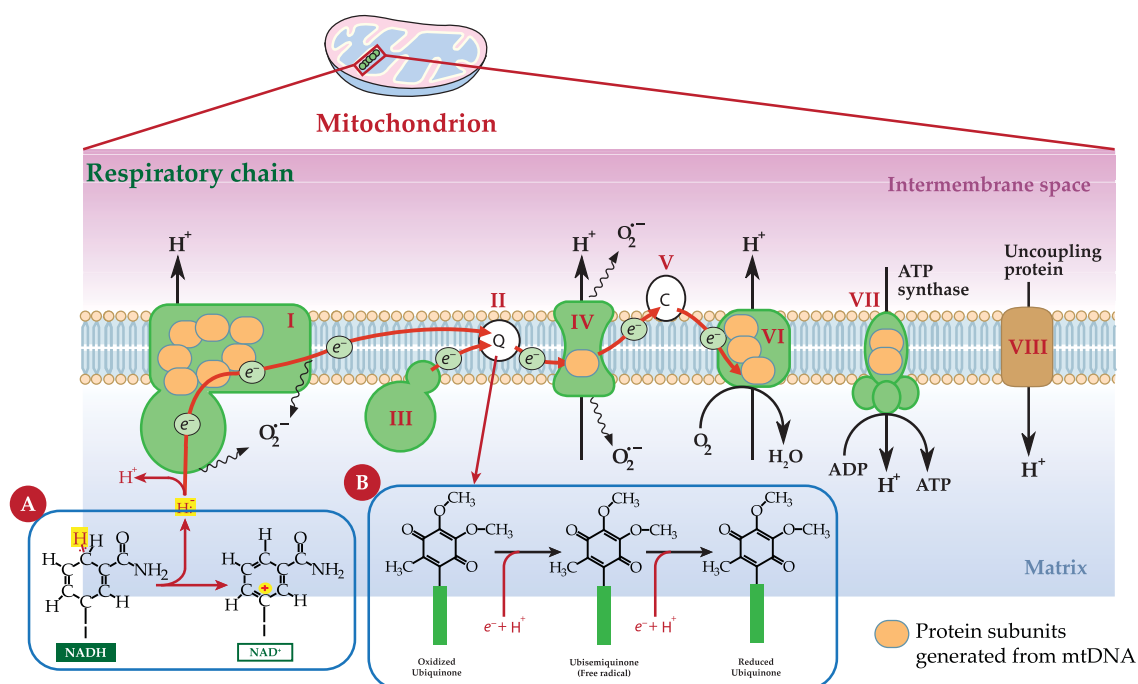


Figure 1.2: **Oxidative Phosphorylation process and the Electron Transport Chain in the inner mitochondrial membrane.** The respiratory chain complex that make the ETC are: I.) Complex I (NADH: dehydrogenase), II.) ubiquinone/ubiquinol complex, III.) Complex II (Succinate dehydrogenase), IV.) Complex III (ferricytochrome-c oxidoreductase), V.) cytochrome - c, VI.) Complex IV (cytochrome - c oxidase), VII.) Complex V (ATP synthase), and VII.) Uncoupling protein. A.) Electron donation by NADH. B.) Quinone electron transport (illustration adapted and modified from [33]).

1.3.2 Generation of Reactive Oxygen Species

The movement of electrons along the ETC subunits results in the generation of reactive intermediates. These reactive intermediates can inadvertently donate electrons to an oxygen atom resulting in the formation of free radicals or Reactive Oxygen Species (ROS) in the form of superoxide radicals (O_2^-). Measurable ROS productions have been reported in at least nine mammalian mitochondrial enzymes. These mitochondrial sources are: complex - I, complex - III, cytochrome b5 reductase, monoamine oxidases, dihydroorotate dehydrogenase, dehydrogenase of α -glycerophosphate, complex - II, aconitase and α -ketoglutarate dehydrogenase complex. Complex-I [36–38] and Complex-III [39–41] of the mitochondrial electron transport chain have been widely reported as the main source of mitochondrial ROS production and thus the mechanisms of ROS production in these two mitochondrial sites are briefly discussed in the subsequent paragraphs (Figure 1.2). Specifically, the electron transfer between coenzyme Q and Complex III is the main producer of ROS (Figure 1.2B). The extent of ROS production from the OXPHOS process depends on following two factors: a.) Steady state levels of reactive intermediates in the ETC, and b.) the amount of free oxygen available in the inner mitochondrial matrix (also known as the "oxygen tension"). The steady state level of reactive ETC intermediates depends on the energetic states of mitochondria and on the respiratory efficiency of the OXPHOS process. On the other hand, the propensity for a greater free radical production occurs when there is abundance of oxygen in the inner mitochondrial matrix and the extent to which ETC is reduced is high.

ROS generation at Complex I: Earlier studies indicated that isolated complex I can generate superoxide in the presence of NADH [42]. The reaction apparently requires tightly bound ubiquinone because it was inhibited by rotenone, an inhibitor that blocks electron transfer in close proximity to the ubiquinone binding site. Studies with both isolated complex I and submitochondrial particles demonstrated that a ROS producing site

is located between flavin and rotenone and flavin binding site [43–45]. There are also suggestions that ROS is most likely produced by one of the iron-sulfur centers and that there may be more than one superoxide producing in these regions [43, 46].

Three major experimental paradigms have been used for studying the ROS production attributed at complex I: a.) reverse electron transfer (RET), b.) rotenone induced ROS production, and c.) ROS production in normally functioning respiratory chain.

Reverse electron transfer experiments involving submitochondrial particles, became one of the first reported mitochondrial reactions supporting ROS production [44]. RET reactions in the respiratory chain allows electron to be transferred against the gradient of redox potential, from reduced coenzyme Q to NAD^+ , instead of oxygen. To proceed with this thermodynamically unfavorable process, it has to be coupled with utilization of energy of membrane potential.

Rotenone induced ROS production is observed in mitochondria oxidizing NAD^+ -linked substrates such as pyruvate and glutamate [47, 48]. In mitochondria, rotenone-induced ROS production requires very high degree of reduction of redox carriers upstream of rotenone binding site [49]. Rotenone induced ROS production can be mimicked by inhibiting the respiratory chain down stream of rotenone binding sites, e.g., by complex III inhibitor or cytochrome c depletion. These treatments induce the highly reduced state that is required for ROS production [49].

In the absence of complex I inhibitors, ROS production supported by NAD^+ -linked substrate is stimulated by high membrane potential [50]. These conditions also favor high degree of reduction of redox carriers proximal to the proton pump of complex I. The dependence of ROS production rate on the amplitude of membrane potential is not as steep as in the case of RET, consistent with a more thermodynamically favorable process [50].

ROS generation at Complex III: As discussed earlier, complex III oxidizes coenzyme Q (QH_2) using cytochrome c as an electron acceptor. The oxidation of QH_2 proceeds in a

set of reaction known as Q-cycle coupled with vectorial translocation of protons, thereby generating transmembrane potential. Complex III is capable of robust production of superoxide [39–41]. Generally, it is believed that an unstable semiquinone formed at the Q_0 site is largely responsible for the superoxide formation [39, 51, 52]. According to the classical Q-cycle hypothesis, inhibitors acting at the quinone reducing site, e.g. antimycin A, prevents the transfer of the second electron to Q_i site thereby causing accumulation of unstable semiquinone at Q_0 site [53], and thus generation of free radicals.

An increase in superoxide radical production can result, if the flow of electrons in the ETC is inhibited. For example, if one of the complexes in ETC becomes dysfunctional or has its conformational structure altered, then the flow of electron through ETC can be hindered and proportion of reactive intermediates in the ETC can increase [54]. The inhibition of flow of electron through the ETC can be caused due to the presence of mutated mtDNA. Damaged proteins encoded by mutant mtDNA may be dysfunctional and can hinder the flow of electrons in the ETC, and subsequently increase the formation of superoxide radicals. Due to damaging effects of ROS, nature has evolved an efficient defense system against their actions. Figure 1.3A, describes some of the mechanisms involved in reducing free radicals.

Oxygen is an essential component of life, but in the form of ROS, they are damaging factors to cellular components. There are several pathways by which oxygen can accept electrons *in vivo* as illustrated in Figure 1.3A. Amongst the different molecular states of oxygen radicals (Figure 1.3A), the hydroxyl radicals are most mutagenic and can readily react with any cellular biomolecules. An age-related increase in ROS-induced oxidative damage to DNA, lipids, and proteins have been well documented [55, 56]. Figure 1.3B illustrates an instance of mutagenic interaction of hydroxyl radical with DNA nucleotide resulting in oxidized DNA damage. Oxidized DNA damage can further result in base substitution during DNA replication resulting in mtDNA point substitution mutations.

Furthermore, progressive increase in oxidative damage in mtDNA, and the consequential increase in the genetic repair process with age have been found to cause DNA strand break and subsequent formation of somatic mtDNA rearrangements (like partial deletions and partial duplications) [57–60].

Mitochondria contribute about 20-30% of the steady state level of H_2O_2 in the cellular cytosol [54]. The superoxide radical that is produced in the mitochondria cannot cross biological membrane except in the protonated form, which may constitute a very minor fraction of the superoxide being generated during the cellular respiration [61]. The superoxide radical present in the inter membrane space can diffuse into the cytosol through the pores in the OMM, notably through the voltage-dependent anion channel (VDAC) [54].

Since mitochondria forms the major ROS produce in the cell, there is constant need for protection of cellular macromolecules from the toxic action of these species. Such defense is provided by several low-molecular-weight antioxidants, as well as several enzymatic defense systems (Figure 1.3A). Particularly, ubiquinone and vitamin E have been found to play a very important role in mitochondrial antioxidant defense system. Most important of the mitochondrial antioxidant protection is the tri-peptide glutathione, GSH (L- γ -glutamyl-L-cysteinylglycine), and multiple GSH-linked enzymatic defense systems (Figure 1.3A). Among the GSH-linked enzymes involved with mitochondrial antioxidant defense system are the Glutathione peroxidases (Gpx) 1 and 4. Glutathione peroxidases catalyze the reduction of H_2O_2 and various hydroperoxidase, with GSH as the electron donor. Gpx1 is the major isoform and is localized predominantly in the cytosol, but a small fraction is also present in the mitochondrial matrix. In contrast, Gpx4 is membrane associated, with a fraction present in the mitochondrial inter-membrane space [62]. Gpx4 reduces hydrogen peroxide groups on phospholipids, lipoproteins and cholestrestrol-esters. Gpx4 is considered to be the primary enzymatic defense against oxidative damage to cellular membranes [63]. A recent addition to the member of the family of GSH-linked mi-

tochondrial redox enzymes is the glutaredoxin 2 (Grx2), which has been thought to be present in both mitochondria and cytosol [64].

The mitochondrial thioredoxin system, which comprises of thioredoxin (Trx2) and thioredoxin reductase 2 (TrxR2), is another potential source of disulfide reductase activity required for maintenance of mitochondrial proteins in their reduced state. Thioredoxins catalyze reduction of proteins at much higher rates compared to Grxs [65]. The thioredoxin system can also interact with the peroxiredoxins (Prx), which constitute a novel family of thiol-specific peroxidases that rely on Trx as hydrogen donor for the reduction of hydrogen peroxide and lipid hydroperoxides [66]. One of the Prx isoform, Prx III is exclusively found in the mitochondrial matrix [67]. Upon interaction with H_2O_2 , the redox-sensitive site Cys residue of the Prx III homodimer is oxidized to cys-SOH, which then reacts with the neighboring Cys-SH of the subunit to form the intermolecular disulfide. The disulfide is reduced specifically by Trx2, which is subsequently regenerated by TrxR2 at the expense of NADPH [68]. Despite the presence of such cellular ROS defense system, some quantity of ROS inevitably escape and cause mutagenic damages to important cellular macromolecules such as DNA.

Free radical generation in mitochondria is also regulated by another process arising as a consequence of proton conductance across inner-membrane of mitochondria. OXPHOS process is thought to be incompletely coupled since the proton leak across the inner mitochondrial membrane can occur independent of the ATP synthesis [70]. Two respiration pathways exist: a.) consumption of oxygen in the presence of oxidizable substrate and in the presence of ADP to allow ATP production (state 3 respiration) and, b.) slow respiration in the absence of ADP (state 4 respiration) [70, 71]. The uncoupling protein of the ETC facilitates the proton cycling process. State 4 respiration increases disproportionately as the proton gradient increases, which can be entirely explained by the proton leak across the mitochondrial inner membrane [71, 72].

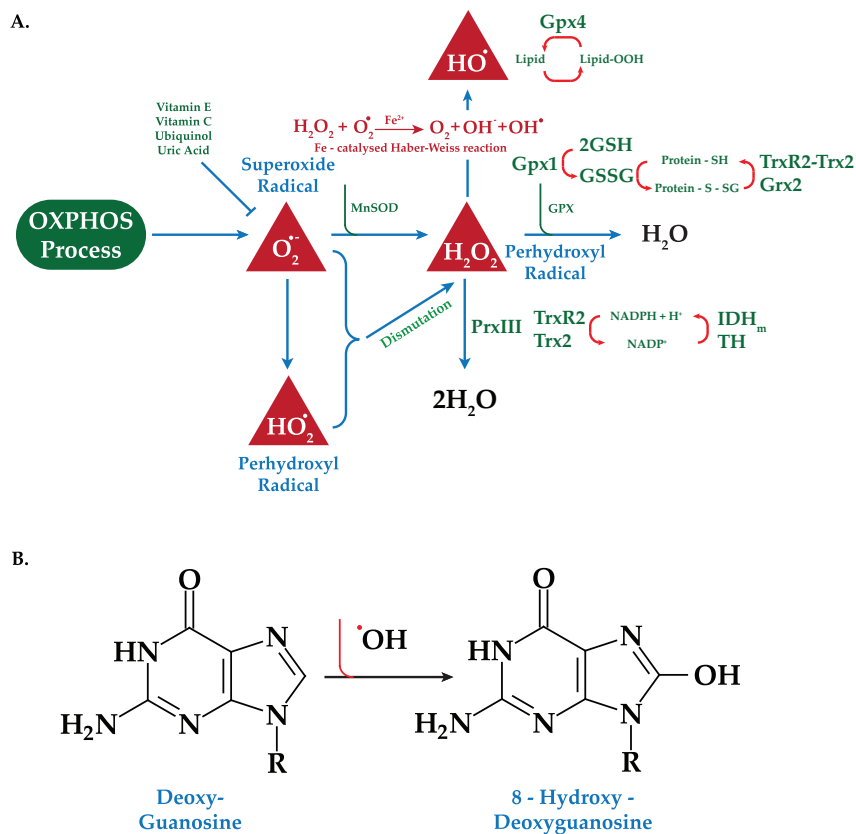


Figure 1.3: **Free radical chemistry and cellular anti-oxidant defense system.** A.) Cellular defense system against the mutagenic ROS actions (illustration adapted and modified from [69]). Figure also illustrates the radical chemistry associated with the interaction of oxygen with free electrons from the ETC to form different free radical species (red triangles). B.) 8-hydroxy-deoxyguanosine (8OHdG) lesion formation in the presence of hydroxyl radical. In event of DNA replication, oxidized DNA nucleotide can mispair, resulting in the formation of point mutation. Predominant mispair that can occur is the GC to AT transitions.

Proton leak is relatively high in perfused rat muscle. It comprises of 35% [73] and 50% [74] of the basal respiration rate and the "futile" proton leak comes at a tremendous energetic cost: roughly about 20-25% of rat's basal metabolic rate [73]. Its persistence throughout evolution, therefore implies a conferred advantage significant enough to necessitate mitochondrial inefficiency.

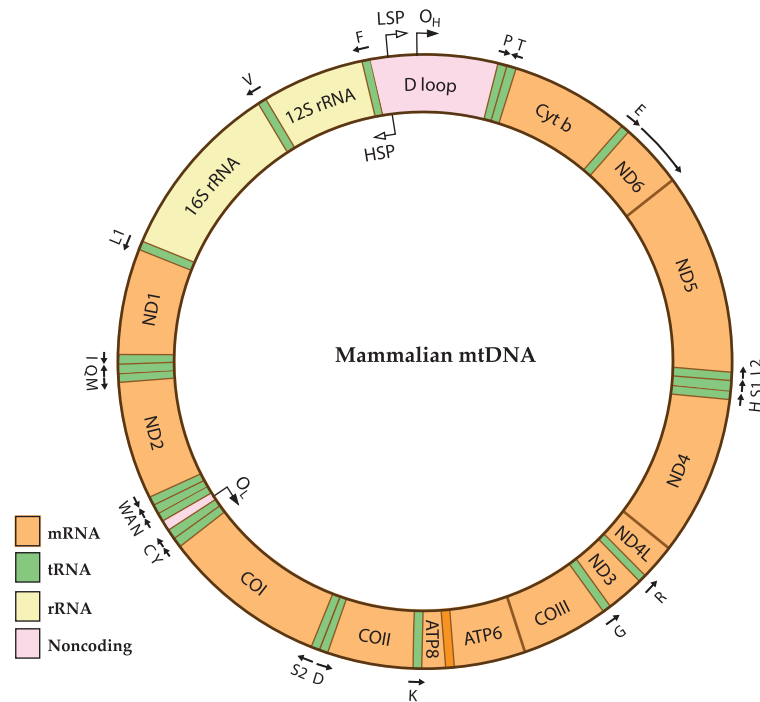
The physiological benefit of proton leak arises due to incompletely coupled OXPHOS, thus permitting the adjustment of energy metabolism to regulate metabolic homeostasis and maintain body functions. In certain specific cell types, the proton leak includes thermogenesis, maintaining carbon flux despite low ATP demand [75]. A general theory explaining the ubiquity of the proton leak process is the "uncoupling to survive" hypothesis

[70]. Since mitochondrial superoxide production is steeply dependent on the proton gradient [76, 77], the uncoupling to survive hypothesis proposes that the existence of proton leak pathway in the mitochondrial respiration process is to minimize the oxidative ROS production by modulating the proton gradient across the mitochondrial inner membrane [70]. Proton leaks can be broadly classified into: a.) constitutive, basal proton conductance and b.) regulated, inducible proton conductance catalyzed by the uncoupling proteins (UCPs).

The mechanism responsible for the basal proton leak is not fully understood, but it is believed that majority of basal proton conductance, is attributed to the abundance of the adenine nucleotide translocase (ANT) [78]. In the case of regulated proton conductance, three proteins in the mitochondrial respiratory chain have been found to modulate the proton leak: UCP1, UCP2 and UCP3. Studies on bioenergetics and inner membrane composition of brown adipose tissue mitochondria led to the discovery of UCP1 [71, 79], which basically dissipates the proton gradient in mammals during non-shivering thermogenesis [80]. Aside from the thermogenesis regulation of the UCP1, it also possesses additional physiological roles of controlling body weight [81] and modulating oxidative damage [82, 83].

UCP2 mRNA expression is ubiquitous in vertebrates, and proteins can be detected in spleen, kidney, thymus, pancreas, central nervous system and macrophages [84, 85]. Expression of UCP3 is confined to skeletal muscle, brown adipose tissue and heart [86, 87]. Plants [88, 89] and birds [90, 91] also express these UCP orthologs. It is well established that UCP2 and UCP3 do not catalyze adaptive, non-shivering thermogenesis [78]. There is evidence that UCP2 [92, 93] and UCP3 [94] are protective against oxidative damage. These UCPs have been suggested to provide local feedback in response to ROS damage and thus lowering the proton gradient across the inner mitochondrial membrane [70]."

Figure 1.4: Mammalian mitochondrial genome. The mitochondrial genome contains genes for encoding 12S and 16S rRNA, the subunits of Complex I (ND1, ND2, ND3, ND4, ND4L, ND5 and ND6), Complex III (Cyt b), Complex IV (COX I, COX II and COX III) and Complex V (A6 and A8) (also refer the [Figure 1.3](#)), and 22 tRNAs (one letter amino acid nomenclature in the figure). The origin of heavy-strand (O_H), the light-strand (O_L) and the promoters for the initiation of transcription from the heavy-strand (HSP) and the light-strand (LSP), are illustrated in the figure by using arrows (illustration adapted and modified from [33]).



1.3.3 Mitochondrial genetics

Each mitochondrion in mammalian cells contains about 2-10 molecules of mtDNA. Several mtDNA molecules are bound to each other and tethered to inner mitochondrial membrane by means of a nucleo-protein complex referred to as nucleoids [95]. In the last decade progress has been made in identifying the composition of nucleoids. Several studies showed that yeast nucleoids include proteins that bind DNA and contain replication and transcription factors such as the mitochondrial transcription factor A (TFAM), helicase TWINKLE, polymerase- γ and mitochondrial single stranded binding proteins (mtSSB) [96, 97]. Other structural proteins are also found in nucleoids [98].

Different types of somatic and germ line cells contain about 10^3 to 10^5 mtDNA molecules per cell [99]. In mammals, mtDNA are maternally inherited because the cells obtain their mtDNA exclusively from oocytes [99]. Human mtDNA is a plasmid-like closed circular double-stranded DNA molecule containing 16,569 nucleosides. Mitochondrial DNA encodes for 12S small and 16S larger rRNAs, 22 tRNAs and 13 polypeptides all of which are involved in the mitochondrial protein synthesis [32].

[Figure 1.4](#) illustrates the details of mammalian mtDNA. The mtDNA are very com-

pact molecule with no introns and intragenic sequences (Table 1.1). The compactness of mtDNA is evident from the overlap between some of the genetic locus and that some of the gene sequence lack termination codons [100]. Most of the genes are encoded by heavy strand of mtDNA (2 rRNAs, 14 tRNAs and 12 polypeptides), while light mitochondrial strand contains genes for one polypeptide and 8 tRNAs. Other than these mitochondrially synthesized components, most of the mitochondrial proteome necessary for the synthesis of ETC complex are generated by the nuclear gene synthesis (≈ 100 proteins) [99]. Furthermore, nuclear gene synthesis also contributes other proteins (≈ 1500 polypeptides) which are necessary for general functioning of mitochondria.

Table 1.1: Comparison between mammalian nuclear and mitochondrial genome. (Table adapted from [101])

Characteristics	Nuclear Genome	Mitochondrial Genome
Size	$\approx 3.3 \times 10^9$ bps	16,569 bps
DNA copy number per cell	23 in haploid; 46 in diploid	Several thousand copies per cell
Number of encoded genes	$\approx 20,000 - 30,000$	37 (13 polypeptide, 22 tRNAs and 2 rRNAs)
Gene density	≈ 1 per 40,000 bps	≈ 1 per 450 bps
Introns	Frequently found in most genes	Absent
Percentage of coding DNA	$\approx 3\%$	$\approx 93\%$
Associated proteins	Nucleosome-associated histone proteins and non-histone proteins	No histones; but associated with several proteins to form nucleoids
Mode of inheritance	Mendelian inheritance	Maternal inheritance

Mitochondrial DNA mutations are primarily generated during mtDNA replication process [102]. A detailed knowledge of mtDNA replication process is lacking, but essential to have a greater understanding of the accumulation and dynamics of mtDNA mutations. Presently, there are two models of mtDNA replication process. In accordance to the widely accepted asymmetric mtDNA replication model, it is generally believed that the mtDNA replication is a bidirectional process, where the mtDNA replication occurring in both the directions, from two different initiation sites of mtDNA (O_H and O_L in Figure 1.4) [97]. The replication initiates at the heavy stand origin (O_H) and proceeds in the clock-wise direction to about two-third of the mtDNA length, where the light strand origin exists (O_L).

Then, the replication of the light strand starts and proceeds in the anti-clockwise direction, until two copies of mtDNA are produced. An alternative mtDNA replication model based on the findings from a two dimensional agarose gel electrophoresis of mtDNA replication intermediates, indicated a strand coupled uni-directional replication from the heavy strand replication origin (O_H) [103]. In both these replication models, the transient single stranded conformation of mtDNA molecule makes the molecule more prone to mutagenic events [104, 105].

1.3.4 Mitochondrial DNA mutagenesis

Damages to mtDNA have been implicated in aging and other mitochondrial degenerative diseases, especially in the energetically active post-mitotic cells [27, 106, 107]. Mammalian mtDNA comprises of about 1-3% of the overall genetic material in a cell (Table 1.1) [107]. Several lines of evidence indicate that the contribution of mtDNA mutations towards cellular physiology is much more significant than what would be expected from its size & content [25, 108]. This arises due to the following reasons: (i) presumably higher mutations in mtDNA than nDNA [30], (ii) Functional compactness of mtDNA molecules (mtDNA has high density of coding sequence; Table 1.1), (iii) mtDNA encodes for some critical polypeptides that are important for cellular OXPHOS process, and (iv) the defects in ETC complexes caused by proteins generated from mutated mtDNA can affect the performance of cellular energetics [25, 107, 108]. Mitochondrial mutations can generally be classified into two types: (a) point mutations and (b) large-scale mtDNA rearrangements. The large scale mtDNA rearrangements can further be classified into partial deletions and partial duplications. It is found that mtDNA deletions are more commonly associated with aging and mtDNA associated disease phenotypes than point mutations [109–111].

Point mutations involve a single base nucleotide replacement. Commonly quantified biomarker of oxidative damage is 8OHdG, which can further lead to a point substitution

upon DNA replication process. Such point substitutions play an important role in the DNA mutagenesis [112, 113], and has been observed to increase with age in tissues of different organisms [114, 115]. Mitochondrial point mutations are predominantly of GC to AT, which is consistent with the mutation spectrum observed with the oxidative lesions [102, 116].

Deletion mutations occur when a portion of DNA template is truncated. A majority of mitochondrial deletions are located between the origins of replication (O_H and O_L), also known as the major arc. These deletions may be caused by polymerase slippage, homologous recombination, double strand break or inefficient repair mechanisms [58]. One of the most commonly found deletions that is associated with human mitochondrial myopathies is the nt4977 deletion, also known as the 'common deletion'. This deletion specifically disrupts the functions of mitochondrial complexes such as Complex I, IV and V [107, 117]. Other types of deletions have been found to be associated with aging skeletal, cardiac and neural tissues [26, 118].

Mitochondrial DNA mutations affect several different tissue types and the clinical manifestations of these diseases are variegated [25, 108]. In these mitochondrial disease phenotypes, pathogenic mtDNA mutation can severely impair the function of post mitotic tissues, which have higher energy demands. The extent and onset of pathology in these organs are variable, depending on the mutation load in the cells of the tissue. Most of mtDNA pathology is generally congenital and several human mitochondrial disorders are linked to different types of mtDNA mutations (point mutations, deletions and mtDNA duplications) [119, 120]. Since the early discovery of the mitochondrial role in these syndromes [25, 121], several pathogenic mtDNA mutations have been found to be associated with mitochondrial myopathies [122]. A mitomap database listing > 290 pathogenic point mutations, deletions and insertions is available at <http://mitomap.org/> [123, 124]. Mitochondrial DNA deletions have been associated with several degenerative diseases

like chronic progressive external ophthalmoplegia (cPEO), Pearson's syndrome and Kearns-Sayre syndrome, Alzheimer's disease, Parkinson's disease, Huntington disease and several other neuro-muscular degenerative diseases [125].

Despite the medical advances in characterizing the mitochondria-associated genetic defects that lead to mitochondrial myopathies, the pathogenic mechanisms for such diseases are still not completely understood. In particular the relationship between the presence of pathogenic mtDNA mutation and the occurrence of specific clinical symptoms of disease phenotypes are still largely unclear. Functional studies of mutant mtDNA associated with overt mitochondrial disease phenotypes have shown that the biochemical defects arising from same respiratory chain complex can lead to very different clinical manifestations [123]. The population of mutants in a cell can vary significantly between different cells and between different tissues within the same individual. The intrinsic genetic variability in the origin and progression of respiratory deficiency is one of the explanations for the existence of such stochasticity. The variable segregation of mutant and wild-type mtDNA between the dividing cells during embryogenesis can also be used to argue for tissue specific pathologies.

1.4 Research motivation - accumulation of mtDNA mutations in aging

1.4.1 Mitochondrial Free Radical Theory of Aging

The free radical theory of aging (FRTA) or the oxidative stress theory of aging was originally proposed by Denham Harman in 1956, which postulates that age-related loss of physiological functions of tissues occur due to progressive accumulation of free radical damage of different cellular macromolecules. The accumulation of such damage eventually determines the life span of an organism [1]. Shortly after the discovery of mtDNA,

Harman modified his original theory to include the role of mitochondria in producing cellular oxidative stress and proposed that the mitochondria are both the source and target of ROS. In subsequent years, the mFRTA was further refined by Miquel and colleagues [126], who suggested that the accumulation of somatic mutations in mtDNA induced by oxidative damages is the main cause of aging and age-related mitochondrial degenerative diseases. The present version of mFRTA is based on the following five premises [127]:

- ★ Mitochondria produce ROS as by-product of OXPHOS.
- ★ Mitochondrial DNA are attached to the inner mitochondrial membrane, where cellular respiration takes place. Due to their proximity to the ROS production site, mtDNA are subjected to an elevated level of free radical attack.
- ★ Mitochondrial DNA, unlike nuclear DNA (nDNA), are not protected by structural proteins like histones, which offers protection to genomic material from free radical attacks.
- ★ Although mitochondria possess some DNA repair capabilities, they are primitive and less developed than nuclear DNA.
- ★ Mitochondrial DNA mutations can further elevate ROS production. This assumption suggests that a small amount of initial mtDNA functional damage can trigger an exponential increase in the ROS production, which further aggravates the mitochondrial damage, constituting a positive feedback mechanism, known as the ROS "vicious cycle" (Figure 1.5).

There exist several ongoing debates in this field, which relate to the issue of causality: are mtDNA mutations merely markers of biological aging, or do they cause the observed decline in physiological function that contribute to the aging process? Some of the assumptions of mFRTA have been proven invalid, while others are still intensely debated due to the contradicting experimental data. An extensive discussion of these debates and

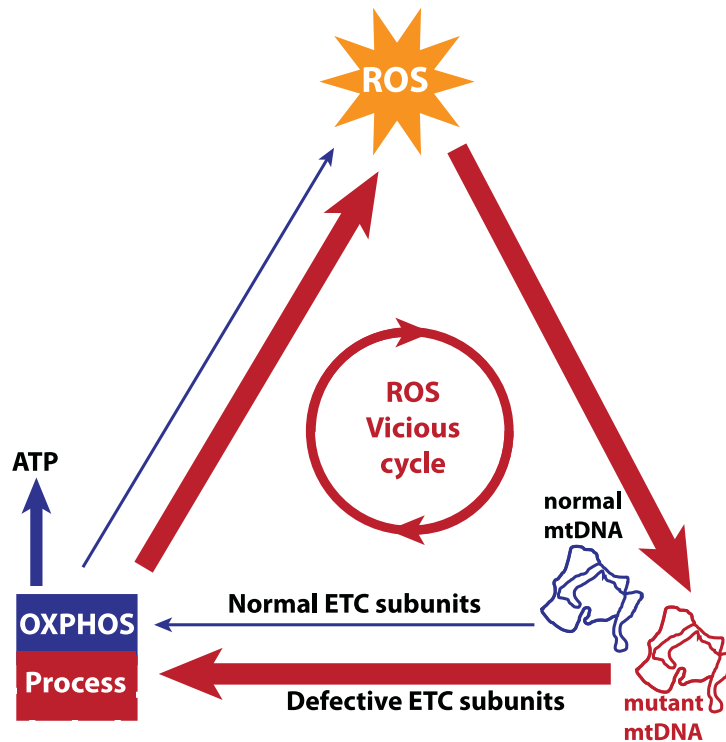


Figure 1.5: **Mitochondrial ROS Vicious Cycle.** A putative positive feedback mechanism between mtDNA and ROS is based on the hypothesis that ROS-induced damaged mtDNA produce defective components of the ETC, thereby increasing electron leakage in the OXPHOS process and ROS production. The vicious cycle is expected to give an exponential expansion of mtDNA mutations over time, which eventually causes the loss of mitochondrial function in generating ATP.

the related experimental evidence are summarized in recent reviews [107, 127]. One of the hypothesis which is important for the present dissertation is the existence of ROS vicious cycle. Earlier support for the existence of ROS vicious cycle was provided by experiments conducted in cultured human neurons, where the disruption of the ETC was done using RNA interference (RNAi). This experiment showed a significant increase in the oxidative DNA damage and ROS production [128]. A similar study done using *in vivo* chemical inhibition of complex I using rotenone also led to irreversible damage of mitochondria and resulted in an increased ROS production [129]. However, to investigate the functional relevance of *in vivo* point mutation accumulation, transgenic mouse models carrying error prone mtDNA polymerase (mutator mouse model) have been created [116, 130, 131]. These mice have been found to accumulate high mutation burden in most tissues even at a very young age, but interestingly, did not show any evidence of increased protein carbonyl levels, a biomarker of oxidative protein damage resulting from ROS induced dam-

age. Similarly, age dependent increase of mtDNA oxidative damage was not observed in these transgenic mice [132]. Thus, studies involving transgenic mouse models, which can be viewed as a direct test of mitochondrial theory of aging, has produced inconsistent results, thus challenging the core premise of mFRTA.

1.4.2 Mouse models - human aging phenotypes

Several of the earlier studies connecting mtDNA mutations and aging have been predominantly correlative, with unclear causality. Consequently, this resulted in persistent debates over the basic premises of mFRTA. Despite these controversies, the direct evidence of the role of mitochondria as both source and target of ROS mediated pathology, leading to mammalian aging has been provided by several transgenic mouse model studies [133].

One important progress in this direction is the generation of mutator mice with tissue specific disruption of nuclear gene encoding mitochondrial TFAM, and thus having reduced mtDNA expression. A TFAM knockout mice, specific to dopaminergic neurons were developed, which showed Parkinsonism phenotype by age of 20 weeks [134]. In another model, knock-in mouse were engineered to specifically express human catalase in peroxisomes, nucleus and mitochondria. Significant increase in the maximum and medium life-span (up-to 21%) was observed in the transgenic mice expressing catalase specifically in their mitochondria (MCAT mice), but neither in the nucleus nor in peroxisomes [135]. Furthermore, the MCAT mice have been shown to accumulate less than 50% of mtDNA point mutation in their heart cells compared to age-matched wild-type cohort, supporting the idea that ROS affects mtDNA mutation dynamics and the ROS mediated mtDNA damages affect murine life-span.

As mentioned earlier, mutator mouse models carrying proof-reading deficient version of mtDNA polymerase- γ (POLG), have been used to investigate mtDNA causality in aging. The first study considered tissue-specific mutator mice which only expressed the

proof-reading deficient form of POLG in heart [136, 137]. These mice had increased levels of mtDNA point mutations and deletion, and developed clinical signs of congestive heart failure as early as 4 weeks after birth. Following on this work, tissues of "whole body" mutator mouse models had 3 to 5 fold increase in mtDNA point mutations and mtDNA deletions [116, 130, 131]. The increased mtDNA mutation burden in tissues of POLG knock-out mice are associated with a number of aging phenotypes like hair loss, kyphosis, osteoporosis, reduced fertility and anemia. These mice had significantly reduced life-span compared to wild-type cohorts.

Pathogenic mutations in nuclear DNA encoding TWINKLE helicase has been identified in some human mitochondrial myopathies [138]. Transgenic mice expressing mutant TWINKLE isoforms modeled after these mutations displayed progressive localized mitochondrial respiratory deficiency, particularly in the musculature (skeletal and heart) and brain tissues. A mild myopathic syndrome was observed in these mouse models at about 1 year of age [139].

Another mouse model carrying mtDNA deletions has been created via a gene targeting protocol in mouse embryonic stem cells. These mice, referred to as "mito-mouse" [140], carry a 4,696 bps deletions in their mtDNA. This deletion removes six tRNA and seven structural genes from the mitochondrial genome. Mito-mice exhibited COX deficiency in muscle fibers when the deletion burden was higher than 85%. Mosaic respiratory deficits were noted predominantly in heart and kidney. These mice were also anemic and eventually died from renal failure by 200 days of age.

1.4.3 Mutation accumulation dynamics - significance of stochasticity

An important perspective towards understanding the role of mtDNA mutagenesis in aging relates to the mtDNA mutation dynamics. The mechanism by which mtDNA mutations accumulate over time is yet to be clearly elucidated. Comprehending the dynamics

of mtDNA mutation accumulation in somatic cells forms the primary thesis of the present dissertation. Two important observations discussed below motivated the development of a theoretical framework in this dissertation, for the better understanding of mtDNA mutation dynamics and its role in the aging process.

Uncertainty in age-dependent mtDNA mutation accumulation dynamics

Two recent experiments on mtDNA point mutation burden in mouse heart tissue by independent investigators have indicated contrasting dynamics of mitochondrial point mutation accumulation. Using standard polymerase chain reaction (PCR), cloning and sequencing techniques, Larsson's group demonstrated that the accumulation of somatic mtDNA point mutation in both wild-type and POLG mutator mouse heart tissue, to be approximately a linear function of age (Figure 1.6A). However, using a highly sensitive mutation quantification method, known as the Random Mutation Capture (RMC) assay, Loeb's group obtained point mutation burden data shown in Figure 1.6B. The authors subsequently suggested that the accumulation of mtDNA point mutations increase exponentially with age, consistent with the ROS vicious cycle theory.

The RMC assay developed by Bielas and Loeb provided a sensitive estimation of the mtDNA point mutation frequency [141]. However, as observed from the Figure 1.6B, protocol produces data that has significant amount of variability associated with them. The presence of such high variability in the measurement data, introduces a significant challenge in understanding the underlying mtDNA mutation accumulation dynamics. Thus, understanding the dynamics of mtDNA mutation accumulation from such noisy data requires a careful interpretation of different sources of noise, which can arise from both the intrinsic cellular stochasticity associated with the underlying mtDNA mutation process and stochasticity arising from the measurement protocol. The inherent stochastic nature of mtDNA point mutation process and challenges faced due to the quantification of extremely low mutation burden makes it imperative to optimize the experimental protocol

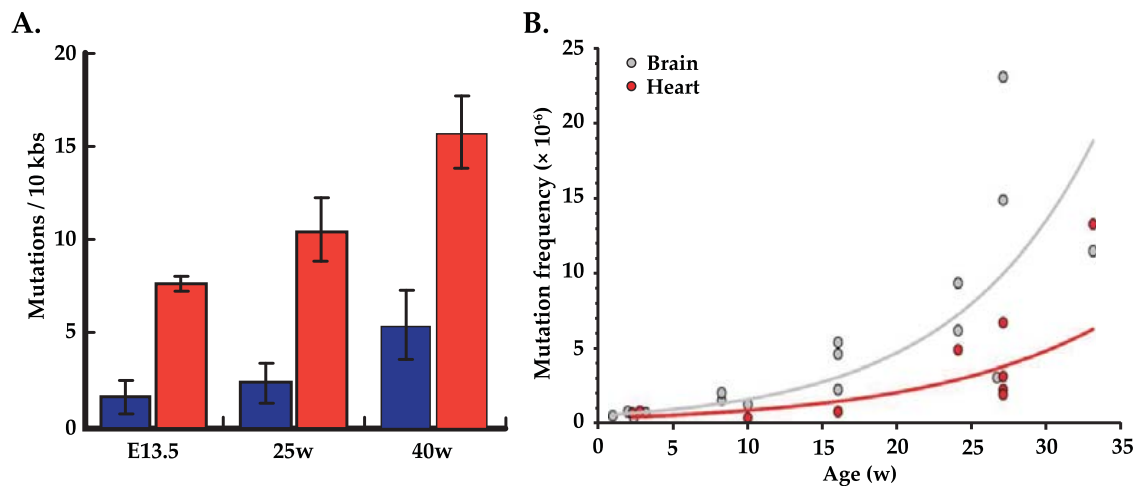


Figure 1.6: **Technological challenges in quantifying somatic mtDNA point mutations.** (A) Linear age-dependent accumulation of mtDNA point mutation as measured using the conventional PCR cloning method [132]. The plot shows mtDNA point mutation accumulation in mouse heart tissue. The data has been sampled from mouse at embryonic day of 13.5 days, and 25 weeks and 40 weeks post natal. B.) Age-dependent accumulation of mtDNA point mutation as measured by the RMC method [116], appears to be exponential. This plot illustrates the age-dependent (in weeks) accumulation of mtDNA point mutation accumulation in heart and brain tissue of mice. These plots are adapted from Trifunovic *et al.*, 2005 [132] (Panel A) and Vermulst *et al.*, 2007 [116] (Panel B), respectively.

and to perform critical analysis of experimental data.

The search for mtDNA mutations that accumulate with aging was initially inspired by the fact that certain type of mtDNA rearrangements and point mutations are profoundly deleterious, when present in proportions that is at appreciable levels compared to the wild-type mtDNA, and are commonly associated with overt diseases [25]. Interestingly, similar mutations are also present in healthy individuals, albeit at much lower levels than those seen in patients with mitochondrial diseases [27, 110, 142]. Existence of such low mutation burden in tissue of normal aging individuals presents an important challenge in understanding the role of these mtDNA mutations in cellular aging. Despite the low abundance of pathogenic mutations in aging tissue (the overall mutation burden is generally below any functionally relevant level), the mutation load in tissues were found to be consistently increasing with age. Another intriguing aspect was how such low abundance of mtDNA mutations ever becomes relevant in the context of tissue aging, considering cells contain several hundred to thousands of copies of mtDNA.

Clonal expansion of mtDNA mutation — role of cellular stochasticity

Single cell analyses have demonstrated the mosaicity of mtDNA mutation burden in different cells of the same tissue, indicating that mutant mtDNA often coexists with the normal mtDNA, a condition referred to as heteroplasmy, and that the level of mtDNA mutations and their functional significance often varied dramatically among different cells and consequently among different tissues [143, 144]. It was also discovered that pathogenic mtDNA mutations can only cause respiratory chain dysfunction, when they are present above a certain threshold level referred to as 'phenotypic threshold' [145], which is > 60% for single large mtDNA deletion [146], and >90% for certain point mutation in tRNA genes [147] (also see Figure 1.7). In these cells, once pathological mutation(s) has occurred, it initially has low functional importance due to inherent redundancy of mtDNA in a cell. However, the same mutation(s) can subsequently reach functionally relevant levels due to still poorly understood mechanism known as the "clonal expansion process", where the progenies of a single mutant mtDNA molecule out-populate wild-type mtDNA and eventually comprise a large portion of cellular mtDNA pool, as shown in Figure 1.7. Different types of tissue analysis demonstrated mosaic respiratory chain deficiency due to clonal expansion process, in heart [111], hippocampal neuron [148], choroids plexus [149], mid-brain dopaminergic neurons [27], and colon [29].

Studies on skeletal muscle of rats, rhesus monkeys, and human have shown that focal accumulation of mtDNA deletions colocalize not only with respiratory deficient regions [150, 151], but also with areas of fiber atrophy and splitting, suggesting the importance of stochasticity of mtDNA mutagenesis in age-associated progression of muscular sarcopenia [152]. Such focal accumulation resulting from clonal expansion of few mutant mtDNA in a population, has been found to cause degeneration of post mitotic tissue like heart, brain and skeletal muscle tissue [27, 110, 153, 154], and this aspect is especially important in post-mitotic tissues, since the connectivity is of high relevance and the presence of only

a few non-functional cells can potentially disturb the performance of the entire organ. It is therefore unequivocally important to comprehend the origin and role of stochastic factors in *de novo* mtDNA mutagenesis as well as understand the exact dynamics of clonal expansion process.

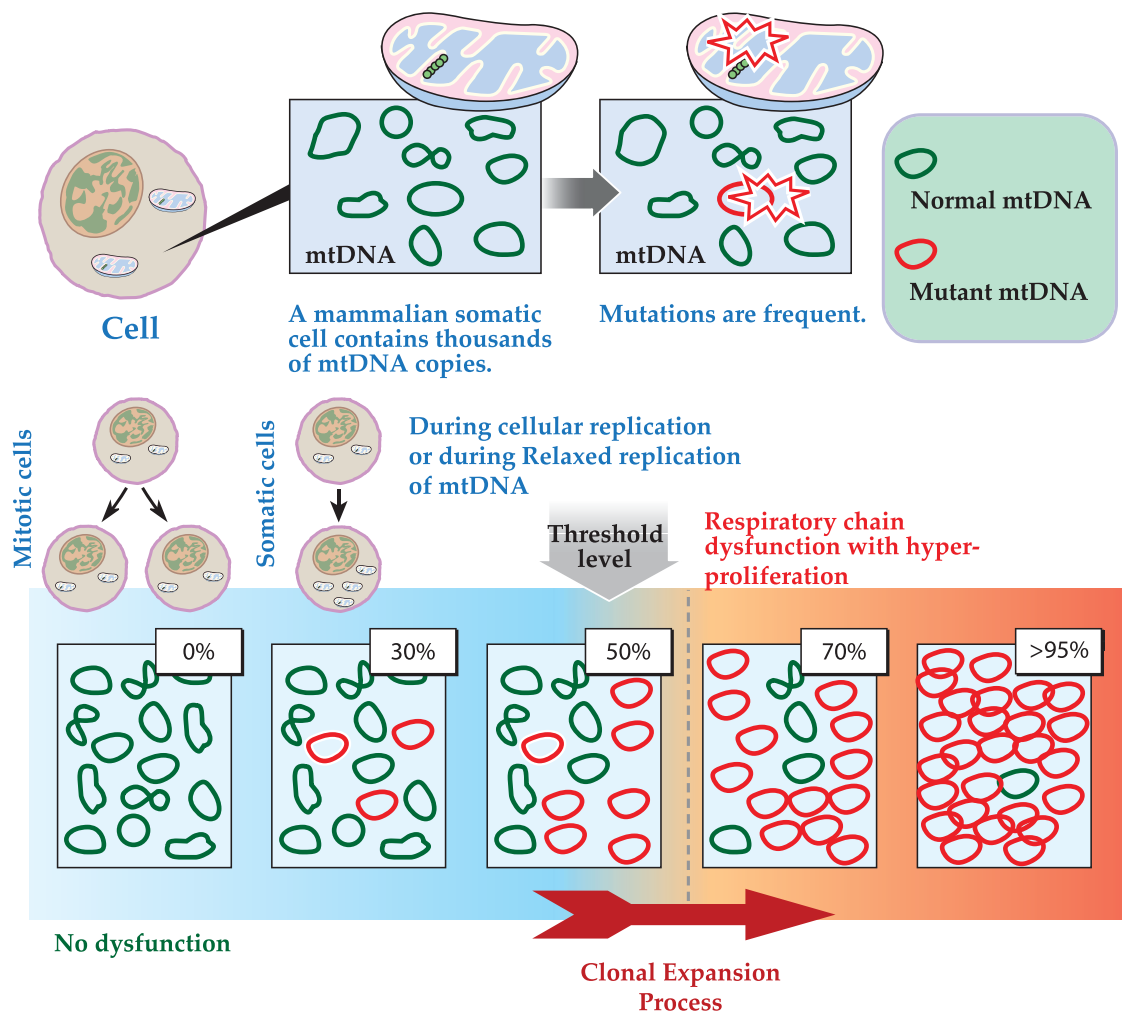


Figure 1.7: **Stochastic origin and clonal expansion dynamics of mtDNA mutation.** Mitochondrial DNA mutations occur infrequently within cells due to presumably mutagenic interaction of mtDNA with ROS. Presence of single functional mtDNA mutation in a cell is likely to have no influence of the cellular energy deficit. The mtDNA replication occurs in both mitotic cells (like epithelial cells, blood cells, etc) and post-mitotic cells (like neuron cells, muscle cells). The respiratory dysfunction can arise due to clonal expansion of a single or few mtDNA templates in the cell. A heteroplasmic mtDNA mutations only causes respiratory chain dysfunction if they are present above a certain minimal threshold level [145] (illustration adapted and modified from [33]).

Aside from the mouse models of aging, which is considered to be a significant evidence for the role of mtDNA mutagenesis in aging, there exists ample amount of evidence indicating mtDNA deletions to have functional consequence in aging process, as is reflected

from the correlation that exist between the deletions and the cellular biochemical parameters. For example, studies of neurons in the substantia nigra region of human brain, have demonstrated a loss of mitochondrial respiration and demise of cells with high levels of mtDNA deletion (levels much greater than functional wild-type mtDNA), whose numbers increases with age and in diseases phenotypes like Parkinson's diseases [27, 154]. Other studies involving single muscle fiber samples from rats, rhesus monkey and humans support a strong correlation between deletions, ETC abnormality and age associated sarcopenia [150]. Studies involving histochemical analysis followed by laser capture micro dissection (LCM) and PCR analysis of single cells/fibers have revealed that large scale deletions are confined to individual cells or fibers, while the neighboring cells do not harbor any deletions [26], leading to the conclusion that mitochondrial deletions originate from discrete stochastic events and reach to high mutation burden levels due to the clonal expansion of one or few mtDNA templates expanded over the passage of time.

1.5 Thesis scope & organization

While there exists abundant support in literature regarding the direct role of mitochondria and mtDNA in cellular physiology and aging, there is still a large knowledge gap on the mechanistic perspective of the origin and accumulation dynamics of mtDNA mutations during aging. Based on this impetus and considering that the cellular processes, especially the genetic processes are essentially non-deterministic in nature, and due to the large amount of mosaicity observed with mtDNA mutation burden in different cells of the same tissue, a stochastic modeling framework will be developed in this work to elucidate the mutation dynamics of mtDNA in aging cells. In particular, this dissertation focuses on the development of stochastic models to capture the role of different sources of variability, arising from the random processes of mtDNA turnover (replication and degradation) and the associated *de novo* mtDNA mutagenesis and due to the randomness involved with the

mutation quantification assays, for understanding the dynamics of mtDNA mutation accumulation, and to further elucidate the consequence of cellular heterogeneity of mtDNA mutation burden to aging tissues.

In this dissertation, different *in silico* models of mtDNA maintenance processes will be developed to capture the age-dependent dynamics of different types of mtDNA mutations like point mutations and deletions, that are relevant in the context of cellular aging. The construction of these models is intended to encompass only the most relevant features of biological processes, with minimal number of modeling assumptions, i.e. they are minimal models. Further, different statistical modeling techniques will be used to capture the variability associated with mtDNA mutation measurement protocols and based on this analysis, experiments are designed to minimize the measurement variability and subsequently improve the measurement accuracy.

The organization of the present thesis is as follows: **Chapter 2** provides a overview of the importance of mathematical modeling in unraveling the complexities of biological system and further provides a brief outline of past mathematical modeling in the context of aging, with specific emphasis on the models with mtDNA mutation dynamics. Further, this chapter provides details on the significance and relevance of cellular stochasticity in capturing the dynamics of cellular systems. **Chapter 3**, provides an overview of modeling and analysis framework developed for capturing the stochastic origin and accumulation of mtDNA mutation burden in mouse heart tissue. In **Chapter 4**, different hypotheses related to the clonal expansion process of mtDNA mutations are analyzed. In addition, parametric sensitivity analyses are performed, to obtain insights on the dynamics of mtDNA mutation accumulation in clonal expansion process.

One important age-related condition, which is directly correlated to the focal accumulation of mtDNA mutations in skeletal muscle tissue, is *sarcopenia*. **Chapter 5** outlines the development and analysis of a multi-scale modeling framework (Stochastic and Deter-

ministic models), developed for understanding the stochastic origin and subsequent clonal expansion of mtDNA mutations in sarcopenic skeletal muscle tissue. In this chapter, parametric sensitivity analysis affecting the development and progression of sarcopenia is also studied.

Statistical modeling framework to estimate the measurement variability associated with the Random Mutation Capture assay necessitating low DNA template dilution will be discussed in [Chapter 6](#). This chapter also outlines the development of an optimization protocol for reducing the measurement data uncertainty.

Finally, one of the most critical challenges experienced in the development of the present stochastic modeling framework relates to obtaining sensitive and robust estimates of model parameters. Thus, [Chapter 7](#) outlines some of the novel methods of parameter estimation framework for inferring rate constants associated with stochastic biochemical systems. Finally, [Chapter 8](#) presents an overview of the findings from this thesis and summarizes the major contributions. Future research directions are also discussed here.

Systems biology of aging — an overview

2.1 Mathematical modeling in biology

2.1.1 Historical perspective

Molecular biology has been instrumental in uncovering multitudes of biological information, like genome sequence and properties of other biological macromolecules such as RNA and proteins. Advances in quantitative experimental approaches will without doubt continue to provide significant insights into the functionality of different biological components. However, insights into the functioning of biological systems, based on these experimental data and based on purely intuitive understanding might not be practical [155]. Understanding the functionality of complete biological system requires the unraveling of complex network of biological interactions, and human intuition may be incapable of doing this. Cells, tissues, organs and organisms are parts of bigger system, whose specific functionality is screened by evolution and thus a system-level understanding is an important component of the biological research [156]. Computational biology and systems theory through systems biology presents an avenue towards this end, by which a comprehensive, quantitative understanding of biological systems can be obtained by studying how different components of biological systems interact with each other.

The usage of systems theory or computational modeling in biology is arcane; notably in the 1960's a number of researchers took modeling approach to search for general biological principles governing the behavior and evolution of biological systems in a way similar to that associated with the physical laws of non-living systems [157]. It was dur-

ing this transfer of ideas from physics to biology and with the perception that biological systems are special cases of physical systems, that the criticism culminated in most comprehensive debate related to the limitations of adapting modeling approach in understanding the behavior of complex biological systems [158]. The founding element of the present day systems biology dates back to Nobert Weiner, who pointed out the importance of systems perspective in understanding complex biological systems in his book 'Cybernetics, on Control and Communication in the Animal and the Machine', published in 1948 [159]. The first noteworthy numerical simulation in biology was published in 1952 by British neurophysiologists and Nobel laureates, Hodgkin, A. L. and Huxley, A. F., who constructed a mathematical model that explained the action potential propagating along the neuronal axes [160]. Their model described cellular function emerging from the interaction between two different molecular components, potassium and sodium channels, which can further be seen as the inception of computational systems biology [161]. In 1960, Denis Noble developed the first computer model of the heart pacemaker [162]. Another notable advance is in 1970, when Jacob and Monod proposed a feedback regulatory mechanism on a molecular level (Monod kinetics) [163], to investigate the regulatory mechanisms of protein interaction of allosteric enzymes.

Other prominent attempts that were made during this time, which are still classified under the umbrella of cybernetics are the Biochemical Systems Theory (BST) [164], developed in the late 1960s, and a related approach, Metabolic Control Theory (MCT), proposed in the mid 1970s [165]. BST and MCT are two ways to create simplified mathematical models of biological systems. Such models represent systems at and around a steady state functional levels. Although they were conceived as modeling approaches, both BST and MCT have resulted in large number of tools and theories for analyzing biological systems. In subsequent decades, researchers have been increasingly using different modeling/systems approach to elucidate different mechanisms of complex biological systems.

2.1.2 Characteristics of modeling

Observation of natural processes, especially biological processes, is frequently confronted with many complex processes, which cannot be explained with elementary principles, and the outcome of these processes cannot be predicated reliably based on experience. Mathematical modeling and computer simulations can aid in understanding the internal structure and dynamics of these complex processes. Engineering sciences are good instances of how mathematics has been effectively used in resolving the physical laws that govern different natural processes. Presently, it could be argued that the biologists are finding themselves in a similar situation to engineers, many decades ago, when they faced the need to use mathematical analyses to analyze and resolve complex dynamic systems for which empirical perception was inappropriate [166]. In modeling, the underlying strategy is to represent the natural systems by a set of state variables governing different processes and then investigate the relationship among these variables within a model (Figure 2.1).

The development of any mathematical models in science has two main aims, (i) understanding; and (ii) prediction. A key question in the mind of any modeler is, what are the minimal ingredients needed to make the results of a model match with experi-

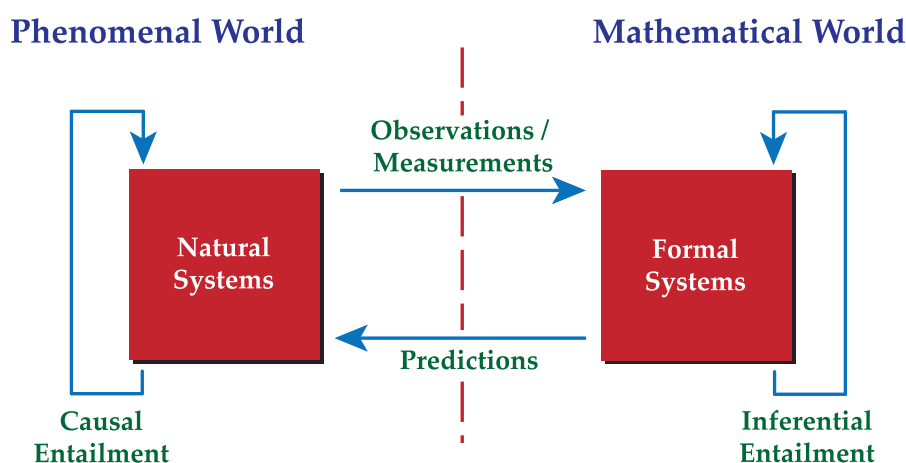


Figure 2.1: **Synergism between the natural system and formal system.** If the modeling relation brings both systems (Natural and Formal) into congruence by suitable modes of prediction and measurement cycles, the resulting outcome would describe the *natural principle* (illustration adapted and modified from [166]).

ments? Which are the dominant effects, which are secondary, and which are negligible? Generally in many aspects, the simpler the model, the greater the understanding would be. For example, if one were to incorporate all the currently available data on the heart, from physical geometry down to genetic pathways, into a mathematical model, then the resulting complexity would be such that one has simply replaced one complex biological system with another complex model that one cannot still understand. Simple mathematical model formulations have also other benefits e.g., they tend to have only few unknown parameters, and thus are easier to fit to data, and there is less issue of "overfitting".

A model only represents certain specific aspects of reality. The purpose of modeling study is only to answer specific questions. For example, if the purpose of modeling study is to predict system output, it is sufficient to obtain precise input-output relations, while neglecting the internal structure of the system (black box modeling). However, if the modeling intention is to elucidate the function of an object, then its structure and the relations between its parts must be described realistically. One way to do this is to formulate a model that is generally applicable to all the systems (like, Michaelis-Menten kinetics for enzymes), the other way is to conceive special formulations specific to objects in study (like, 3D structure of protein, specific sequence of genes). The mathematical models can be kept as simple as possible to allow easy implementation and comprehensible results, or it can be modeled realistically, to include finer details of the process being modeled. Again, the choice of the modeling approach (black box vs. detailed) depends on the final use of the model and there are generally more than one model that can serve the purpose.

2.1.3 Advantages of modeling

Mathematical representation of biological processes drives conceptual clarification. Models gain their reference from comparison with the experimental data. Therefore, their benefits are highly dependent on the experimental performance. Nevertheless, there are sev-

eral advantages of adopting modeling approaches for understanding biological processes [155, 156]:

- ★ A well-validated mathematical model can aid in characterizing different underlying mechanisms of complex biological processes and further provide useful predictions to highlight different insights and knowledge gaps existing in understanding of the functionality of these systems. With the development of such predictive models, mathematical models can be significantly beneficial in identification of novel drug targets or assist with dosing assessment of drug compounds. Highly focused, problem-directed modeling approaches have been shown to have a significant impact on drug-target validation and clinical development decisions [155].
- ★ Performing simulations are relatively cheap compared to the experimental counterparts and simulations can assist in experimentation. Since, model simulations can be repeated any number of time, under many different experimental conditions, simulations can be used for sorting out only those conditions which produces interesting system behavior, subsequently these conditions can be tried out using experimental methods.
- ★ With development of an adequate model, one can generate data that is beyond the present-day experimental capabilities. Also, perturbations can be imposed, that is beyond the feasible regime of real systems and can be imposed in such a manner that other components of the system is undisturbed, which is usually not feasible in real systems.
- ★ Modeling results can be presented in precise mathematical form that would allow generalization and greater understanding of the process (e.g. enzyme-substrate interaction using Michaelis Menten kinetics).
- ★ Modeling works exerts no harm to the experimental subjects, like animals or plants.

2.1.4 Different kinds of models

Modeling in computational systems biology can be broadly categorized into two distinct branches: (a) data driven modeling, which extracts the hidden pattern from experimental data and presents different hypotheses and, (b) first-principle modeling, which tests different hypotheses with *in silico* experiments, providing further predictions to be validated by *in vitro* and *in vivo* studies. In this dissertation, the second approach is used for addressing some of the intriguing biological questions related to mammalian aging.

A finer classification of the taxonomy of biological modeling methods, indicates a wide spectrum of mathematical abstraction with different resolution or accuracy. As seen in [Figure 2.2](#), at the highest level of abstraction, models are built to represent the interaction scaffold describing cellular components and their connectivity. At the next level of cellular modeling, the connectivity contains information on how biological information flows from one component to another. In low level models, the reaction kinetics of molecular interaction, affinities of network components, dynamics of the change in concentration of network components can be studied. At this level, the dynamics of the change in concentration of the states can be studied using Ordinary Differential Equations (ODE). However for modeling biochemical reactions at the level of cellular genetics, where the state concentration usually exists at nano molar level, the intrinsic cellular stochasticity cannot be neglected and as shall be described in the subsequent section, the assumption of mass action kinetics for modeling the system dynamics breaks down and has to be modeled using stochastic chemical kinetics. At the lowest levels of resolution, atomistic interaction and its consequence on the system behavior can be studied using atomic scale modeling.

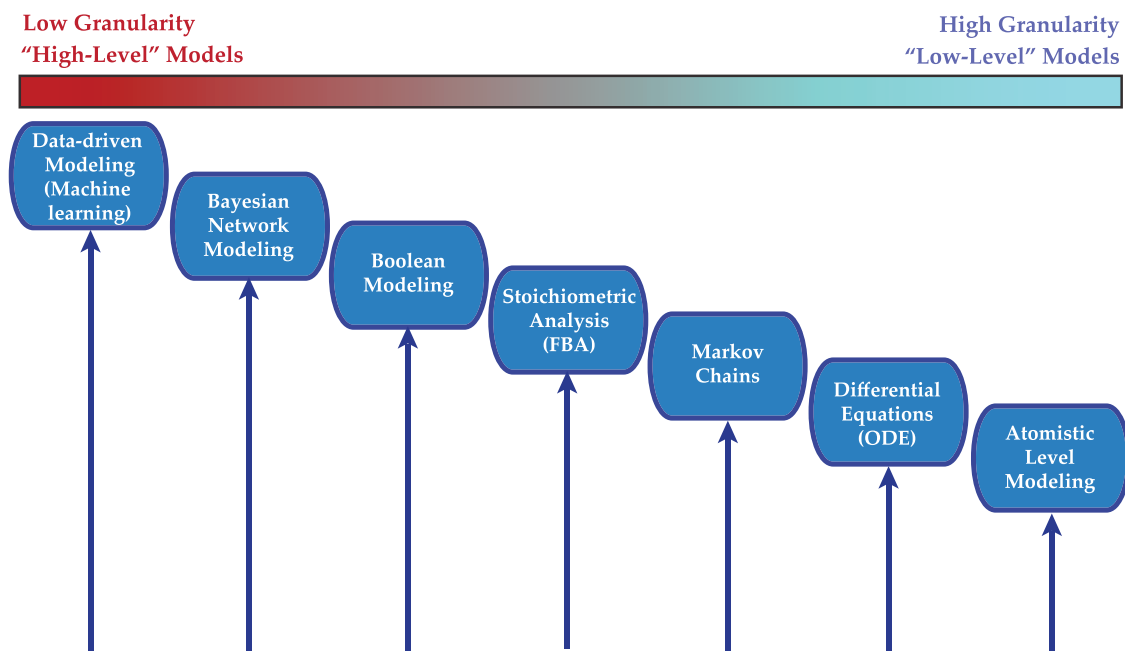


Figure 2.2: **Taxonomy of biological modeling in systems biology.** Different modeling techniques used in systems biology has been arranged in the increasing level of granularity or resolution.

2.2 Origin of cellular stochasticity

2.2.1 Stochasticity in biology

The importance of stochastic fluctuations in gene regulation have been known for more than three decades [167]. But the ability to measure gene expression profiles at single cell resolution have recently renewed the interest in stochastic modeling of biochemical reactions that underlie the genetic networks [168]. There is compelling evidence in the literature that the outcomes of cellular events in both the prokaryotic and eukaryotic organisms are non-deterministic in nature [169–171]. A good example demonstrating biochemical stochasticity comes from the genetic transcription process in *Escherichia coli* (*E. coli*) [172]. In the experiment by Elowitz and Leibler, the expression levels of green fluorescent protein (gfp) in *E. coli* were measured in single cells by flow cytometry and the study showed that cells in a clonal population (isogenetic population), grown in uniform environmental condition can express vastly different protein levels. Furthermore, the synthesis and degradation of different biological macromolecules like proteins and RNAs are also sub-

jected to similar type of stochastic dynamics [173].

Kinetics of conventional macroscopic reaction system is modeled using system of ODE, with an implicit assumption that chemical concentration varies continuously with time. Two important characteristics associated with such systems are: a.) number of molecules of states in the reaction system is much larger than the thermal fluctuations in concentration, and b.) for the different reactions in the system, the number of reactions is large within each observational interval. At low concentrations of reacting species and at slow reaction rates (such as the eukaryotic genetic regulatory processes), the conventional deterministic chemical kinetics (mass action kinetics) may not describe the dynamics of the system of coupled reactions accurately [174, 175]. For such systems, it has to be recognized that individual chemical reactions occur discretely and are separated by time intervals of random length and both the characteristics described above fails for smaller reaction systems like the genetic circuit. Thus, in small, low-rate chemical systems, it is necessary to pay close attention to the fact that changes in the chemical population levels occur as integral number of molecules, and are occasioned by essentially random distinct reaction events.

Cellular noise can be categorized into two sources: *intrinsic noise* and *extrinsic noise* [176]. As discussed above, the intrinsic noise arises from the low copy number of biomolecules associated with the intra-cellular processes, while the extrinsic noise arises due to the fluctuations in extra-cellular milieu of the cell (environmental changes) [176]. It has been observed that a simple transcriptional regulation depends heavily on transcriptional rates and the concentration of the promoter and repressor molecules, and a combination of these opposing processes determine the stability of the biochemical network [172, 177]. Thus, noise in cellular processes of living cells are characterized by a complex web of genetic networks and the execution of this web depends on the faithful signal propagation along the genetic pathways. The presence of noise and fluctuations in the genetic pool can quiet

often determine the fate of an organism within a given population [178, 179].

2.2.2 Stochasticity in aging process

Similar to other biological systems/processes, one of the critical challenges associated with understanding of the underlying mechanisms in aging process, relates to the heterogeneity associated with the aging process, where longevity can be achieved by different combinations of genetic, epigenetic, environmental and stochastic elements. It is increasingly evident that the multi-factorial process of aging does not occur only at the cellular level, but it also occurs in the level of tissue-, organ- and or at organism-level at different rates. Thus, an aging body can be considered as mosaic of tissue and organs and within tissue, mosaic cells displaying different levels of cellular morbidity [180].

Thus, even in the same organism, the aging process appears to follow different trajectories in different organs, tissues and even in cells, which are believed to arise from the accumulation of unrepaired damages at different rates [181–183]. A typical example is brain, where different regions of brain, such as cortex, hippocampus and cerebellum, have shown different levels of neuro-degeneration and inflammation in the same subject [109, 184]. Similarly, analysis of substantia nigra region of human brain of normal aged subjects and patient's with Parkinson's disease revealed a high degree of mosaicity of dysfunctional neuron cells [27, 154]. This mosaicity is also highly dynamic, which can change significantly with time and this is due to complex non linear interactions among different components of the system, resulting in phenotypes that are not easily predictable by considering individual components [185].

Stochasticity may be an explanation for the marked variability observed in the life spans of organisms chosen from an iso-genetic population, reared under uniform environmental conditions, and the cell to cell variability of the mtDNA mutation burden in a tissue. Analysis of the life span distribution in different short-lived species such as *Cerati-*

tis capitata, fruit flies *Drosophila melanogaster* and *Anastrepha* species, indicated that, despite the population being isogenic, their life span distribution exhibited a typical bell-shaped distribution. Each of the individuals from such a population is intrinsically different in their susceptibility to death. This suggests that some individuals are more frail because of intrinsic and extrinsic variations, and consequently may have higher mortality compared to its age-matched cohorts.

One striking aspect in which stochasticity have been found to have a significant role in the biology of aging is the wide variability in both the time of onset and rate of apparent deterioration within an isogenic population of *Caenorhabditis elegans* (*C. elegans*) reared under uniform environmental condition. In a recent work characterizing the stochastic and genetic factors influencing the tissue specific decline with aging, Herndon *et al.* demonstrated the prevalent variability of cellular and organ level functionalities with age, using GFP tracking method in isoclinal *C. elegans* population [185]. This study indicated that different cell types exhibited different rates of aging, and the cellular stochasticity is largely responsible for age related decline of multiple tissues. A similar scenario has also been reported in the case of human centenarian study, in which a wide dispersion in the frequencies of human mortality was observed [186].

2.3 Systems Biology & aging research

Aging of an organism is both a manifestation and a consequence of complex changes in the structure and function across all the levels of biological organization from simple biological macromolecules to the whole body systems [6, 9]. Like any other complex systems, the behavior of aging systems cannot be predicted completely by discrete analysis of its individual components. This reductionist approach has traditionally been the norm in biological and biomedical research. Although this traditional methodology has unquestionably yielded significant and insightful progress in understanding the fundamental mechanisms

of aging, it is nevertheless becoming clear that a more integrated approach is required [9, 11, 187].

Given that the process of aging is manifested across the entire spectrum of life, from cellular level to the entire ecosystem, it is natural and compelling that many of the ideas, techniques and the way of thinking that fall under the domain of systems biology or computational modeling should be used in addressing important challenges of aging [9]. With such integrated framework, it will then be possible to develop viable strategies for addressing multiple challenges of human aging. However, a critical roadblock in using this framework in aging research is the development of multi- and trans-disciplinary approaches, where ideas and concepts derived from other disciplines like mathematics, physics and engineering can be convoluted in generating significant insights.

One of the areas where modeling of aging processes will be beneficial, is in building integrative models that is well suited to take into account the comprehensive view of different interactions involved in a system. This is particularly important since many of the key processes involved in cell maintenance and metabolism do not act in isolation, but act in a coordinated manner. When the activity of one enzyme changes, all the other connected metabolites and enzyme activities may be affected [11]. In some cases, there may be redundancy in system pathways, which may provide buffering capacity against damages, whereas in other cases, the effect of damage may also be propagated. Quantitative modeling can also be beneficial in understanding the actions of genes that affect the rate of aging. Over the past decades, several genes have been identified that affects aging in yeast, nematodes, fruit flies, and mice, and there is a growing interest in genes affecting human longevity [188, 189]. Experimental data are revealing that the interaction of several of these genes within pathways that control the rate of aging, and there is increasing evidence that the most important gene locus corresponds to those that affects basic cellular processes, such as insulin and insulin-like growth factor (IGF) signaling, which are

strongly conserved across the species [188, 190]. Nevertheless, the understanding of different interactions and their effects are still largely sparse. Upon that, these studies should also take into account the intrinsic stochastic nature of gene regulatory networks.

Another important area where the quantitative modeling of aging process can play a significant role, which also forms the central theme of the present work, is the understanding of the role of unrepaired biomolecular faults in aging process. According to the disposable soma theory [5, 6], aging is caused due to gradual accumulation of unrepaired random molecular faults, resulting in increasing fraction of damaged cells and eventually leading to functional impairment of tissue and organs. The idea of aging as a result of macromolecular damage is straightforward and is supported by sufficient amount of experimental evidences. However, its understanding presents a number of significant challenges since [2]: a.) it predicts that there are multiple mechanisms that may cause aging, b.) it also predicts that the process of aging is inherently stochastic, and c.) since aging is an outcome of multiple mechanisms, a high level of complexity is expected. Therefore, the current challenge for aging research is to integrate knowledge from fundamental physiological properties of different species as they age with data from molecular and cellular investigations, and to provide a more comprehensive framework that accommodates these approaches to study the complex process of aging [187]. For all the above said reasons, there is an exceptional need of computational approach in aging research to be integrated with different experimental approaches to elucidate the underlying aging mechanisms.

2.4 Models of cellular aging - an overview

Similar to the plethora of theories explaining the mechanisms of aging, there also exists a wide range of computational models in aging research. These models range from detailed modeling of individual intra-cellular mechanisms to higher level systems modeling required for elucidating aging mechanisms. Since the motivation of present work is in un-

Understanding the role of molecular mechanisms in cellular aging, this section outlines past models of intracellular mechanisms and their impact on enhancing the understanding of different mechanisms involved in aging.

2.4.1 Models of intra-cellular mechanisms

Several models are currently available that focus on individual intra-cellular mechanisms. Largely, these models are concerned with telomere shortening, accumulation of somatic mutations, and accumulation of defective mitochondria.

Telomere shortening

Early models attempting to explain the replicative senescence of human fibroblast was based on the end-replication problem [191]. Later models, included more specific features of telomere shortening process, like, Rubelj and Vondracek modeled abrupt telomere shortening due to DNA recombination or nuclease digestion [192]. Most recent models have included features like acceleration of telomere shortening due to increase in oxidative stress [193, 194]. These simulations indicated that increasing the level of ROS in cells led to fewer cell divisions.

Somatic mutations

The role of somatic mutations in aging was proposed several decades ago, when the experiments indicated that irradiation shortened life span in animal models and induced aging phenotypes [195]. Due to the advances in methodologies for quantifying DNA modification and repair, the research in understanding the role of somatic mutations in aging has become particularly active. Szilard proposed a model [196], which assumed that in a diploid organism, pair of genes had to receive a mutational hit simultaneously, in order to render a cell dysfunctional. A deterministic model was developed to account for the accumulation of recessive mutation in human fibroblast and its effect on life span of diploid

fibroblast population [197]. Further, a stochastic model based on the same processes was developed to consider the possibility of a synergistic interaction between mutations [198]. These simulation results indicated that models used for simulating somatic mutation theory did not well characterize the experimental data on the life span attributes of the human fibroblast population, and based on these results authors further questioned the suitability of somatic mutation theory in explaining the cellular aging process.

Chaperone models

Cellular chaperones have a significant role in maintaining protein homeostasis. It is observed that an important class of chaperones known as the heat shock proteins (HSPs) is impaired with age also the functions of several other chaperone declined with age. Aberrant/damaged proteins accumulate with age and are implicated in several important age-related conditions (e.g. Alzheimer disease, Parkinson's disease, cataract). Therefore, the balance between damaged proteins and available free chaperones may be greatly disturbed during ageing. There have been several models focusing on different roles of HSPs on cellular physiology. However, to date, there is only one model that has looked at role of HSPs in aging process [193]. This model addressed how the HSPs are upregulated after an increase in intra-cellular stress and can be used to investigate the role of stress on protein homeostasis.

2.4.2 Role of mitochondria in aging

Deterministic modeling

Early models on the role of mtDNA somatic mutations in aging and degenerative diseases were primarily deterministic. These early deterministic models were developed to study the role of mtDNA 'vicious cycle' in the origin and accumulation of mtDNA mutations in cells. Kowald and Kirkwood developed a mitochondrial population dynamics model

to simulate the mtDNA mutation dynamics caused due to free radical attacks [199]. The population dynamics model tracked for the mutation dynamics and its influence on the mtDNA content in individual mitochondrion. The model developed in this work, considered mitochondria as isolated independent units, with fixed count of mtDNA and without the existence of mitochondrial fusion-fission process. Such method of modeling produces some physiological in-consistency in the model predictions such as: local enhanced mitochondrial biogenesis and the associated enhancement of the mtDNA mutations. The model developed specifically included features of ROS induced 'vicious cycle' and the decrease in the cellular bioenergenesis due to the accumulation of mtDNA mutations. Nuclear retrograde response for the mitochondrial biogenesis was modeled as a linear function of mtDNA mutation burden in each mitochondrion and the ROS production rate arising from the damaged mitochondria was modeled as a non-linear function of damage. An ODE formulation is used to describe the mitochondrial population dynamics of mtDNA mutation burden, where the evolution of different mutation levels (mitochondria with different damage extent) was captured. The model comprised of 7 ODEs with 13 model parameters tracking the evolution of mitochondria with different mutant load, and two algebraic equations representing the radical generation and the associated mitogenesis. Due to the unavailability of the basal mtDNA mutation rate, different mtDNA mutation rates were considered to study the effect of its magnitude on the cellular mtDNA mutagenesis. Model prediction indicated increasing accumulation of mtDNA mutations with increasing mutation rate. Unlike what is observed in the experimental data [26, 28], in which only few cells ever get to very high mutation burden, the simulation results in this work indicated that all cells had high mutation burden and eventually underwent apoptosis. Most of the results predicted from this model, only have qualitative resemblance with the experimental data. The authors did not show any explicit quantitative comparison of their simulation results with the experimental data.

In different models describing the network theory of aging [8, 200], the same authors further integrated the details related to several of their earlier models into one comprehensive modeling framework. Specifically, they attempted to integrate the Free radical theory and the Protein Error theories into one modeling framework [200]. In this work, the authors attempted to model the interaction of several mitochondrial network mechanisms that may be responsible for aging process. The model included the homeostatic maintenance of error propagation in different mitochondrial macromolecules as a consequence of opposing and competing reactions of free radical attack and the consequent ROS scavenging reactions. On a global perspective, model specifically simulated the accumulation dynamics of mtDNA mutations with age, by including most of the processes related to the OXPHOS machinery, like protein synthesis, turnover process related to proteolytic degradation and recycle of aberrant mitochondrial proteins, RNA synthesis, Free radical generation, ATP synthesis, and ROS scavenging dynamics. This model again used an ODE formulation, predominantly the Michelis-Menten (MM) kinetics, to obtain the dynamical evolution of various mitochondrial metabolites (e.g. the MM kinetics was used to specifically model the DNA transcription and translation machineries related to mtDNA). Such comprehensive detailing makes the model quite exhaustive with around 31 ODE equations having approximately 42 parameters. Most of the results predicted from this work were shown to only have qualitative resemblance to the experimental data. Like their previous works, what is mainly lacking in this work also is a quantitative comparison of simulation results with the experimental findings.

In Kowald and Kirkwood; 2000 [201], the same authors further discussed different hypotheses available in the literature to explain the mtDNA mutation clonal expansion process. They pointed out different limitations associated with several of these hypotheses and proposed the delayed degradation of defective mitochondria as a possible mechanism to explain the clonal expansion process. This hypothesis is also referred to as survival of

the slowest hypothesis [202]. The central idea of this hypothesis is that the mitochondria harboring mutant mtDNA has lower OXPHOS activity and thus these organelles are relatively immune of further oxidative damage arising as a consequence ROS generated during energy synthesis. Thus, since the mitochondria having mutant mtDNA harbor lesser membrane damage compared to wild-type mitochondria, the mutant mitochondria are spared from the mitophagic degradation and they accumulate to result in clonal expansion [201]. In other words, the frequency of functional mitochondrion getting removed from the population is much higher than the organelle having mutant mtDNA, thus resulting in its clonal expansion. An implicit modeling assumption related to this hypothesis was that the mitochondria are disconnected from each other and existed as independent members in the population. In addition to the original hypothesis, authors of this work also included an additional assumption of damaged mitochondria growing and dividing slower compared to functional organelles, due to the deficit of energy. Based on the original hypothesis, authors developed the mitochondrial model as discrete independent entities. However, this assumption is inconsistent with what is observed *in vivo*. There is abundant evidence indicating the existence of frequent fusion-fission process, providing an efficient complementation of the internal metabolites of mitochondria, including mtDNA nucleoids [34, 140, 203–205]. Consistent with their earlier model [199], the ODE model developed in this work tracked for the population of mitochondria with no damage and with different degrees of damage (number of mutated mtDNA copies in mitochondria). The model developed here specifically included features of ROS induced 'vicious cycle' and the consequent decrease in the cellular bioenergenesis. The model highlighted an extensive framework of mitochondrial maintenance processes involving: mtDNA synthesis, ATP synthesis, free radical generation during the OXPHOS, damage arising to mtDNA and the mitochondrial membrane damage and subsequent degradation of mitochondria. Different kinetic models including the Michelis-Menten and Hill type kinetics have been adapted in

this work. Most of the mathematical constructs used in this work was borrowed from their earlier works [8, 199, 200]. Similarly, to validate the hypothesis, authors borrowed the mitochondrial population dynamics model, which was basically 7 ODE equations having 20 parameters associated with them. Using this model, authors tested several hypotheses like the influence of mitochondrial turnover rate, influence of growth difference factor in damaged mitochondria and the influence of cell division rate on the clonal expansion dynamics related to different mitochondrial population. Although the authors demonstrated many qualitative comparisons with the observations related to mtDNA clonal expansion dynamics related to post mitotic cells of normal aging and disease phenotypes with experimental observation, this work also produced very little quantitative validation to the actual experimental data.

In a recent work by Mao *et al.*, the authors used proteomic analysis to quantify the load of oxidized nuclear encoded proteins in mitochondrial membrane, to determine the instantaneous mutation load prevailing in the mitochondria [206]. Quantification of protein contents in different mouse organs was obtained using 2D-gel electrophoresis and subsequent mass spectrometry measurements. These measurements were used to estimate the over-expression levels of nuclear encoded respiratory chain subunits. The model used in this work was a simplification of the earlier model using the hypothesis of the survival of the slowest (SOS). The ODE model developed in this work tracked for the changes in the concentration of mutant, wild type mtDNA and free radical produced during the OXPHOS process. The model consisted of 3 ODE equations for tracking the dynamics of each of the species in consideration and the model consisted of 9 modeling parameters. The replication rate of the mtDNA templates was modeled as a linear function of the change in molecular count of mtDNA (both wild-type and mutant) below a homeostatic mtDNA template level in somatic cells. Combined with differential protein expression level data obtained from the proteomic analysis, the model was further used for estimating the un-

derlying mutation rate of mtDNA mutation process using least square parameter estimation. This model however presents a considerable flaw in its prediction, where all cells in the tissue became homoplasmic with mutant mtDNA in 3 years (average natural life span of mouse). However, it is generally appreciated that in a tissue there exists substantial amount of mosaicity in mtDNA mutation load among cells and only a few cells ever have high mutation burden [111, 207].

Capps *et al.* [208] developed a deterministic model based on Chinnery and Samuels earlier stochastic model (details in the next section) [209], to characterize the regulation of nuclear retrograde response to different levels of heteroplasmy observed in post mitotic cells. Like the Chinnery and Samuels's stochastic model [209], in this work also authors have attempted to characterize the role of mtDNA relaxed replication in influencing the levels of wild-type and mutant mtDNA population found in post mitotic cells. The outcome obtained from the deterministic work was compared with the earlier simulation results [209, 210]. The deterministic model developed in this work was subjected to extensive analysis for understanding the behavior of mtDNA replication to different scenarios such as steady state dynamics (asymptotic behavior of the mtDNA states) and the role of cellular stochasticity in the mtDNA mutation dynamics. Also, the influence of the presence of pathogenic mutations on the development of mtDNA homoplasmy was studied. The following three cases were considered for the study using the deterministic framework: a.) the maintenance of mtDNA copy number in the presence of nuclear retrograde signaling, b.) the maintenance of wild type mtDNA in presence of mutant mtDNA and, c.) the contribution of pathogenic mtDNA mutations towards the retrograde signaling and further manifestation in the mitochondrial homoplasmy. The authors further proposed possible experiments that could be designed to test different hypotheses.

Stochastic modeling

Kowald *et al.* have developed a stochastic model mtDNA deletion accumulation in post-mitotic tissue based on a different hypothesis, the survival of tiny (SOT) [211]. The motivation for developing a stochastic model for studying the mutation dynamics in a tissue was attributed to the non-uniform accumulation of mtDNA mutations observed in the tissue of normal aging and mitochondrial disease phenotypes [212]. The model comprised of basic mitochondrial maintenance processes of mtDNA replication, degradation and *de novo* mtDNA mutation generation. In this work, the replication rate of mtDNA was modeled as a negative feedback function, such that the replication rate of mtDNA declines with increasing respiratory-functional mtDNA copy number. The feedback response was modeled as Hill type kinetic function. The replication propensities of mutants were modeled such that the shorter mutant mtDNA replicates faster. The degradation propensity of different mtDNA templates were modeled with uniform probability, i.e. the removal of mtDNA templates from the population was assumed to be independent of the type (wild-type or mutant). Similar to the earlier model on the mitochondrial population dynamics [199], in this work the dynamics of deletion generation of different size was also modeled as a progressive process, i.e. the larger deletions was considered to either arise from normal mtDNA templates or from templates having smaller deletions. Such assumption is problematic as, the creation of deletions of different sizes are independent from each other and directly arise from the normal wild-type mtDNA templates [58, 103, 121, 213]. However, unlike their earlier works [8, 199–201], the authors have assumed a well mixed population of mtDNA arising from the frequent fusion-fission process. The stochastic simulation of the mitochondrial turnover process was simulated using the Gillespie's Monte Carlo sampling method [175, 214]. Like the results from their earlier deterministic formulation, the simulation results of the stochastic model also indicated that most the cells harbor mutant at the end of 400 days of mouse life span, despite the model using con-

servative estimate of mtDNA deletion rate [212, 215]. The results of the stochastic model also indicated a high sensitivity of the model predictions to mtDNA turnover rates. The model also specifically incorporated the increase in the oxidative burden of mtDNA in the presence of functional mtDNA mutations (assumption of the existence of ROS 'vicious cycle'). The simulation results related to the spectrum of mtDNA deletions were compared with the experimental data on mtDNA deletion spectrum obtained from the mouse heart tissues of different age.

Instead of building large, complex models, Chinnery and Samuels developed a minimal stochastic model of mtDNA turnover process to understand how mtDNA relaxed replication (a process in which mtDNA turns over independently from the cellular division, such that the mtDNA molecules replicate at random making one or more copies of mtDNA, while still conserving the total number of mtDNA population in the cell) leads to a random intra-cellular drift of mtDNA mutants [209]. They assumed a well mixed mtDNA population with two different mtDNA alleles with a fixed proportion in each cell simulations (one was wild-type mtDNA and the other allele was mutant mtDNA, harboring neutral mutations, i.e. mutations which do not affect cellular respiration) [209]. The dynamics of the intra-cellular heteroplasmy of these alleles was tracked using the stochastic model. In this work, a Langevin type kinetic formulation [209] was adapted to model the processes involved in mtDNA turnover: mtDNA degradation and replication. The evolution of the mtDNA alleles were tracked at constant time interval, assuming that no *de novo* mutations occurs during the mtDNA replication process. The simulation results indicated a significant fluctuation in the population levels of different alleles (or the heteroplasmy level), hinting the predominant role of stochasticity in deciding the fate of cells with regards to the mtDNA mutation burden. The authors further studied the fixation dynamics of different mtDNA alleles, to understand the role of cellular stochasticity in mtDNA mutagenesis and subsequent accumulation of mtDNA mutations. In this work,

the authors also modified their stochastic model of mtDNA turnover process, to study the role of pathogenic mtDNA mutations. In this model, only the rate governing the mtDNA replication process was modified to include a piece-wise linear function of optimal count of wild-type mtDNA template in the cell, to simulate the nuclear retrograde response in the presence of pathogenic mtDNA mutations. The modeling results showed that, mutant mtDNA proportion increases significantly compared to the wild-type mtDNA genome, but the levels of mtDNA copy number are much below (several orders of magnitude) the levels observed in the experimental data [26].

Based on the above model of mtDNA turnover process [209], Elson *et al.*'s proposed a mechanism for the clonal expansion of mtDNA mutation in post-mitotic human tissues, in the absence of any replicative advantage of mtDNA and since the mutation considered in this work was non functional (does not effect the OXPHOS conditions of the cells), the authors also proposed that the clonal expansion dynamics may occur in the absence of nuclear retrograde feedback [210]. The authors proposed several arguments to challenge some of the commonly used hypotheses for explaining clonal expansion process, such as the SOT and SOS hypotheses. The same Langevin type kinetic formulation was again adapted to model the processes involved in mtDNA turnover processes. However in this work, the replication of wild-type mtDNA copy was assigned with certain probability corresponding to the basal mutation rate, to give rise to mutant mtDNA. The evolution of the mtDNA alleles (wild-type and mutant mtDNA) were tracked at constant time interval. Simulation results indicated that in some of the cells, the proportion of mtDNA mutants increased to very large levels, similar to what is seen with clonal expansion process. Further, the results hinted that during the simulations of first decade of human life almost every cell had pure wild-type mtDNA population. This prediction presented a significant challenge with Elson *et al.*'s model in predicting the mtDNA mutation dynamics associated with short lived organisms like mice (average life span ≈ 3 years) [210]. Elson *et al.*'s

model may not be able to predict the COX deficiency observed in post-natal somatic tissues of mice [26, 216, 217]. The authors further studied the influence of changing the basal mutation rate on mtDNA mutation. The results indicated that, changing the mtDNA mutation rate had a marked effect on the number of cells becoming COX deficient. Further, the influence of mtDNA copy number on the mtDNA mutation burden was studied. Simulation results indicated that cells having higher mtDNA copy number harbor lesser mtDNA mutations. In contrary to this finding, experimental evidence suggest that tissue with higher oxidative capacity (larger mtDNA count per cell) has higher mutation load associated with them [139]. The stochastic model developed by Elson *et al.* is highly sensitive to the mtDNA turnover rate. In this model the authors have used a turnover rate of 10 days [209, 210, 218]. It would be interesting to see the behavior of their model for lower value of mtDNA turnover rate. The simulation predictions are likely to be significantly different compared to their original work (will be further discussed in the subsequent chapters of this dissertation). Elson *et al.* further characterized the seeding mutations for the clonal expansion process. Their analysis indicated that the earliest *de novo* mutations occur during the late childhood or during the early adulthood. Subsequently, the mutations start to expand clonally. Despite being a hypothesis work, very little experimental validation is provided. However, authors determined optimum values of some of the parameters (like mutation rate) to match the experimental data on the fraction of COX deficient cells in human muscle. Although the model addresses aspects related to mtDNA mutation clonal expansion process, the authors did not discuss anything about the mtDNA hyperproliferation process, which is generally found to be associated with mtDNA clonal expansion process [26, 29, 140, 205]. The shortcomings of the random drift hypothesis proposed by Chinnery and Samuels groups in explaining the mtDNA clonal expansion process will be dealt in more details in [Chapter 4](#).

Similar to the Elson *et al.*'s model on random intra-cellular drift of mtDNA mutations

in post mitotic tissues [210], Collier *et al.* [219], demonstrated that pure stochastic drift of mtDNA mutations and segregation of mutant mtDNA between the dividing cells of mitotic tissue is sufficient to produce mutant homoplasmy observed in mitotic tissues [29]. The stochastic model developed in this work simulates the simple mitochondrial maintenance process such as mtDNA replication, degradation and mutation arising during cellular division. The tissue was assumed to contain approximately 50 mtDNA undergoing 1000 generations of cell division. Similar to the Elson *et al.*'s work on the mtDNA turnover dynamics in post-mitotic tissue [210], the model developed in this work was also based on Langevin formulation. The mtDNA replication was simulated to increase the mtDNA count to twice the nominal mtDNA count in the cell at the end of cell cycle. Further, the mtDNA were partitioned randomly between the dividing cells and the number of mutants generated during a single generation of cell division was simulated by sampling from a Poisson distribution having a mean corresponding to the basal mutation rate. Using this model, simulations were performed to analyze the dynamics and accumulation of mtDNA mutations in both normal aging epithelial cells and with tumor cells in human, which was further validated with the experimental data. Although the model addressed different aspects related to mtDNA mutation clonal expansion process, the authors did not discuss anything about the mtDNA hyperproliferation process, which is generally found to be associated with mtDNA clonal expansion process [26, 29, 140, 205]. Another interesting results associated with this work also is the mutation load in cells with different mtDNA counts. Unlike the experimental observation [139], simulation of cells with higher mtDNA count were observed to have lower mutant burden in them.

Modeling perspectives

There are several shortcomings in the previous modeling works, which the present dissertation will attempt to address. Most of the above deterministic modeling framework were constructed based on several modeling assumptions and based on large number of ODE

equations and associated modeling parameter space. As discussed earlier, with increasing complexity of the mathematical model, there are several challenges associated with its usage, such as: a.) the understanding of the underlying biological processes based on such complex models becomes significantly difficult, b.) with increasing model complexity, the number of unknown model parameters also increases and consequently several of the associated parameters may be unknown and the parameter estimation based on such complex model is highly likely to over fit the experimental data. In fact, one of the potential criticisms related to the usage of such complex models, is that most of the parameters used in these models are chosen *ad hoc*, without having sufficient physiological significance. In many of these cases, few model equations and the associated modeling parameters are chosen to serve as tuning variables to obtain desired output profiles.

In most of the earlier modeling works developed for addressing the mtDNA mutation accumulation dynamics, one of the important aspect which is sparingly considered is the sensitivity analysis of different parameters associated with the mtDNA turnover process. One of the parameters which was observed to be very sensitive to the mtDNA mutation accumulation dynamics is the turnover rate, as the mtDNA *de novo* mutations are directly tied to the replication process. Thus an accurate estimation and understanding of the mitochondrial turnover rate is essential for reliable modeling of the mitochondrial mutation dynamics. In this dissertation, different parametric sensitivities related to mtDNA turnover process will be analyzed, to get a better understanding of different sensitive parameters which requires to be estimated more critically for getting more reliable model predictions. Also, in most of the earlier modeling works, effort in validating the model outcome with the experimental data have been relatively sparse. In this work, sensitive validation of the simulation results will be prioritized. Additionally, in all of the previous simulations, the origin and accumulation of mtDNA mutations is primarily modeled only during the post natal stage of an organism. The consequence of developmental cell

lineage on the mtDNA mutation accumulation is rarely studied. The enormous quanta of developmental cell divisions and the associated mtDNA replication is likely to contribute significantly to the post natal mtDNA mutation burden, and thus in the present modeling framework the mtDNA turnover process and the segregation of mtDNA between the dividing embryonic cells during animal development will be considered.

Stochastic drift in mitochondrial DNA point mutations: a novel perspective *ex silico*

3.1 Introduction

Although there is reasonable evidence for an age-dependent increase in mtDNA mutations, the dynamics by which these mutations accumulate is still largely unclear. Inferring dynamics and more importantly, the mechanism by which mtDNA mutations accumulate critically depends on accurate quantification of oxidative and mutational burden, which poses significant experimental challenges [127]. Many of these challenges stem from the limitations associated with experimental protocols in measuring oxidative damages and mutational frequency [220, 221], which typically exist at extremely low magnitude. Also, another requirement for addressing these uncertainties is a better understanding of the inherent stochasticity of cellular processes [178]. The accumulation of mtDNA mutations likely involves complex stochastic factors, such as the inherent random nature of mutations and related cellular processes in the context of aging.

As discussed in [Chapter 1](#), published reports show conflicting results regarding the levels of oxidative damages and mutation dynamics of mtDNA during aging [116, 132, 222, 223]. Earlier estimation of mtDNA mutation quantification was obtained using PCR amplification of short fragments of DNA. However, recently it has been shown that due to the underlying intrinsic error rate of polymerase enzyme, these PCR methods introduce significant amount of spurious mutations [224, 225], thus over estimating the mtDNA mutation frequencies [224, 226]. A highly sensitive method based on the random mutation capture (RMC) assay has been developed for the quantification of mtDNA mutation fre-

quency [141]. This method is based on restriction enzyme digestion and amplification of mtDNA molecules carrying mutations at the corresponding recognition site [116]. Application of the assay to wild-type mice revealed mtDNA mutation burdens that were two orders of magnitude lower than previously determined using PCR-cloning and sequencing protocols [130, 131]. This indicates that PCR artifacts may have been a major contributor of errors in the past reports. Furthermore, quantification of age-dependent accumulation of point mutation burdens using the RMC assay in wild-type mice suggested an exponential increase, apparently supporting the existence of a 'vicious cycle' in the mutation accumulation [222, 227]. However, the low levels of burden suggest that point mutations may not be a major determinant of lifespan [116] and it is difficult to see how a positive feedback mechanism could set in at such a minuscule level of point mutation burden.

In this chapter, the aim is to address these challenges using a systems approach by the way of constructing mathematical models that encompass the most relevant biological processes and also features related to experimental protocols to comprehend the origin and consequence of mutation variability that arises in individuals of a mouse population. Additionally, this chapter attempts to better understand the influence of intrinsic stochasticity of the mutation process on the variability observed in the experimental data. Such understanding may reveal possible causes of disagreements amongst published reports and further facilitate optimization of experimental design. In this study, an *in silico* stochastic mouse model is constructed using the Chemical Master Equation (CME) [228]. Here, the accumulation of point mutations in mtDNA is simulated to arise as a consequence of what is believed to be a minimal process required for the maintenance of mtDNA integrity.

3.2 Methods

3.2.1 Chemical Master Equation

Consider a well-stirred constant volume Ω , which is in thermal equilibrium at some constant temperature T , with N chemical species $\{S_1, \dots, S_N\}$ interacting with each other through M reaction paths $\{R_1, \dots, R_M\}$. Let the vector $\mathbf{x}(t)$ denote the number of molecules of each chemical species in the system, where $\mathbf{x}(t) = \{x_1(t), x_2(t), \dots, x_N(t)\}$, given that the system was in state $\mathbf{x}(t_0)$, at some initial time t_0 . The changes in the molecular content of different species in a systems as consequence of a chemical reaction, is characterized mathematically by two quantities. The first is the change in state vector $\boldsymbol{\nu}(t) = \{\nu_{1j}, \dots, \nu_{Nj}(t)\}$, where ν_{1j} is the change in the molecular population of S_i caused by any one R_j reaction fires. The other entity characterizes the propensity function of a reaction such that,

$$a_j(\mathbf{x}) \cdot dt = \text{probability, given } \mathbf{x} \text{ that one reaction } R_j \text{ happens in infinitesimal time interval } [t, t + dt). \quad (3.1)$$

Equation 3.1, forms the fundamental premise of the stochastic chemical kinetics, since everything else in this theory follows from it via the laws of probability [175]. The physical rationale of **Equation 3.1** can be obtained for the interaction mechanisms of unimolecular and bimolecular reactions. Higher order interactions can be considered as subsets of different unimolecular and bimolecular interactions [214, 229]. If a unimolecular reaction $S_1 \rightarrow \text{products}$ is considered, there exists a *specific reaction probability rate constant*, c_j , such that $c_j dt$ gives the probability that the reaction will occur in the next infinitesimal time dt . Thus, it follows that if a system currently has x_1 molecules of component S_1 , then the probability that one of the molecule will undergo a reaction R_j in next dt is $x_1 \cdot c_j dt$. Thus, the propensity function for the unimolecular reaction is $a_j(\mathbf{x}) = c_j x_1$. Similarly, for a bi-

molecular interaction of the form $S_1 + S_2 \rightarrow products$, the propensity of the bimolecular reaction R_j will occur in the time interval dt is $a_j(\mathbf{x}) = c_j x_1 x_2$. However, for a bimolecular reaction of type $S_1 + S_1 \rightarrow products$, the propensity is $a_j(\mathbf{x}) = c_j \frac{1}{2} x_1 (x_1 - 1)$.

Evaluation of the *specific reaction probability rate constant*, c_j , from the first principles is significantly challenging, requiring specific assumptions about how reaction R_j physically occurs [175, 214]. It has been found that for a unimolecular reaction, c_j , is numerically equal to the reaction-rate constant k_j of the conventional deterministic chemical kinetics. However, for a bimolecular reaction c_j is equal to k_j/Ω , if the reactions involve reactant of different species, and $2k_j/\Omega$, if they are of same species.

Due to the probabilistic nature of Equation 3.1, the exact estimation of the $\mathbf{x}(t)$ is not possible and one has to infer the probability of states as $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$, referred to as the Chemical Master Equation (CME) [175, 228]:

$$\frac{\partial P(\mathbf{x}, t | \mathbf{x}_0, t_0)}{\partial t} = \sum_j a_j(\mathbf{x} - \boldsymbol{\nu}_j) P(\mathbf{x} - \boldsymbol{\nu}_j, t | \mathbf{x}_0, t_0) - a_j(\mathbf{x}) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \quad (3.2)$$

The density function $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$, denotes the probability that the system assumes the state configuration \mathbf{x}_j at time t , given the initial condition \mathbf{x}_0 at time t_0 . A close inspection of Equation 3.2 reveals that the CME is actually a set of coupled ODEs, with each ODE characterizing the dynamics of every possible combination of reactant molecules. It is therefore not surprising that the CME can only be solved analytically for limited simple cases [174, 175].

3.2.2 Gillespie's algorithm

Since the CME (Equation 3.2) can rarely be solved analytically, the rationale behind stochastic simulations is to generate several independent realizations of the system dynamics and study the resulting statistics. The key is to generate simulated trajectories of $\mathbf{x}(t)$ is a joint probability distribution function $p(\tau, j | \mathbf{x}, t)$ [214, 229], of two random variables of time to

next reaction (τ), and the index of next reaction (j), given that the system is currently in state $\mathbf{x}(t)$. The formulation of this joint probability density function in terms of different reaction propensities can be obtained as [229]:

$$p(\tau, j | \mathbf{x}, t) = a_j(\mathbf{x}) \exp(-a_0(\mathbf{x}) \cdot \tau) \quad (3.3)$$

where, the total propensity $a_0(\mathbf{x})$ is defined as,

$$a_0(\mathbf{x}) = \sum_{j=1}^M a_j(\mathbf{x}) \quad (3.4)$$

Equation 3.3 forms the basis of the stochastic simulation algorithm (SSA) [175, 229]. Most commonly used Monte Carlo approach for generating samples of τ and j , according to the joint distribution $p(\tau, j | \mathbf{x}, t)$, is the Gillespie's direct approach [175, 229]. This approach is based on the standard inversion method obtained using the Monte Carlo Theory [175].

In the Gillespie's Direct method, two random numbers r_1 and r_2 are drawn from the uniform random number distribution and the following transformation are done to obtain the two counters τ and j :

$$\tau = \frac{1}{a_0(\mathbf{x})} \cdot \ln\left(\frac{1}{r_1}\right) \quad (3.5)$$

and,

$$j = \text{the smallest integer satisfying: } \sum_{j'=1}^j a_{j'}(\mathbf{x}) > r_2 \cdot a_0(\mathbf{x}) \quad (3.6)$$

Thus, the Gillespie's SSA can be summarized as:

1. Initialize the systems at $t = 0$ with rate constants c_1, \dots, c_M and initial molecular population of each species, x_1, \dots, x_N .
2. For each reactions, $j = 1, \dots, M$, calculate the propensity functions ($a_j(\mathbf{x})$) based on the current state of the system $\mathbf{x}(t)$.
3. Calculate the total propensity function $a_0(\mathbf{x}) = \sum_{j=1}^M a_j(\mathbf{x})$.

4. Sample time to next reaction, τ according to Equation 3.5.
5. Increment the simulation time as, $t = t + \tau$.
6. Sample the next reaction index, j as a discrete random quantity with point probability estimated according to Equation 3.6.
7. Update the current state, according to reaction j , i.e. update $\mathbf{x} \Rightarrow \mathbf{x} + \nu_j$.
8. If $t < T_{max}$, return to step 2.

3.2.3 *In silico* mouse model

The *in silico* mouse model accounts for the accumulation of mtDNA point mutations across two stages of mouse life: development and postnatal (Figure 3.1). In this study, the number of wild-type mtDNA (W) and mutant mtDNA (M) molecules were tracked for each cell in whole mouse heart ($\sim 2.5 \times 10^7$ cells) and liver tissues ($\sim 4 \times 10^8$ cells) [230]. Each mutant mtDNA molecule was assumed to contain only a single mutation in the *TaqI* recognition site (TCGA), following the RMC experimental design [116]. The probability of finding two or more mutations at the same site was assumed to be negligible [141]. Similar the earlier minimal stochastic models [209, 210], the model simulated two mtDNA-related maintenance processes: mitochondrial turnover, comprising of relaxed replication and degradation of mitochondria, and *de novo* point mutation arising during mtDNA replication, based on a minimal conservative assumptions. First, the mtDNA population of each cell was assumed to exist as a well-mixed pool due to fast fusion and fission dynamics of mitochondria [34]. Second, due to the low overall mutation burden, point mutation burden was assumed to remain below the level of functional significance (i.e. no nuclear retrograde signaling [209, 231]). While the latter assumption is conservative, simulations indicate that the incorporation of functional effects of mutations into the model, by assuming that mutant mtDNA are functional and cells respond to a decrease in the number of

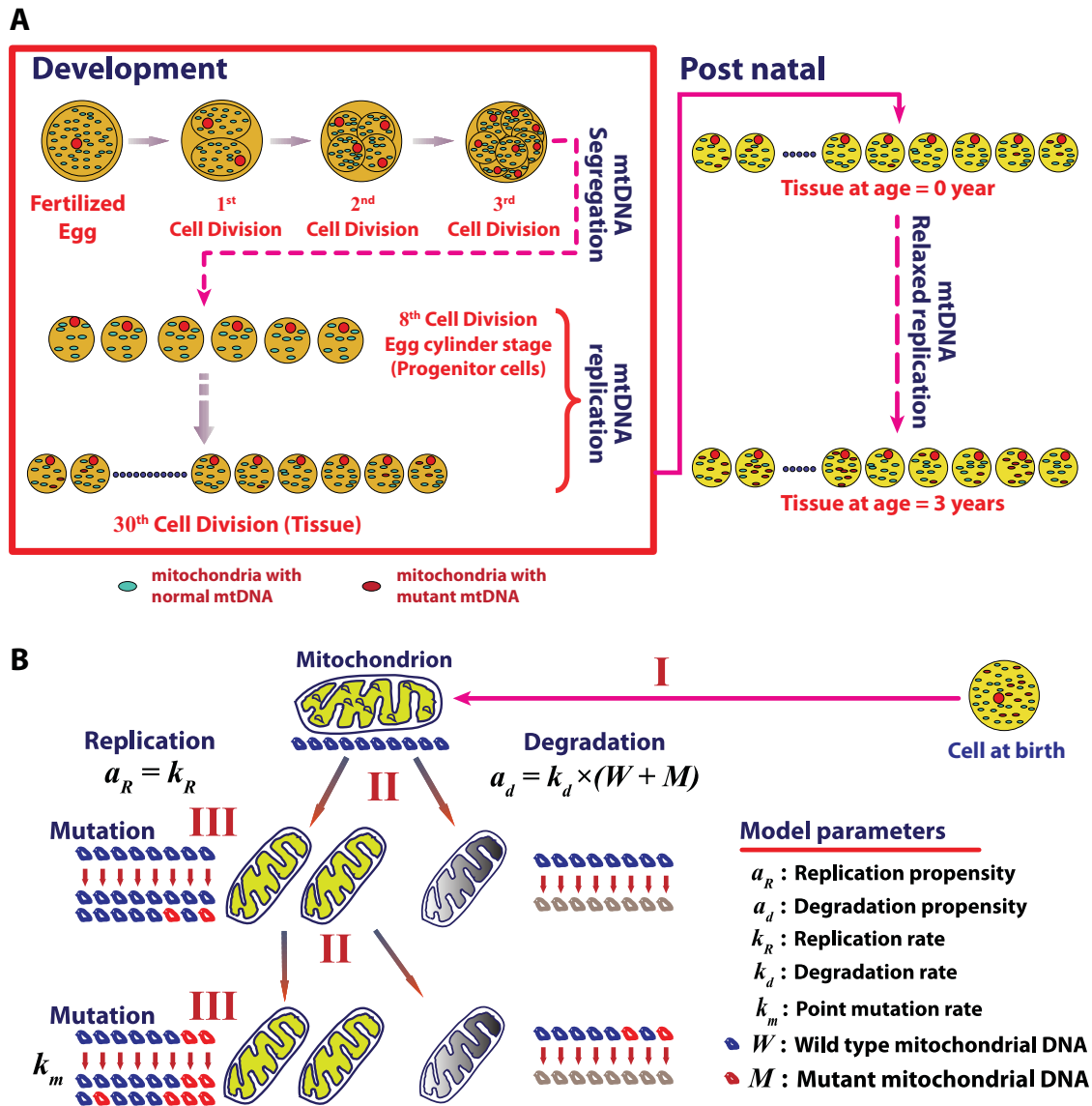


Figure 3.1: **Stochastic mouse mtDNA turnover model.** (A) The *in silico* mouse model simulates the point mutation load of mtDNA in cells of a tissue such as heart and liver during development and postnatal stages. (B) Stochastic drift of point mutations in cells results as a consequence of mtDNA maintenance processes. Three sources of randomness are captured: (I) a random selection of a mitochondrion with ten mtDNA molecules from a well-mixed population, (II) a random replication or degradation of a mitochondrion, and (III) random occurrences of *de novo* mtDNA point mutation during replication.

wild-type (WT) mtDNA by increasing replication, did not result in any significant changes to the mutation burden (see Section 3.2.7).

Following experimental evidence, each mitochondrion was assumed to carry 10 mtDNA molecules and these mtDNA were assumed to undergo replication and degradation due to mitochondrial turnover [232]. In a single turnover event (Figure 3.1B), ten molecules of mtDNA were chosen randomly from a well mixed population of mtDNA in a cell and

are either degraded or replicated according to the CME described below. The selection of ten wild-type and mutant mtDNA molecules from the population was described as a hypergeometric random sampling following the probability distribution: [233]

$$f(x) = \frac{\binom{W}{x} \binom{M}{10-x}}{\binom{W+M}{10}} \quad (3.7)$$

where x represents the number of wild type mtDNA chosen for replication or degradation.

De novo point mutation can occur during replication of mtDNA due to mis-pairing associated with ROS-induced mutagenic lesions such as 8OHdG [102] or as random errors arising due to finite POLG fidelity [234]. Consequently, each replication of a wild-type mtDNA had a finite probability, given by the mutation rate constant (k_m), to produce a mutant. Here, the number of *de novo* mutant mtDNA among x mtDNA was randomly chosen from a Binomial distribution: [233]

$$g(y) = \binom{x}{y} \cdot k_m^y \cdot (1 - k_m)^{x-y} \quad (3.8)$$

where y denotes the number of *de novo* mutations resulting from replication of x wild-type mtDNA.

Based on these probabilities, the *in silico* mouse model was formulated as a CME in which each mtDNA-related process: replication without mutation, replication with *de novo* mutations and degradation, was described as a jump Markov process with the following state transitions:



The first two transitions reflect regular replication, the third represents a *de novo* mutation,

and the last pair represent degradation.

A modified version of the SSA was used in this work for simulating *in silico* mice tissues based on CME in (Equation 3.10). The density function $P(W, M; t)$ denotes the probability

$$\begin{aligned}
\frac{\partial P(W, M; t)}{\partial t} = & k_R \cdot \sum_{x=0}^{10} \left\{ \frac{\binom{W-x}{x} \binom{M-(10-x)}{10-x}}{\binom{W+M-10}{10}} \right\} \cdot P(W-x, M-(10-x); t) \\
& - k_R \cdot \sum_{x=0}^{10} \left\{ \frac{\binom{W}{x} \binom{M}{10-x}}{\binom{W+M}{10}} \right\} \cdot P(W, M; t) \\
& + k_R \cdot \sum_{x=0}^{10} \sum_{y=0}^x \left\{ \frac{\binom{W-(x-y)}{x} \binom{x}{y} k_m^y (1-k_m)^{x-y} \binom{M-(10-(x-y))}{10-x}}{\binom{W+M-10}{10}} \right\} \\
& \cdot P(W-(x-y), M-(10-(x-y)); t) \\
& - k_R \cdot \sum_{x=0}^{10} \sum_{y=0}^x \left\{ \frac{\binom{W}{x} \binom{x}{y} k_m^y (1-k_m)^{x-y} \binom{M}{10-x}}{\binom{W+M}{10}} \right\} \cdot P(W, M; t) \\
& + k_d \cdot \sum_{x=0}^{10} \left\{ \frac{\binom{W+x}{x} \binom{M+(10-x)}{10-x}}{\binom{W+M+10}{10}} \right\} \cdot (W+x+M+(10-x)) \\
& \cdot P(W+x, M-(10-x); t) \\
& - k_d \cdot \sum_{x=0}^{10} \left\{ \frac{\binom{W}{x} \binom{M}{10-x}}{\binom{W+M}{10}} \right\} \cdot (W+M) \cdot P(W, M; t)
\end{aligned} \tag{3.10}$$

of a cell to contain W and M number of wild-type and mutant mtDNA, respectively, given the initial conditions of the states (not explicitly stated here for brevity, refer to Equation 3.2). The parameters k_R , k_d and k_m represent the specific probability rate constants for mtDNA replication, degradation and *de novo* point mutations, respectively. The terms in the curly braces describe the hypergeometric sampling of mtDNA from the population.

Particularly, the first two terms of the CME above represent mtDNA replication without mutations, the second pair of terms corresponds to replication with *de novo* mutations, and the last two terms represent the degradation of mtDNA. The CME can be solved numerically using a Monte Carlo approach following the SSA. Consistent with the above section on SSA algorithm, the implementation of the modified SSA is described below:

1. Compute the propensities of replication and degradation processes as a function of W and M at time t .
2. Based on the propensities, generate random samples of (τ, j) as in the SSA algorithm [214].
3. Select ten mtDNA molecules randomly from the population (hypergeometric sampling) for either replication or degradation. Each replication of a wild type mtDNA can result in a mutant mtDNA with a probability given by the mutation rate constant (k_m).
4. Update W and M based on events in steps 2 and 3 and increment the time t by τ .
5. Repeat steps 1 through 4 until the desired end time.

To predict mtDNA mutation burden in a single organ or tissue, millions of such simulations were performed to capture the mtDNA dynamics of all cells in a tissue.

Simulations were performed using an IBM high performance computing cluster with 112 Intel 1.6 GHz processors. The simulation code was compiled using GNU FORTRAN compiler G77 (v4.1.1) and ran on a CentOS Linux platform. On average, a single simulation of a heart tissue (~25 million cells) from development to 3 years of age required approximately 3 hours.

3.2.4 Simulations of mouse development

The embryonic cell divisions begin after fertilization of an oocyte. Mouse oocytes harbor a large number of mitochondria ($\sim 1.5 \times 10^5$ mtDNA) [235], which allow the zygote to multiply initially without the need to replicate mtDNA [236, 237]. Mouse embryos with dysfunctional mitochondrial replication are able to proceed through the implantation and gastrulation stages, but eventually die, presumably when the mtDNA synthesis becomes necessary to maintain ATP level [238, 239]. Furthermore, the total mtDNA number in mouse embryo does not increase until the late stage of blastocyst, which is roughly the 7th to 8th cell divisions in development (i.e., 4.7 to 5.5 days post coitum (d.p.c)) [236–238]. During these stages, mtDNA are segregated among the dividing progenitor cells (Figure 3.1A). Consequently, each progenitor cell of the developing embryo has only few copies of mtDNA at the early egg-cylinder stage [236, 237].

Table 3.1: Model parameters used in the simulations of the *in silico* wild-type mice.

Parameters	Units	Values	Comments	References
W_0	<i>molecules</i>	580	Initial value of wild type mtDNA during start of development	[236, 237]
M_0	<i>molecules</i>	0	Initial value of mutant mtDNA during start of development	
k_d	d^{-1}	2.3377×10^{-3}	Degradation rate of mtDNA	[240]
k_R^{dev}	<i>molecules</i> d^{-1}	465	Maximum replication rate of mtDNA during development	
k_m^{dev}	$rep^{-1} d^{-1}$	1.0×10^{-7}	Mutation rate of mtDNA during development (POLG fidelity)	[137, 234, 241]
N_{cyc}	—	22	Number of developmental cycles	[237, 242–244]
$(W + M)_{ss}$	<i>molecules</i>	3500	Homeostatic set-point of the mtDNA population (Heart cells)	[230, 245]
k_R^{PM}	<i>molecules</i> d^{-1}	0.8182	Maximum replication rate of mtDNA during post natal stage	
k_m^{PM}	$rep^{-1} d^{-1}$	1.6×10^{-6}	Mutation rate of mtDNA during post natal stage (POLG fidelity and Oxidative burden)	[137, 221, 234, 241, 246, 247]
N_{cell}	—	2.2443×10^7	Number of cells (Heart)	[230, 248]
α_{POLG}	—	200	POLG allele fidelity factor	[137, 234, 241]

In order to account for the mtDNA segregation without replication during the initial cell divisions, the developmental simulations started from the end of the 8th stage (5 d.p.c) with an initial wild-type mtDNA count of roughly 580 molecules per cell ($W = 580$, $M = 0$) [236]. Mitochondrial DNA replication is tied to the cellular division to maintain a steady state number of total mtDNA after each division [97]. Mouse development lasts until 20 d.p.c [244] with a doubling time of roughly 15.5 hours [243]. The mtDNA replication rate was estimated assuming that mtDNA doubles its population every 15 hours, while still undergoing degradation. Here, a cell division was simulated to occur when the total number of mtDNA count reached twice the steady state homeostatic count (Table 3.1). The segregation of wild-type and mutant mtDNA between the daughter cells was assumed to occur at random, without any selective advantage according to a hypergeometric distribution: [233]

$$f(x) = \frac{\binom{W}{x} \binom{M}{n-x}}{\binom{W+M}{n}} \quad (3.11)$$

where x in Equation 3.11 denotes the number of wild-type mtDNA in one of the daughter cells after segregation, and n is the total number of mtDNA in a single daughter cell (i.e., $n = (W + M)/2$). During development, polymerase- γ , the care taker of the mtDNA replication fidelity, is the main contributor for point mutations in mtDNA, with negligible oxidative activity and damage [234, 241].

3.2.5 Simulation of postnatal stage

After birth, many tissues like heart do not undergo further cellular division. However, mtDNA in these tissues are still continuously turned over independent of cellular division, a process called "relaxed replication" [232]. The functional significance of relaxed replication in postmitotic tissues like heart and brain is to maintain a healthy population of mtDNA to satisfy the cellular energy requirements [105, 232]. The postmitotic simula-

tions continue from cells produced at the last stage of development (Figure 3.1A), in which each cell maintains mitochondrial biogenesis to balance degradation. The mutation rate in this stage was a summation of contributions from oxidative damage and POLG-related error.

3.2.6 Simulation of POLG mutator mouse models

The *in silico* mouse model is also used to simulate POLG mutator heterozygous ($POLG^{+/mut}$) and homozygous ($POLG^{mut/mut}$) mice by changing the rate of *de novo* point mutations. Mutator mice carry a proofreading-deficient allele of POLG which has 200 times the error rate of the wild-type enzyme [137, 234]. Thus, in the simulations of POLG mutator mice, the model formulation remains the same in all aspects with the exception that the POLG error rate corresponding to the mutant allele was assumed to be 200 times higher (Table 3.2, 3.3). In heterozygous POLG mutator mouse, the replication of mtDNA molecules is carried out by either wild-type or mutant allele with equal probability.

3.2.7 Model parameters

Model parameters are compiled from published data for mice and it was ensured that they are consistent with the current literature and the state of the art measurement techniques. The basic model parameters for wild-type mice are listed in Table 3.1, while more detailed information on the rest of parameters of the POLG mutator mice is given in Tables 3.2 and 3.3.

Mitochondrial DNA turnover: Defective mitochondria are believed to be eliminated to maintain cellular homeostasis [250, 251]. Autophagy has been suggested to be the main pathway for the mitochondrial turnover; this process is also termed as mitophagy. Autophagy is the process of catabolism of cellular components such as cytosol organelles and protein aggregates by a double-membrane structure known as the autophagosomes [252]. In mammalian cells, it is believed that autophagosomes develop from a special-

Table 3.2: Model parameters used in the simulations of the *in silico* POLG heterozygous ($POLG^{+/mut}$) mice.

Parameters	Units	Values	Comments	References
k_d	d^{-1}	2.3377×10^{-3} (Heart)	Degradation rate of mtDNA	[240]
		4.0706×10^{-3} (Liver)		[249]
k_R^{dev}	$molecules\ d^{-1}$	465 (Heart)	Maximum replication rate of mtDNA during development	
		530 (Liver)		
$k_{m,WT}^{dev}$	$rep^{-1}\ d^{-1}$	1.0×10^{-7}	Mutation rate of wild-type mtDNA allele during development (POLG fidelity)	[137, 234, 241]
$k_{m,mut}^{dev}$	$rep^{-1}\ d^{-1}$	$1.0 \times 10^{-7} \times 200$	Mutation rate of POLG mutant mtDNA allele during development (POLG fidelity)	[137, 234, 241]
N_{cyc}	—	22	Number of developmental cycles	[237, 242–244]
$(W + M)_{ss}$	molecules	3500 (Heart)	Homeostatic set-point of the mtDNA population (Heart cells)	[230, 245]
		4000 (Liver)		
k_R^{PM}	$molecules\ d^{-1}$	0.8182 (Heart)	Maximum replication rate of mtDNA during post natal stage	
		1.6262 (Liver)		
$k_{m,WT}^{PM}$	$rep^{-1}\ d^{-1}$	1.6×10^{-6}	Mutation rate of wild-type mtDNA allele with oxidative error	[137, 221, 234, 241, 246, 247]
$k_{m,mut}^{PM}$	$rep^{-1}\ d^{-1}$	2.17×10^{-5}	Mutation rate of POLG mutant mtDNA allele with oxidative error	
N_{cell}	—	2.2443×10^7	Number of cells (Heart)	[230, 248]
	—	4.1871×10^8	Number of cells (Liver)	

ized double-membrane extension of the endo-plasmic reticulum (ER), known as the omega-somes [253]. The double membrane, also known as the isolation membrane, surrounds the cytosolic cargo and its edges seal to sequester the contents into the autophagosomes. This then fuses with a lysosome, allowing the degradation of both the cargo and inner bilayer of the double autophagosomal membrane, which can then be reused or catabolized for the energy synthesis.

Two types of macro-autophagic process have been identified. A.) Non-selective autophagy, which occurs on nutrient deprivation to supply cells with essential metabolic components and energy until nutrient source is restored to cellular milieu. B.) By con-

trast, cargo-specific autophagy occurs when the cell is prevailing under the nutrient-rich environment. In such conditions, cell triggers the autophagic response, only to mediate the removal of superfluous of damaged organelle and protein aggregates in the cells. It is believed that the damaged mitochondrial is removed from the mitochondrial population by a cargo-specific autophagic process known as mitophagy [251].

In yeast [254] and mammalian cells [255] mitophagy is preceded by mitochondrial fission [256], which divides the elongated mitochondria into pieces of manageable size for encapsulation and subsequent mitophagic process. Beyond quality control, mitophagy is also believed to be required for steady turnover of mitochondrial population, for adjustment of mitochondrial number to account for the changing cellular metabolic requirement [257, 258]. Mitophagy process requires specific labeling of the mitochondria and their subsequent recruitment for the isolation membrane. For example, in a well-studied case of yeast mitophagy, it is believed that this process occurs when the outer mitochondrial protein autophagy-related 32 (Atg32), binds with the isolation membrane protein Atg8 [259, 260]. Given this understanding of basic mechanism of label induced targeting of mitophagy process, it is still largely unclear of how mitochondria are specified for removal to calibrate the mitochondrial population density according to the cell's metabolic demand.

Mitochondrial biogenesis in conjunction with elimination of mitochondria by autophagy is necessary to regulate the changes in steady-state mitochondria number that are required to meet cellular metabolic demand. Since the present dissertation is primarily concerned with maintenance and replication of mitochondrial DNA, replication mechanisms associated with the mtDNA is delineated in the subsequent sections.

Past studies have demonstrated that the mtDNA synthesis can occur independent of the cell cycle process [261, 262]. The mtDNA are replicated more than once or not at all during the cells life, yet, mtDNA copy number is maintained roughly constant throughout the generations of cell growth and divisions [97, 99]. Mitochondrial DNA replication re-

quires a large number of factors for maintenance of genetic replication process. The exact mechanisms involved in the mtDNA replication process are still largely unknown. However, it is widely believed that there are about six nuclear encoded proteins involved in the mtDNA replication process. Transcription process is essential for the mtDNA replication process, and thus a RNA primer is necessary to initiate the mtDNA replication process. The mitochondrial transcription factor A (mtTFA) and RNA polymerase are primary elements involved in the mtDNA maintenance and replication. The processing of the RNA primer is carried out by the RNA processing complex (RNase MRP). Endonuclease G also participates in the RNA primer processing. Finally, the replication of mtDNA by polymerase- γ (POLG) is assisted by mitochondrial single stranded DNA binding protein (mtSSB). The presence of all these factors are believed to be essential for the maintenance of mtDNA replication process, and the deficiency in any of the aforementioned factors could essentially disrupt the mtDNA replication process [97, 99].

Experimental evidence indicates that some of the aforementioned factors are critically regulated in response of mtDNA copy number. For instance, POLG-B is altered in response to changing mtDNA copy number, suggesting the potential role of POLG-B in mtDNA copy number regulation [97]. Furthermore, it has also been widely thought that the level of mtDNA is maintained based on the ATP deficit in the cell [97, 105]. An increase in the mitochondrial biogenesis is induced as a consequence of adaptation to impaired metabolism or oxidative stress, such as those associated with aging and degenerative diseases [97]. Partial depletion of mtDNA or impairment of mitochondrial metabolism have indicated an enhanced expression of factors of both the mitochondrial and nuclear origin [263].

The proteins associated with the mtDNA replication process are packed tightly in structures known as nucleoids, which are also associated with the inner mitochondrial membrane. In the last few years significant progress has been made in identifying the

composition of these structures. Several studies have indicated that yeast nucleoids includes proteins that bind DNA and are associated with replication and transcription such as mtTFA, helicase TWINKLE, POLG and mtSSB and several other proteins associated with the mitochondrial metabolism [98]. mtDNA POLG and mtTFA are abundantly and uniformly distributed throughout the cellular cytoplasm. The mtDNA replication is believed to preferentially synthesize in the peri-nuclear region of the cell, due to immediate recruitment of nuclear encoded mtDNA replication machinery. Alternatively, there may exist a nuclear tethering mechanism, in which some of the structural entity inherent to the nuclear membrane may be present for maximizing the transport efficiency of newly synthesized mitochondrial transcripts [99, 262].

Mitochondrial DNA degradation rate (k_d): Cellular organelles like mitochondria are normally degraded by the autophagy process, where an entire organelle is engulfed by a lysosome and undergoes lytic degradation [264]. The half-life of mouse mtDNA molecules can be studied *in vivo* using isotopic deuterated water $^2\text{H}_2\text{O}$ [240]. The decrease of isotopic deoxyadenosine in mtDNA after discontinuation of $^2\text{H}_2\text{O}$ treatment can be used to determine the turnover of mtDNA [218, 240], providing a highly sensitive measurement of mtDNA degradation rate constant for the model, k_d .

Hepatocytes of liver are mitotically quiescent and stop differentiating at the end of the postnatal growth period (~60 days in the rats) [249, 265]. While under normal conditions these cells have a very long life span (~ 400 days) [249], they can become mitotic in response to hepatic stress or injury [249, 265]. Thus, in simulating the liver tissue, the slow cellular turnover was approximated using an elevated mtDNA turnover (Tables 3.2, 3.3).

Mitochondrial DNA replication rate (k_R): The mtDNA copy number is maintained during cell growth and divisions [99]. The mtDNA replication should occur to balance the degradation. There exist evidence supporting the existence of a retrograde signaling between mitochondria and nucleus to regulate the mtDNA content based on cellular bioen-

ergetics [231]. This suggests that mitochondrial biogenesis may be initiated as soon as the mtDNA copy numbers in a cell falls below a certain homeostatic set-point value. In the point mutation model, we have used a constant biogenesis (i.e. without retrograde signaling), but the main conclusions of the work remain the same even with retrograde signaling (see Figure 3.2).

The constant mtDNA replication rate was deduced based on the homeostatic mtDNA copy number in a cell and the degradation rate of mtDNA. Thus, the replication constant k_R is given by:

$$k_R = k_d \cdot (W + M)_{ss} \quad (3.12)$$

where $(W + M)_{ss}$ represents the homeostatic level of mtDNA population in the cell (Table 3.1).

In the second model of mtDNA biogenesis, a Hill-type cooperative equation was used to simulate the retrograde signaling. The Hill-type equation has been widely used in the modeling of biological system involving switch-like behavior arising from cooperativity of enzymes [245]. The functional effect of mutations was simulated as a retrograde signaling by means of a Hill-type kinetics, insofar as the Hill-feedback responds to a drop in the number of wild-type (and hence functional) mtDNA (Equation 3.13 below). In other words, a mutant mtDNA was considered to be entirely dysfunctional. Notice that the maximum rate of replication by the Hill-type kinetics is twice that of the constant biogenesis parameter. Using the Hill-feedback, our simulations showed that wild-type mtDNA population in a cell rarely drops large enough to trigger an increased replication by this feedback and the simulations of the two model assumptions were in agreement as shown in the Figure 3.2. The parameters of the Hill-type equation, K_H and n can be used as a set point level of mtDNA numbers and the sharpness of the switch response, respectively. The replication propensity is then a composite of: a.) a maximum replication rate, balancing the degradation rate of mtDNA, and b.) a negative Hill term capturing the retrograde

response of the nucleus. Hence, the replication propensity for replication of mtDNA is given by:

$$k_R = \nu_{\max} \cdot \left(1 - \left\{ \frac{W^n}{K_H^n + W^n} \right\} \right) \quad (3.13)$$

where ν_{\max} is the maximum replication rate computed by,

$$\nu_{\max} = \frac{k_d \cdot K_H}{0.5} \quad (3.14)$$

Due to the low frequency of the point mutation, the difference between the two replication models is negligible (Figure 3.2). The constant biogenesis replication model was used in the present point mutation model, as this minimizes the number of model parameters and assumptions.

Mitochondrial DNA point mutation rate (k_m): *In vivo*, 8OHdG level ranges from 0.3 to 4.2 lesions per 10^6 bases in nuclear DNA [221, 246, 247]. However, such lesions make only about 10 to 20% of the complete damage spectra [116, 266]. Therefore, the actual frequency of point mutation rate may be as low as 1.5 and as high as 42 lesions per 10^6 DNA bases per replication. In this work, we have made a conservative assumption that the oxidative

Figure 3.2: **Results of choice of mtDNA replication model.** Effect of different choices of point mutation model on the average mutation burden. The average mutation frequency reported in the plot represents the mutation burden in the heart tissues of wild-type mice and was recorded at the end of 36 months. Comparison of the average mutation burden obtained using two different replication models: (i) constant biogenesis, and (ii) biogenesis with the Hill-type kinetics.

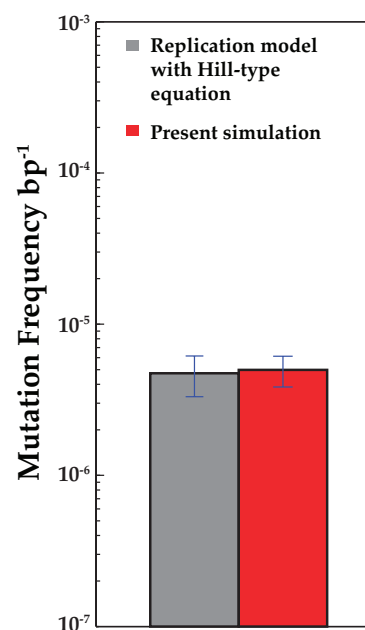


Table 3.3: Model parameters used in the simulations of the *in silico* POLG heterozygous ($POLG^{+/mut}$) mice.

Parameters	Units	Values	Comments	References
k_d	d^{-1}	2.3377×10^{-3} (Heart)	Degradation rate of mtDNA	[240]
		4.0706×10^{-3} (Liver)		[249]
k_R^{dev}	$molecules\ d^{-1}$	465 (Heart)	Maximum replication rate of mtDNA during development	
		530 (Liver)		
k_m^{dev}	$rep^{-1}\ d^{-1}$	$1.0 \times 10^{-7} \times 200$	Mutation rate of mtDNA during development (POLG fidelity)	[137, 234, 241]
N_{cyc}	—	22	Number of developmental cycles	[237, 242–244]
$(W + M)_{ss}$	molecules	3500 (Heart)	Homeostatic set-point of the mtDNA population (Heart cells)	[230, 245]
		4000 (Liver)		
k_R^{PM}	$molecules\ d^{-1}$	0.8182(Heart)	Maximum replication rate of mtDNA during post natal stage	
		0.9351 (Liver)		
k_m^{PM}	$rep^{-1}\ d^{-1}$	2.17×10^{-5}	Mutation rate of wild-type POLG allele with oxidative error	[137, 221, 234, 241, 246, 247]
N_{cell}	—	2.2443×10^7	Number of cells (Heart)	[230, 248]
	—	4.1871×10^8	Number of cells (Liver)	

damage to mtDNA was of similar magnitude to nuclear DNA, consistent with our earlier observations [267]. While some reported values of 8OHdG adducts are determined to be an order of magnitude higher than what is found in nuclear DNA [220], our simulations indicate that such a high damage level was unlikely as this will lead to mtDNA mutation burden much in excess of those quantified by RMC assay (Figure 3.12) [116].

In addition to the oxidative damage, the fidelity of polymerase- γ also contributes to *de novo* point mutations during replication. The polymerase is responsible for the replication and proof reading of newly synthesized strands with a reported error rate between 1×10^{-7} and $1 \times 10^{-6}\ bp^{-1}\ replication^{-1}$ for the wild-type enzyme [234]. Therefore, the overall mutation rate was considered as a sum of oxidative damage and POLG-related errors, giving a range of mutation rate between 1.6×10^{-6} and 4.3×10^{-5} mutations per base pair per mtDNA replication. A conservative value (lowest) of $1.6 \times 10^{-6}\ bp^{-1}\ replication^{-1}$ was chosen for wild-type mouse simulations.

3.2.8 Computation of mtDNA mutation frequency

The point mutation burden (mutation frequency) per base pair is determined using,

$$\Delta_{sim} = \frac{M_{tot}}{(W_{tot} + M_{tot}) \cdot 4bp} \quad (3.15)$$

where W_{tot} and M_{tot} are the total number of wild-type and mutant mtDNA molecules in the tissue, respectively. The length of *TaqI* recognition site used in the RMC assay is 4 bp [116]. Consistent with the original work [116], the probability of a molecule harbouring two or more mutations in the same *TaqI* site was considered to be negligible.

3.3 Results & Discussion

3.3.1 Statistical features of the RMC assay.

In silico wild-type (WT) mouse population of 1100 individuals was generated starting from embryo up to three years of age, the approximate life span of mice (Figure 3.1). The overall point mutation frequency in $\sim 2.5 \times 10^7$ cells of whole heart tissues was recorded at the end of each cell division during development and every fortnight during the postnatal stage. Figure 3.3 illustrates the percentile and distribution function of the mutation frequency arising from two important sources of variability related to the quantification of mtDNA point mutations. The probability density functions indicate the distribution of mutation frequencies in the population as a function of time. Each contour on the percentile plot represents the maximum mutation frequency of a given percentage of the population harbors (e.g. 99% of mice harbor mutation frequencies up to and including the level indicated by the 99th percentile curve).

The main source of randomness is the intrinsic stochastic nature of the aging process, which arises from the mtDNA maintenance processes (Figure 3.1B). Note that the intrinsic stochasticity prevailing in the *in silico* population has a long tailed non-Gaussian density

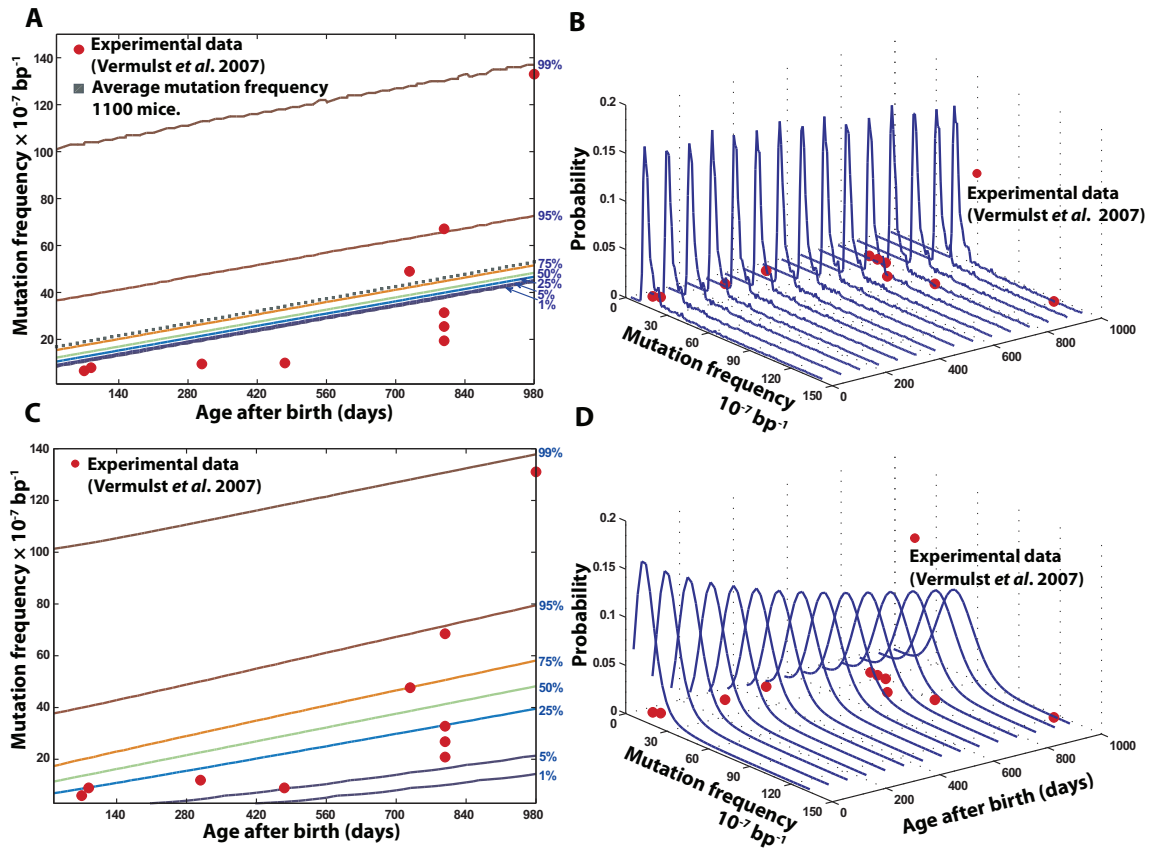


Figure 3.3: **Stochastic determinants of age-dependent dynamics in the observed mtDNA point mutation frequency.** Heart tissue simulations provided the distribution of mutation frequency among 1,100 *in-silico* wild type mice. (A, B) The percentiles and probability distribution functions of the mutation frequency arising from the intrinsic stochasticity of cellular processes alone. The dotted line indicates the evolution of the average mutation frequency of 1,100 mice, which grows linearly with time. (C, D) The percentiles and probability distribution functions of the mutation frequency in the RMC assay of *in silico* wild-type mouse population. The apparent variability arises from the combined effect of intrinsic stochasticity and the (hypergeometric) sampling variability in the RMC protocol (details in the main text).

function (Figure 3.3), indicating that a small fraction of the population harbors a significantly higher mutation burden. Cell-to-cell variability of mtDNA mutation load is also observed as a result of the random processes (Figure 3.4). Figure 3.5 illustrates the evolution of mtDNA states (W and M) in two cardiomyocytes during the postnatal stage of a mouse. Random fluctuation of wild-type mtDNA can be seen in the population with regular bursts and decay of mutant mtDNA. Furthermore, it is interesting to observe that despite the significant cell-to-cell variability of mutation load being large (Figure 3.4), the average accumulation of mtDNA mutation in tissue remains linear after birth (Figure 3.3A). Also, the variance remains roughly constant during the mouse life span, indicating that

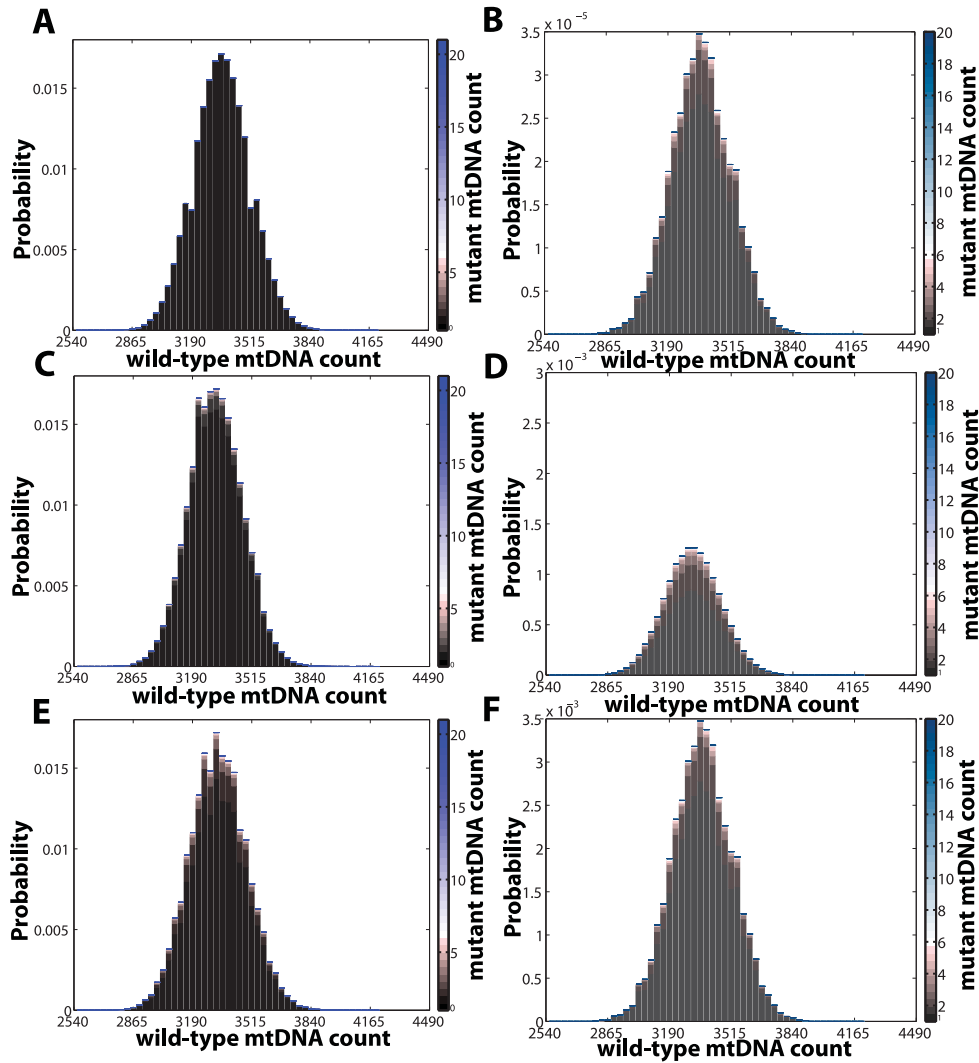


Figure 3.4: Point mutation distribution in cells of heart tissue from different *in silico* mice. Stacked distribution plots of the normal and mutant mtDNA counts (W, M) in an *in silico* mouse heart. Each plot represents the simulation outcome of a single heart tissue. Subplots on the left pane represent the complete distribution of all states in the cells and those on the right pane illustrate the states distribution for all the cells having at least one mutant mtDNA ($M > 0$) (Note, the frequency of the cells having mutant mtDNA in a wild-type tissue is two orders of magnitude lower than the cells from the tissue of *POLG* mutator mice). Dispersion of mutation load in the cells has an increasing trend amongst the three different mouse models with the *POLG*^{mut/mut} mouse having the highest dispersion of mutant mtDNA states in the cells. The stochastic nature of mtDNA turnover has a significant contribution in the mutation load dispersion. (A, B) Distribution of mtDNA states in the tissue of wild-type mouse. (C, D) Distribution of the mtDNA states in the tissue of *POLG*^{+ / mut} (heterozygous) mouse. (E, F) Distribution of the mtDNA states in the tissue of *POLG*^{mut / mut} (homozygous) mouse.

the mtDNA mutation variability among individuals is acquired in the tissue before birth.

However, for comparison with data derived from the RMC assay, a second source of variability has to be considered due to the intrinsic statistical properties of the assay protocol. This is because the determination of point mutation burden by the RMC assay in-

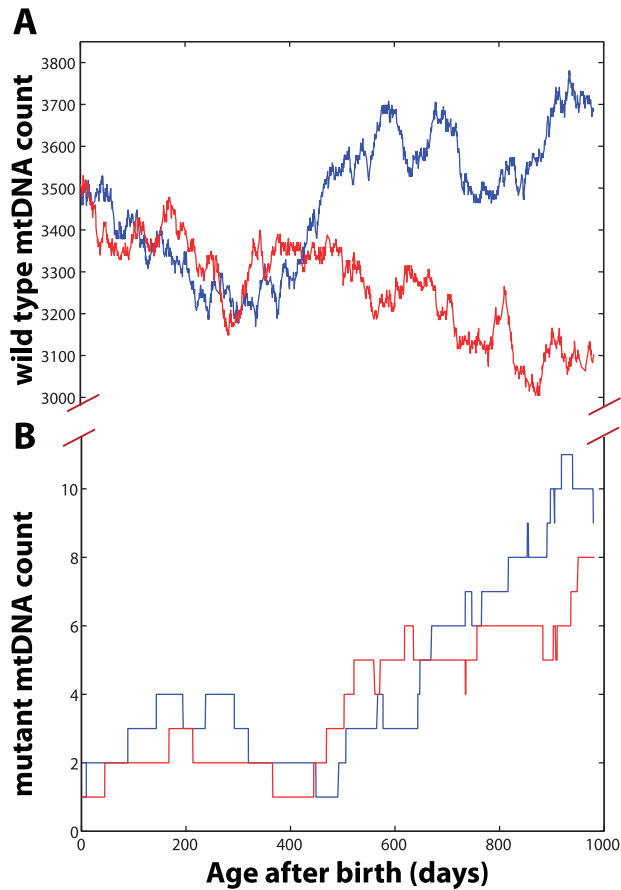


Figure 3.5: **Stochastic evolution of mtDNA states.** (A) represents the stochastic evolution of the wild-type mtDNA, while (B) illustrates the stochastic changes in the mutant mtDNA population. Red and blue curves indicate the outcomes of two independent realizations.

volves drawing a random sample of mtDNA copies ($\sim 840,000$) from tissue homogenates [116]. This sampling procedure introduces additional variability that becomes significant due to the low overall count of total mtDNA mutations. This statistical feature of the RMC protocol can be described as sampling from a hypergeometric distribution [233]:

$$f(m | W_{tot}, M_{tot}) = \frac{\binom{M_{tot}}{m} \binom{W_{tot}}{n-m}}{\binom{W_{tot} + M_{tot}}{n}} \quad (3.16)$$

where m denotes the number of mutant mtDNA molecules present in a random sample of mtDNA of size n ($n = 840,000$ mtDNA molecules in this case). Thus, for low mutation frequencies and sample sizes, the RMC protocol introduces significant additional variability in the data. For example, in heart tissue homogenate containing 10^{10} molecules of mtDNA with a mutation frequency of $10^{-6}/bp$ (a total of 4×10^5 mutant mtDNA), samples of 840,000 mtDNA drawn from the same homogenate will have a mean value of 3.36 mutants with a standard deviation of 1.83 molecules or 54.6% coefficient of variance from

the RMC sampling alone.

The compounded effect of the two sources of variabilities (intrinsic aging related and RMC assay) can be expressed by,

$$\phi(m) = \sum_{W_{tot}, M_{tot}} f(m|W_{tot}, M_{tot}) \times h(W_{tot}, M_{tot}) \quad (3.17)$$

where $h(W_{tot}, M_{tot})$ denotes the underlying probability distribution of mtDNA mutations predicted by the mouse model simulations and $\phi(m)$ is the overall probability function of measured mtDNA mutations. Importantly, the additional variability associated with the sampling of mtDNA in the RMC protocol causes the mutation frequency variance to increase as a function of the average mutation frequency (Figure 3.3C), a result expected from a hypergeometric distribution [233]. This is particularly relevant here because of the age-dependent increase in mean mutation burden and the fact that the distribution describing the mutation process is long-tailed (Figure 3.3A, 3.3B). When this underlying mutation dynamics is sampled using the RMC assay, the resulting data will exhibit an age-dependent increase in variance. Due to low number of samples (typically $n < 5$ per age group), it is highly probable to choose a sample for which the data is best approximated by a non-linear, possibly exponential model (Figure 3.3C). However, this apparent exponential increase is not actually a feature of the underlying mutation dynamics, which may in fact be linear (Figure 3.3A). This has important implications for the interpretation of the available experimental data.

In accordance with the interpretation reached in the original experimental work [116], the variance in the *in silico* data as well as the experimental data for low n -values appears to suggest an exponential dynamics supporting the 'vicious cycle' theory [222, 227]. However, on more careful consideration (Figure 3.3), the apparent exponential increase of the mutational burden could actually be an artifact of: (a) intrinsic stochasticity of aging process (Figure 3.3A, 3.3B), coupled with (b) the random sampling variability introduced by

the statistical properties of the RMC protocol (Figure 3.3C, 3.3D). Experimentally, it is not possible to carry out 100s or 1000s of repeats and it is therefore difficult to distinguish between a truly exponential and a linear increase of age dependent point mutation burden. In summary, while the RMC assay is able to quantify extremely low levels of mutations, its discrete nature (in terms of mutant mtDNA count) introduces significant challenges in data analysis and interpretation. The interpretation of the data can be flawed if the statistical properties of the RMC assay are not considered. Taking both processes into consideration, the fundamental mtDNA maintenance processes modeled by the *in silico* mice were in excellent agreement with the published data (Figure 3.3C). However, the last data point of mutation burden from an old mouse (980 days) deviated from the mean of *in silico* mouse population (p -value = 0.064), suggesting that other processes not predicted by the present model may be involved during the last months of life (e.g., inflammation or other disorders that can accelerate oxidative DNA damage [268]).

3.3.2 Transgenic mouse studies.

Transgenic mouse studies on POLG mutator mouse have recently shed some light on the role of mtDNA in aging [116, 130, 131]. However with these mutator models, many open questions still remain about the role of mtDNA mutation in aging. For example, only the homozygous mutator mice exhibited accelerated human-aging-like phenotypes (e.g., anemia, alopecia, kyphosis) and shortened lifespan, while the heterozygous mice have no obvious aging phenotypes, despite significantly elevated mutation burdens [131].

After successfully validating the *in silico* mouse model against wild-type mouse data, further simulation of 1,100 hetero- and homozygous POLG mouse heart and liver tissues were performed by elevating the baseline POLG error rate to 200 times that of wild-type [137, 234]. In this case also, an excellent agreement of the *in silico* results were found with the reported mutation burdens from two different laboratories [116, 131] (Figure 3.6 and

Figure 3.7). As with the wild-type mice, the point mutation increase was linear with age (Figure 3.7). Again, mitochondrial turnover and *de novo* point mutations alone were sufficient to explain the accumulation of mtDNA point mutations. These results indicate that even at the elevated levels of point mutations ROS-mediated acceleration of point mutations with age was not necessary to explain the data presented in [130, 131]. This is consistent with additional experimental observation suggesting that the levels of ROS in POLG mice are not significantly elevated in the mutator mice [130]. Crucially, no modification of mtDNA maintenance rate constants was required to reproduce the experimental data [130, 131]. That is, one does not have to resort to assumptions such as the existence of a vicious cycle or other possible feedback mechanism [199, 201].

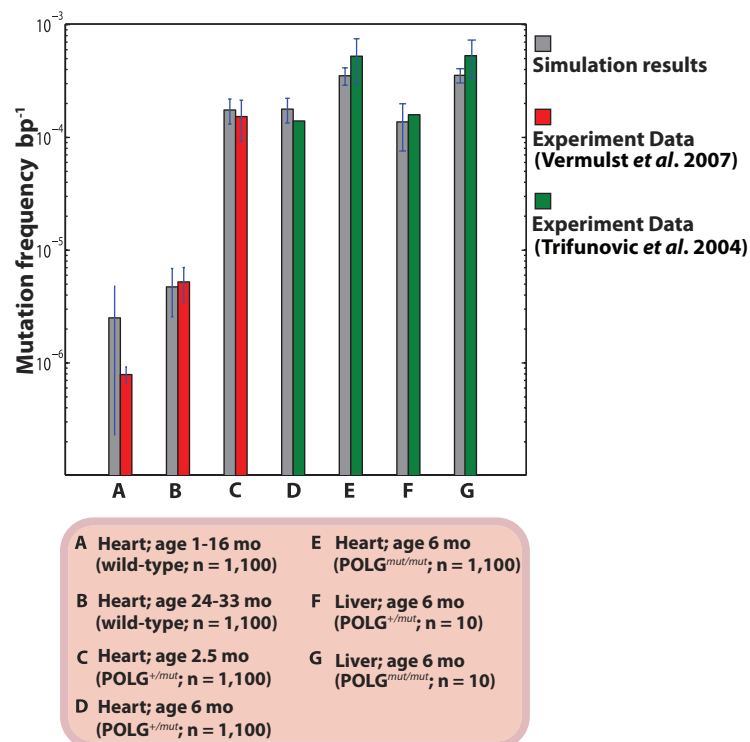


Figure 3.6: Average mtDNA point mutation frequencies in WT and POLG mutator mice. The variances in the *in silico* mouse data represent the intrinsic stochasticity only (without the RMC sampling variability).

3.3.3 Rate of mitochondrial turnover

Since mutations predominantly occur during mtDNA replication as a consequence of either oxidative lesions or intrinsic polymerase errors [133], mitochondrial turnover rate is a critical parameter determining *de novo* mtDNA mutation rate [269]. The mtDNA turnover

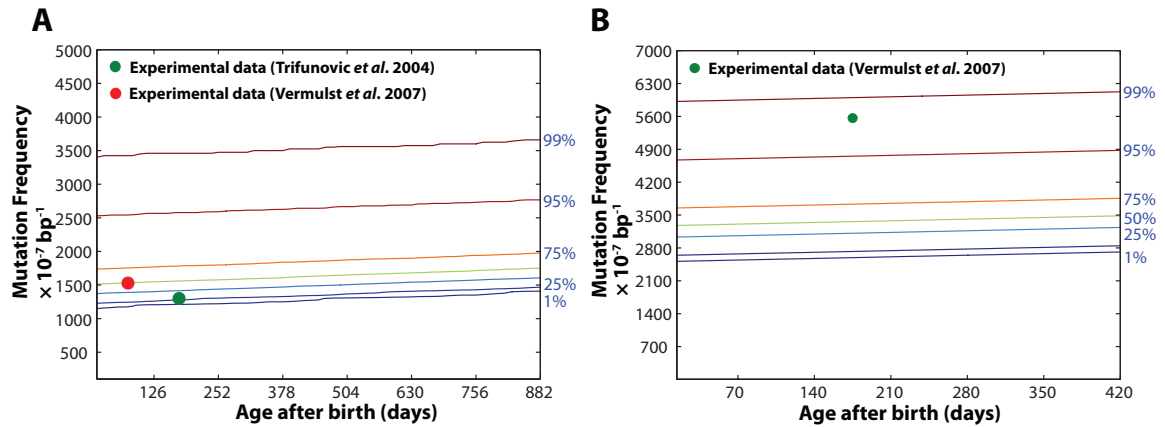


Figure 3.7: Stochastic determinants of age-dependent point mutation dynamics in mutator mice. Comparison of the observed mtDNA point mutation frequency in a population of 1,100 *in-silico* POLG mouse heart tissues (Heterozygous and Homozygous). The mutation frequencies were recorded every fortnight up to 36 months (heterozygous) and 14 months (homozygous). (A) The percentile curves of the mutation frequency in the RMC assay of *in silico* $POLG^{+/mut}$ mouse (heterozygous) population. The apparent variability arises from the genetic variations intrinsic to the aging process and the hypergeometric sampling variability in the RMC protocol (details in the Methods section). (B) The percentile curves of the mutation frequency in the RMC assay of *in silico* $POLG^{mut/mut}$ mouse (homozygous) population. Unlike the wild-type, the uncertainty arising due to the combined effect of the two sources of variability does not increase with time. The variance remains roughly constant with age and this is primarily due to the high point mutation load prevailing in the cells at birth, which only increases relatively marginally with age.

rate may also impacts clonal expansion dynamics of mtDNA mutations [269, 270]. Mitochondria do not generally exist as disconnected organelles but undergo rapid fusion and fission, forming large, complex reticular structures [34]. Therefore, it is not trivial to define what constitutes "turnover of a mitochondrion". Definitions are possible based on turnover rates of its constituents like proteins [271–273], mtDNA [218, 240, 274] or lipids (cardiolipin) [218, 275, 276]. Experimental data on mitochondrial turnover are limited and divergent. Reported half-lives range from 2 to 350 days depending on reference molecules and measurement techniques [218, 240, 273, 274, 276]. Half-life values of about 2-17 days, derived from bulk protein turnover data, are often cited [271, 273, 275]. If these short half-life values are assumed to be indicative of the mtDNA turnover rates, they would have dramatic impact on age-dependent mtDNA mutation accumulation dynamics [201, 209, 210, 212, 269]. Using the modeling analysis, it can be argued that mtDNA turnover cannot be as rapid as is often envisaged based on the protein turnover data.

Mitochondrial protein turnover rates are usually measured using pulse chase experi-

ments [271, 273, 275]. Isotope-labeled protein precursor molecules are administered, followed by tracking labeled mitochondrial protein. The first-order decay rate for the loss of isotope-label yields mitochondrial protein half-lives. However, such data must be interpreted with caution. For one, mitochondria contain approximately 1200 different polypeptides [277], with half-lives ranging from little over an hour to more than 10 days (Table 3.4). The presence of intra-mitochondrial proteolytic systems [278] indicates that mitochondrial proteins are also degraded independent of organelle turnover resulting in protein-specific half-lives [279]. In contrast to proteomics approaches, pulse chase experiments utilizing bulk protein will result in half-lives that reflect the average turnover of the population of proteins in mitochondria [273]. Therefore, protein turnover data from such experiments may be dominated by the turnover rate of few abundant proteins. This suggests that even with the best experimental practices and even when artifacts, such as precursor reutilization [272, 273, 280], are carefully controlled, bulk protein turnover might still overestimate organelle turnover and, in particular, may not be a useful indicator of mtDNA turnover.

The turnover rate of the whole organelle cannot be faster than that of the most persistent macromolecules that the organelle contains. Individual mitochondrial protein half-lives of more than 10 days have been reported (Table 3.4). If these data are accepted, mitochondrial half-life cannot be shorter than 10 days. However, more data from mitochondrial proteomics studies using modern methods would be desirable (e.g.: [281]). Long half-lives have also been reported for other macromolecules in mitochondria. For instance, a half-life of 40 days has been reported for the mitochondrial lipid cardiolipin [276]. Importantly, attempts at direct measurement of mtDNA turnover have yielded half-life values ranging from 10 days up to as low as 300 days, again depending on tissue types and experimental techniques [218, 240, 274]. However, even the highest value of mtDNA turnover (10 days) [218] is significantly longer than the two-day mitochondrial turnover rate estimated based on the bulk protein measurements [273].

Table 3.4: Turnover rates of some mammalian mitochondrial proteins.

Protein	Half life	Organism/Tissue	Reference
Ornithine Aminotransferase	1.9 days	Rat Liver	[282]
Cytochrome Oxidase	5.7 days	Rat Liver	[282]
Carbamyl Phosphate synthetase	7.7 days	Rat	[283]
Malate Dehydrogenase	2.6 days	Rat	[283]
Aminolevulinat synthetase	70 mins	Rat Liver	[284]
Cytochrome c	10.3 days	Rat Liver	[275]
Ornithine transcarbamylase	7.7 days	Rat Liver	[285]
Cytochrome c	11.8 days	Rabbit Heart	[286]
66 Different Mitochondrial Proteins	2 days - 8.5 days	Rat Skeletal Muscle	[279]

The turnover rate of mtDNA determines how fast mtDNA mutations accumulate with age, as each replication event has a finite probability of introducing a new mutation. The mtDNA point mutation burden in postmitotic tissues like mouse heart and brain have recently been reported to be extremely low [116, 131]. If these low levels of mtDNA mutation burden are considered to be true, then these data provide a constraint on the magnitude of the mtDNA turnover rate, because ultimately mtDNA turnover determines the underlying *de novo* mtDNA point mutation rate.

The higher accumulation of mtDNA mutations due to faster mtDNA turnover can be clearly demonstrated using the *in silico* stochastic model of mtDNA maintenance in mouse heart, described above. In these simulations, the mtDNA mutations were assumed to arise as a consequence of the finite fidelity of the mitochondrial DNA polymerase- γ [137, 234] alone, while neglecting the mutations arising from the oxidative DNA damage. This assumption is conservative in the context of this work because it means that the model underestimates actual mutation rates. This model was used to investigate the impact of different mtDNA turnover rates on age-dependent mtDNA point mutation burden in both

wild-type and POLG mutator mice, for which experimental data are previously reported [116, 131]. Simulation results are presented in Figure 3.8.

The model results shows clearly that, as long as neither mutant nor WT mtDNA are preferentially degraded, higher turnover rates always resulted in increased mtDNA mutation burden, since *de novo* mtDNA mutations are tied to mtDNA replication (Figure 3.8). In particular, it was found that the mtDNA mutation burden obtained using short mitochondrial half-life (~ 2 days [273]) was inconsistent with the experimental data [116, 131] (Figure 3.8D). These simulations were repeated using other reported half-lives, ranging from 10 to 300 days [218, 240, 271, 274] and it was further found that a mtDNA turnover

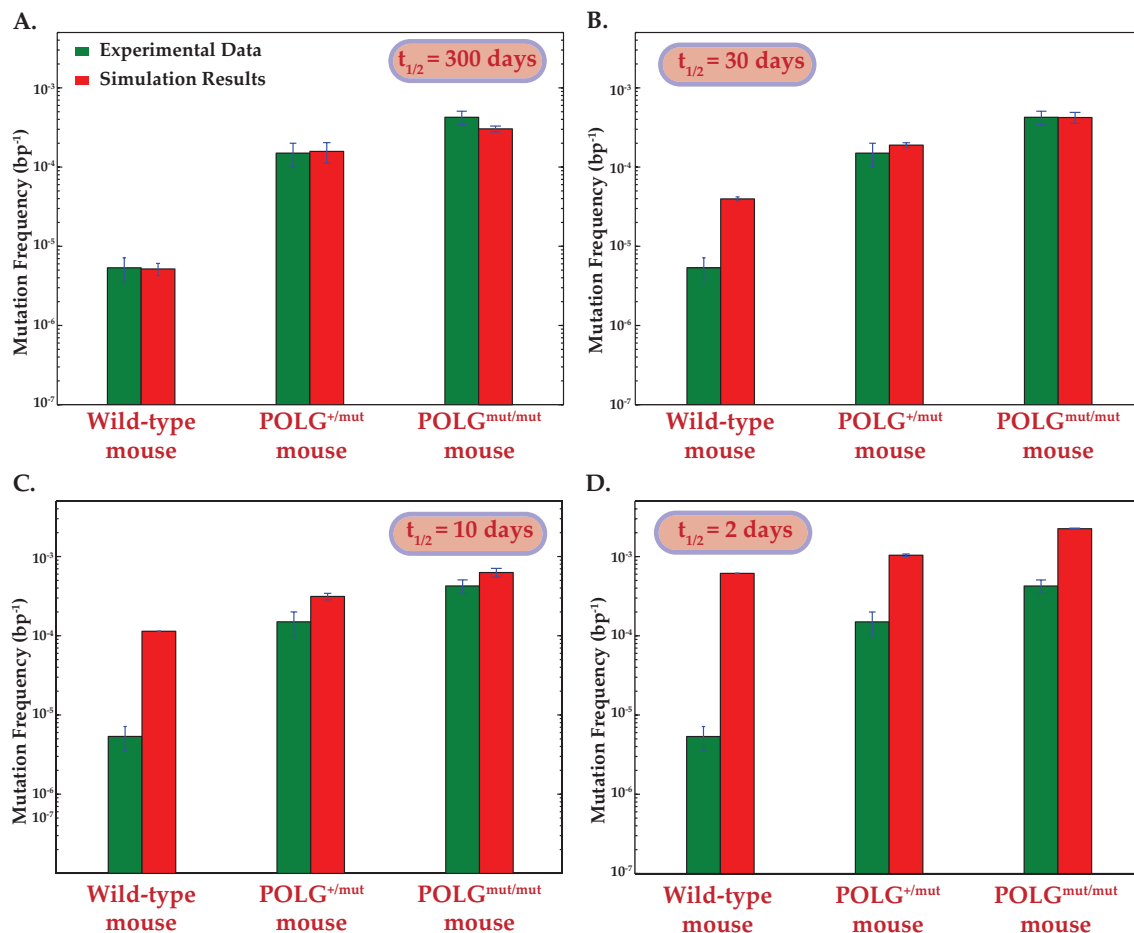


Figure 3.8: **Comparison of mutation frequency (experimental and simulation data) for different mtDNA turnover rates.** Comparison of experimental data and simulation results of point mutation frequency in wild-type and POLG mice ($n = 10$) [116, 131] for different values of mtDNA turnover rates. The frequencies of the accumulated point mutations in mouse heart at the end of 29 months (wild-type mouse) and at the end of 6 months of age for (POLG^{+/mut} and POLG^{mut/mut} mice). Results of point mutation frequencies obtained using mtDNA turnover rates of: A.) $t_{1/2} = 300$ days [240], B.) $t_{1/2} = 30$ days [218], C.) $t_{1/2} = 10$ days [218], and D.) $t_{1/2} = 2$ days [273].

process with a half-life of more than a month is required for having the consistency with the experimentally observed mtDNA point mutation levels.

3.3.4 Significance of animal development.

The stage in an organism's life from which the accumulation of mtDNA mutations starts to become functionally significant (if at all) is unclear. During development, mtDNA replication is tied to the cellular division, and as a consequence, initial mutations may arise as soon as mtDNA replication begins. In fact, the total number of replications during development is comparable to that during the entire adult life. In mice, the heart tissue develops in about 20 days [244]. Considering the degradation rate described in Table 3.1 and the mouse heart to contain $\sim 2.5 \times 10^7$ cardiomyocytes [230, 248] arising from 22 cell divisions (6 progenitor cells), the total number of mtDNA replications needed to maintain homeostatic value of mtDNA (Table 3.1) [230] per cell should exceed 9×10^{10} times during the development. On the other hand, from the degradation rate of mtDNA in postnatal stages (Table 3.1) [240], the number of mtDNA replications events over the three years lifespan of mice is about 1.3×10^{11} . Thus depending on their source (ROS, POLG errors), the development period may carry comparable contributions in *de novo* mtDNA mutations as does the entire adult life. POLG errors have been postulated to be the main cause of *de novo* point mutations in murine embryonic fibroblast [234, 241]. Therefore, the POLG baseline error rate was used as mutation rate during development. Generally, the *in silico* mouse data highlights that mutations occurring in the early embryonic cells have a strong impact on the mutation load at birth (Figure 3.9) and that the variability among individuals is set during development (Figure 3.3, 3.7 and 3.10). Since the mtDNA replication is several folds higher than the degradation during development, *de novo* point mutations generated during the early cell divisions can accumulate very quickly, resulting in a high mutation load at birth in some individuals (Figure 3.9). These results highlight that the

stochastic drift of mutation dynamics during the early developmental cell divisions may be a deciding factor of the organism's mutation trajectory, and also a major contributor of the mutation variability in a population, including isogenetic individuals [185]. The variability generated during development is conserved throughout the organism's life (see Figure 3.3A, 3.7).

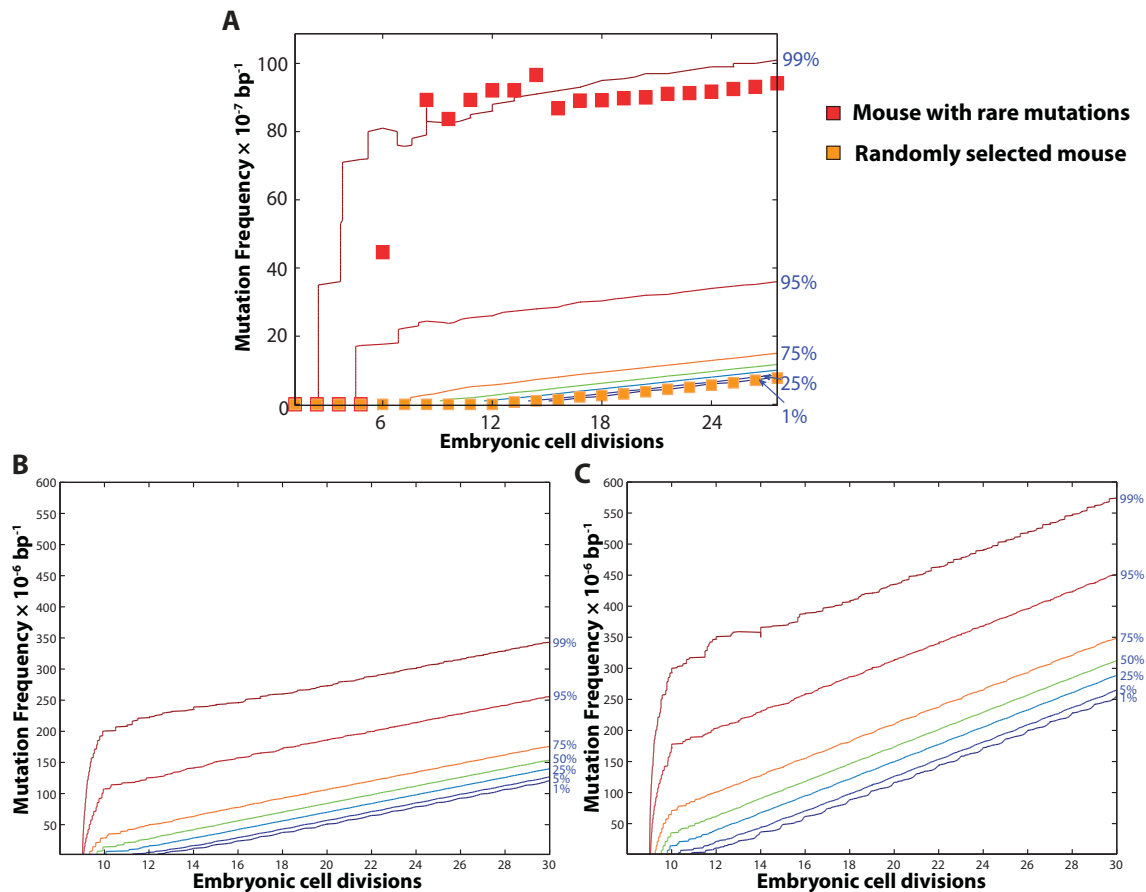


Figure 3.9: Mitochondrial DNA point mutation accumulation during mouse development. Expansion of mtDNA point mutations during heart tissue development from *in silico* wild-type (A), *POLG^{+/mut}* (B) and *POLG^{mut/mut}* mice (C) population ($n = 1,100$). (A) The square symbols show examples of point mutation trajectory from two different mice, one of which suffers from a rare point mutation early in the development, resulting in the amplification of the mutation frequency in subsequent cell divisions. Since replication frequency is several folds higher than degradation during development, *de novo* point mutations generated at early development can accumulate very quickly, resulting in a high mutation load in the tissue at birth. (B) Like in the wild-type cohort, *de novo* point mutations generated in the *POLG^{+/mut}* mice during the early cell divisions can accumulate very quickly, resulting in a high mutation load in the cells at birth. (C) Since the error rate of mtDNA replication in *POLG^{mut/mut}* is much higher than the wild-type mtDNA replication, a significant proportion of the population (>75%) harbors mtDNA mutations at an early stage of development (before the 10th cell division). As a consequence, the resulting mutation load in the tissue is significantly higher than that in the wild-type tissues at birth.

In postmitotic tissues, like heart, mtDNA are continuously turned over independent of cellular division [232]. Although the turnover rate of mtDNA is lower during the post-

natal stage than during development, the higher mutation rate due to oxidative damage (Table 3.1) can lead to 2-3 fold increase in the mutation load between birth and old age in wild-type mice (see (Figure 3.11)). The *in silico* POLG mice however differ from the

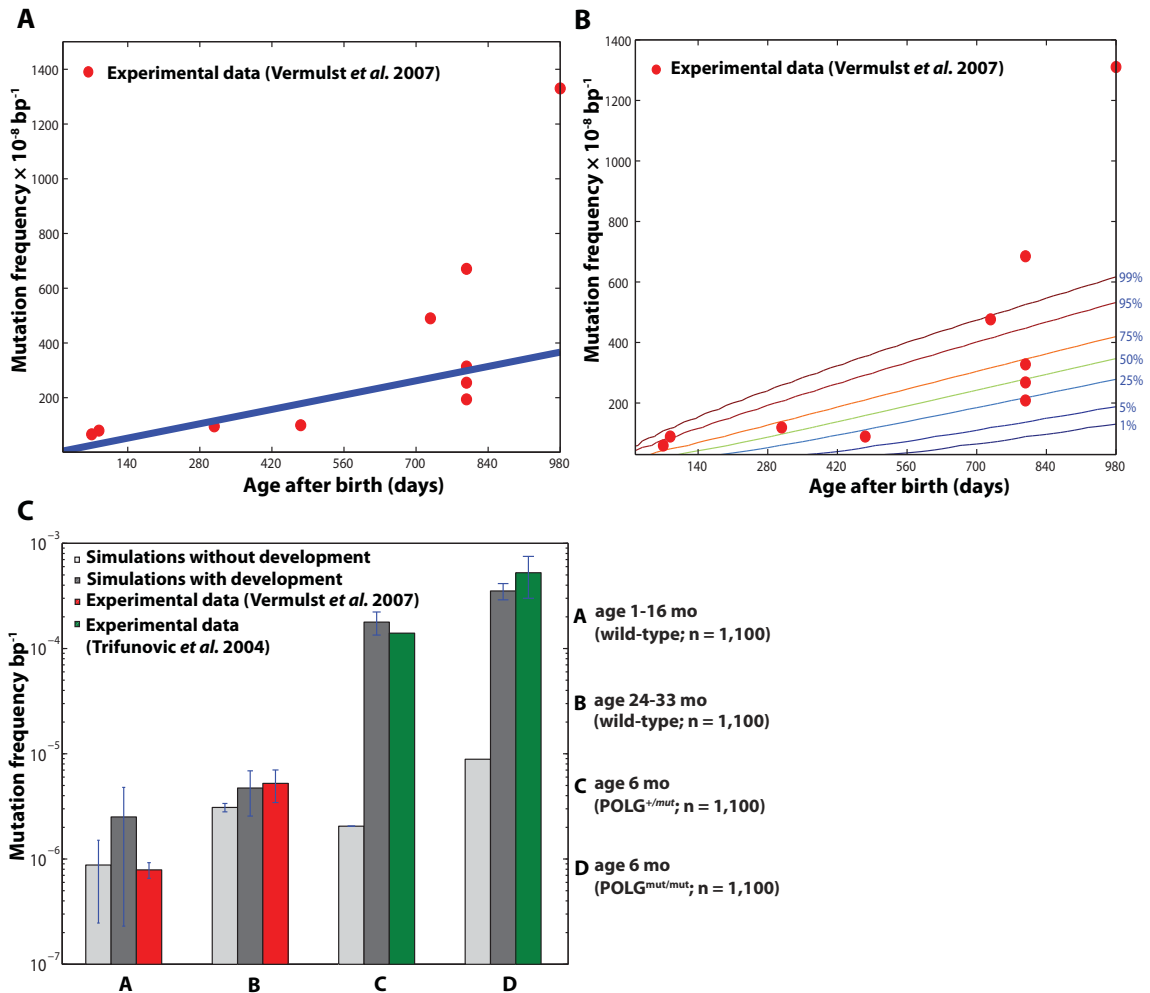


Figure 3.10: Simulations with exclusion of developmental phase. Sources of variability in the observed mtDNA point mutation frequency of 1,100 in *in silico* wild-type mice heart tissues without developmental *de novo* point mutations. The mutation frequencies were recorded every fortnight up to 36 months. (A) The percentile curve of mutation frequency due to the intrinsic stochasticity of aging process in the mouse population. The mtDNA maintenance during life did not cause any observable variability among the mice population, as indicated by the overlapping percentile curves and the sharp distribution. This is in agreement with the previous observation that the genetic variability is inherited from the development and conserved during the adult life. (B) The percentile curve of the mutation frequency in the RMC assay of *in silico* wild-type mouse population. The variability again increases with age due to the increase of the mutation frequency. The model excluding the development was in worse agreement with the experimental data than the trials that included the development. (C) Comparison of the average mutation frequency in mouse heart tissues with and without developmental contribution. Two types of simulations were performed, in which one type included *de novo* mutations during development while the other did not (i.e. no mutations at birth). Although the influence of development in the wild-type mice is rather insignificant at older ages, the exclusion of the developmental stage in the simulations of the POLG mutator mice causes a significant difference in the resulting mutation burden. The variance in the *in silico* mice data represents the intrinsic genetic variability only, without the RMC sampling variability.

wild-type because in these mice, the POLG error is the dominant contributor of *de novo* point mutations, both during embryonic and postmitotic stages (Figure 3.11). Due to faster mtDNA replication (tied to cell division), most of the mutations in mutator mice therefore arise during development (Figure 3.9 B, 3.9 C and Figure 3.11). This is consistent with the experimental data which shows clearly that mutator mice are born with significantly elevated mutation burden [131, 132]. However, during their adult life, the accumulation is relatively lower compared to their development, due to the slow turnover of mtDNA [240]. This further leads to an interesting insight, largely unappreciated in the original work [130, 131, 287], regarding the point mutation load in tissues that remain mitotic (epidermal, stem cells, spleen). Since in POLG mice the point mutation burden of mtDNA is dominated by POLG errors, mutation accumulation in fast dividing cells is expected to be several fold faster than in postmitotic tissues such as heart. This is consistent with the experimental observation in POLG mutator mice, where some of the most prominent pathologies are associated with the fast dividing tissues manifest in the form of alopecia, spleen enlargement and anemia. However it should be appreciated that the later mechanistic hypothesis regarding the mitotic tissue is speculative, because the simulation considered here does not model any fast dividing tissues. Treatment of cell division and selection pressure for mitochondrial turnover might be a promising area of investigation for the future work.

3.4 Conclusions

By thinking carefully about the different sources of stochasticity in each process from early development all the way to experimental sampling, it can be seen that the RMC assay is a major contributor to the overall data uncertainty. In contrast to the original interpretation of the data, the present analysis indicates that the existence of an exponential dynamics in point mutations cannot be inferred with certainty, and thus no contradiction between

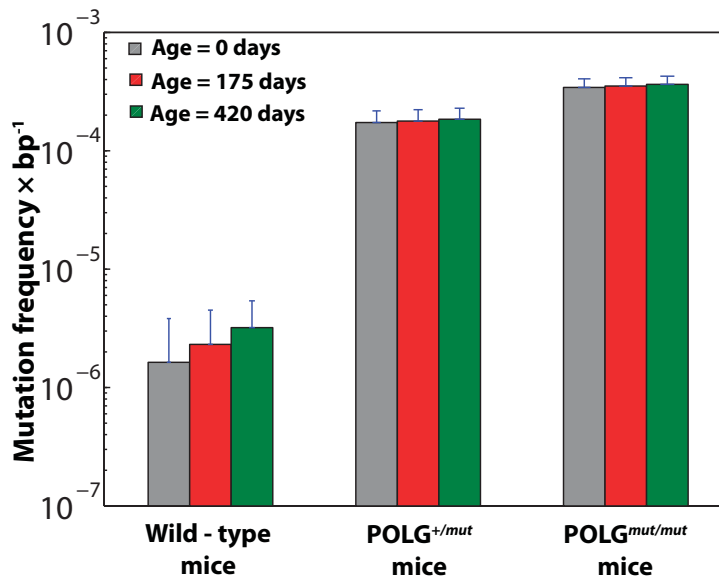


Figure 3.11: Average age-dependent mtDNA point mutation accumulation in wild type and POLG mutator mouse models. Wild-type mice have a low mutation burden at birth, but they accumulate relatively more mutations during their life. On the other hand, the POLG mice harbor a significant mutation load at birth due to error-prone mtDNA replications during development. In the post-mitotic stage however, the relative accumulation (comparing in terms of percentage mutation accumulated during the birth) is significantly lower due to the slow turnover of mtDNA.

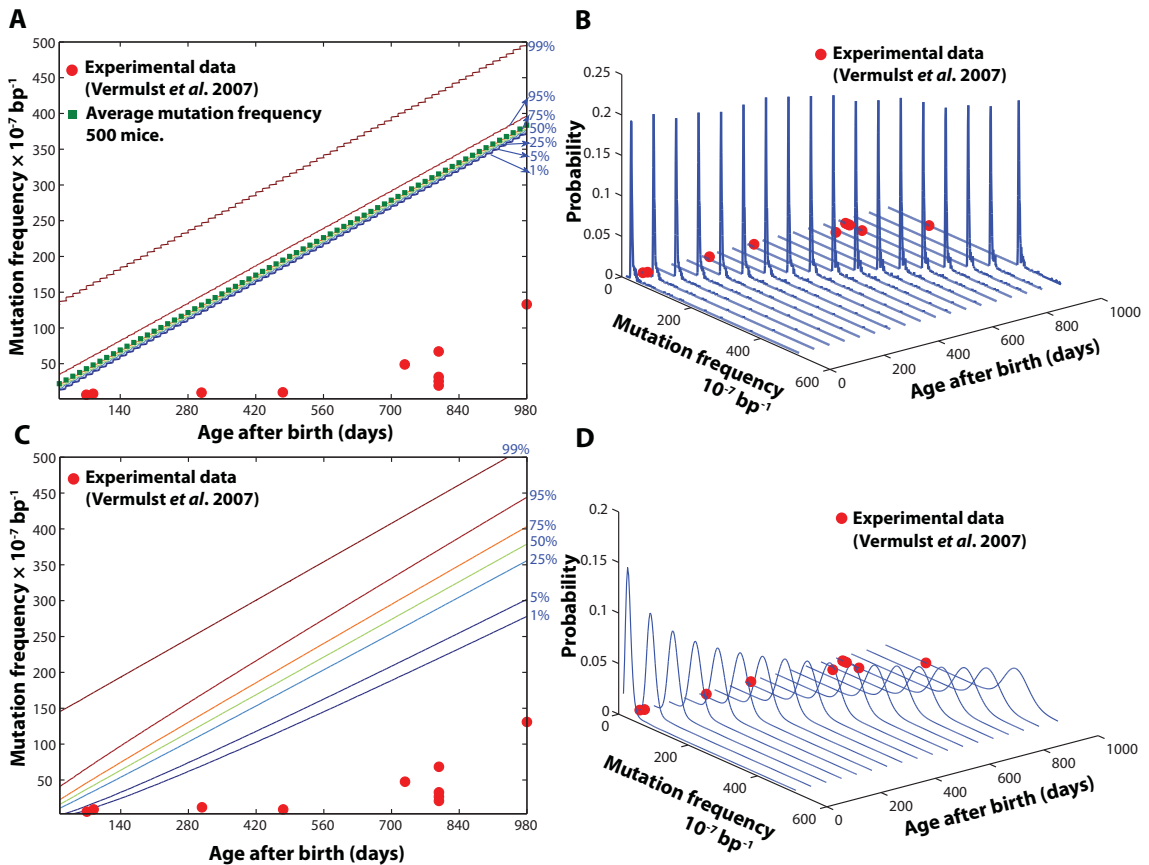


Figure 3.12: Mitochondrial DNA point mutation burden under an elevated oxidative burden assumption. The oxidative burden used in these simulations was elevated to 10 times higher than that used in the present work. The mutation frequencies were again recorded every fortnight up to 36 months. (A, B) The percentile curve and distribution of the mutation frequency due to the intrinsic stochasticity of aging process in the mouse population ($n = 500$ mice). (C, D) The percentile curve and distribution of the mutation frequency estimated by the RMC assay of the *in silico* wild-type mouse population.

the observed point mutation dynamics and the apparent absence of evidence for elevated oxidative stress exists. A detailed, quantitative understanding of the relevant sources of noise also allows optimization of experimental designs, thereby opening avenues for maximizing information return and minimizing cost, time and animal use.

The fact that the reproduction of the POLG mouse data requires no modifications to the wild type model, other than the obvious decrease of the polymerase fidelity, suggests that elevation of the point mutation burden does not trigger fundamentally new processes. In particular, neither mutant replicative advantage nor the elevation of the ROS dynamics resulting from the increase of the point mutation burden is required to explain the POLG data. This is consistent with our current view on the mFRTA [127], showing little evidence for the existence of 'vicious cycle' mechanism. Two further observations related to the POLG mice that have originally been seen as somewhat surprising, can also be explained. The first is the observation that dividing tissues seem to be more severely affected in POLG mice than postmitotic tissues [130, 131, 287]. The second is the fact that most mtDNA mutations in the POLG mice are already present at birth with comparatively little further accumulation during adult life, when compared to its development [131, 132]. Quantitative analysis shows both of these observations to be consequences of the low turnover of mtDNA in postmitotic tissues of adult mice.

Measurements of mitochondrial turnover based on any of its individual components require strong justification and given the dynamical characteristics of mitochondria, the very concept of "whole organelle" turnover may be questionable. Simulation results above indicate that mtDNA half-lives are in the order of months. That is, if (whole) mitochondria organelle turnover parameter is to be at all meaningful, the rate should be much slower than commonly believed (2-10 days) [218, 273]. In this regard there are several studies reporting longer half-lives (>30 days) using cardiolipin, mtDNA and some individual proteins. Additional support for lower mitochondrial turnover rate comes from the

quantification of age-related mtDNA mutations. Using a stochastic computational based on conservative assumptions regarding the mtDNA turnover process and the associated *de novo* mutagenesis, it was found that simulations using a short mtDNA half-life of 2 days grossly overestimated mtDNA point mutations accumulation in murine heart tissue as compared to published values [116, 131]. These experimental data were consistent only with a mtDNA half-life of 30 days or longer. Given the importance of mitochondrial DNA turnover for understanding various aging mechanisms [107, 127, 269], it is an important area worthy of further careful experimental study, preferably by a direct and sensitive experimental estimation of mtDNA turnover. While individual components of mitochondria might turnover rapidly, global mitochondrial turnover rates, if meaningful at all, cannot be faster than mtDNA turnover or turnover of the most persistent macromolecules. Certainly, bulk protein turnover rate should not be used as a surrogate for mtDNA turnover.

Finally, the *in silico* analysis performed in this work reveals the importance of understanding the mtDNA turnover dynamics and the relevance of early development in determining the trajectory of mtDNA mutation burden. This is in sharp contrast to the common assumption that health and diseases are determined predominantly by the genome interacting with the environment. In this work, it has been further demonstrated that *in silico* modeling can contribute significantly to analysis and understanding of experimental data as well as potentially help to design more effective methodology. Furthermore, this approach of "Computer Aided Thought" can contribute towards a fundamentally improved understanding of intrinsically challenging biological problems such as aging.

This chapter highlights the work done in:

- ✦ S. K. Poovathingal, J. Gruber, B. Halliwell & R. Gunawan. Stochastic drift in mitochondrial DNA point mutations: a novel perspective *ex silico*. *PLoS Comp. Biol.*, 5(11): e1000572, 2009.
- ✦ S. K. Poovathingal, J. Gruber, L. N. Lakshmanan, B. Halliwell, R. Gunawan. Is mitochondrial DNA turnover slower than commonly assumed? *Aging Cell*, In review.

Elucidating mechanisms of age-dependent clonal accumulation of mitochondrial DNA mutations — an *in silico* approach

4.1 Introduction

Analyses of post mitotic tissues from both normal aging and individuals with mitochondrial diseases have revealed that a subset of post-mitotic cells in a tissue harbor pathogenic mtDNA mutations in them and the majority of these respiratory deficient cells harbor high levels of clonal mtDNA mutations [27, 154, 288–290], suggesting a clonal expansion of single or few mutational events. In this context an obvious question that arises is: does the mosaic genotype of mutations have to do anything with the aging process? There are extensive evidence available in literature; indicating the association of clonal expansion of mtDNA mutations with the respiratory deficiency of mitochondria and with cellular morbidity, suggesting that some mtDNA mutations play a significant role in the aging process.

Mosaicity of mtDNA mutation burden in tissue can arise from factors such as stochastic drift [209, 210]. Once a pathological mutation has occurred, it initially has low functional importance due to the inherent redundancy of mtDNA copy number in a cell. Subsequently, mutations can reach functionally significant levels (phenotypic threshold; [145]), due to yet poorly understood clonal expansion mechanisms, resulting in a rapid expansion of the mutant mtDNA. Eventually, mutant mtDNA greatly exceed wild type mtDNA population [26]. Characterizing both the etiology of the initial mutation event as well as the stochasticity in the expansion of mtDNA mutation accumulation is critical for understanding the underlying dynamics of mutation accumulation. Understanding these can

provide significant insight towards devising treatment strategies to treat mitochondria related degenerative diseases or retard the progression of aging phenotypes like sarcopenia.

Clonal expansion of mtDNA mutations have been intensely studied in past decades, when it was observed that certain type of mutations out-replicated wild type mtDNA and other types of mutations in a cell [25, 108]. Several mechanistic explanations have been proposed over past decades and some of the frequently cited mechanisms [25, 202, 210, 211] are listed in Table 4.2. Nevertheless, the precise mechanism of mutation expansion is still largely debated and many of these hypotheses have not been tested against the actual experimental data on mtDNA mutations. Since mtDNA deletions are more commonly associated with aging phenotypes, deletion burden data will be mainly used in this work for model validation. The importance of mutation expansion in aging and age-associated mitochondrial diseases calls for more careful study on mutation abundance and mechanisms. Unfortunately, with the existing measurement technologies, it is significantly challenging to obtain relevant measurement indicators for inferring the underlying mechanisms of mtDNA mutant clonal expansion process. In this chapter, different hypotheses are tested using a computational approach, to better understand the mechanisms of the origin and dynamics of clonal expansion of mtDNA mutations. Based on the work in Chapter 3, mtDNA mutation models developed in this chapter is again based on CME and includes specific features of cellular energy logistics in mouse cardiomyocytes and considers only the most relevant biological features of mitochondrial genetic maintenance.

4.2 Methods

Following the stochastic mtDNA point mutation model (Chapter 3; [269]), the *in silico* mouse model developed in this work tracks for mtDNA mutation accumulation in murine cardiomyocytes during two stages of mouse life: developmental and postnatal. The model tracked for wild-type mtDNA (W) and mutant mtDNA (M) in each cardiomyocytes. In

this section, additional modifications of the stochastic model of the mtDNA point mutation accumulation will be discussed.

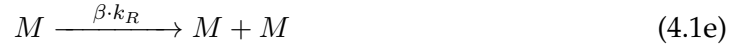
4.2.1 Cell-level modeling details

Consistent with the stochastic model of mtDNA point mutation model (see [Section 3.2](#) in [Chapter 3](#)), the mtDNA turnover process is illustrated in [Figure 4.1](#) and the associated modeling parameters are listed in ([Table 4.1](#)).

The CME formulation [[175](#)] of mitochondrial turnover process ([Figure 4.1](#)) described by the following jump Markov process:

Table 4.1: Model parameters of the stochastic mtDNA turnover process in the cardiomyocytes of *in silico* mouse model.

Parameters	Units	Values	Comments	References
W_0	<i>molecules</i>	1000	Initial value of wild type mtDNA during start of development.	[236 , 237 , 291]
M_0	<i>molecules</i>	0	Initial value of mutant mtDNA during start of development.	
k_d	d^{-1}	2.3377×10^{-3}	Degradation rate of mtDNA	[240]
$\nu_R^{max} _{dev}$	<i>molecules</i> d^{-1}	72293.34	Maximum replication rate of mtDNA during development, obtained using Equation 4.4 , considering $F_M = 0$ and $k_R = k_d \cdot (W + M)_{ss}$	[245]
N_{cyc}	—	22	Number of developmental cycles	[237 , 242–244]
$(W + M)_{ss}$	<i>molecules</i>	3500	Homeostatic set-point of the mtDNA population (Heart cells)	[230]
$\nu_R^{max} _{PN}$	<i>molecules</i> d^{-1}	371.82	Maximum replication rate of mtDNA during post natal stage, obtained using Equation 4.4 , considering $F_M = 0$ and $k_R = k_d \cdot (W + M)_{ss}$.	[245]
k_m	<i>rep</i> $^{-1}$	3.5×10^{-7}	<i>de novo</i> mutation rate of mtDNA	[215]
N_{cell}	—	2.2443×10^7	Number of cells (Heart)	[230 , 248]
β	—	2.25	Replicative gain factor of mutant mtDNA.	
K_H	—	0.15	Hill constant representing the mtDNA phenotypic threshold.	[205 , 245]
n	—	2	Hill coefficient.	[205 , 245]



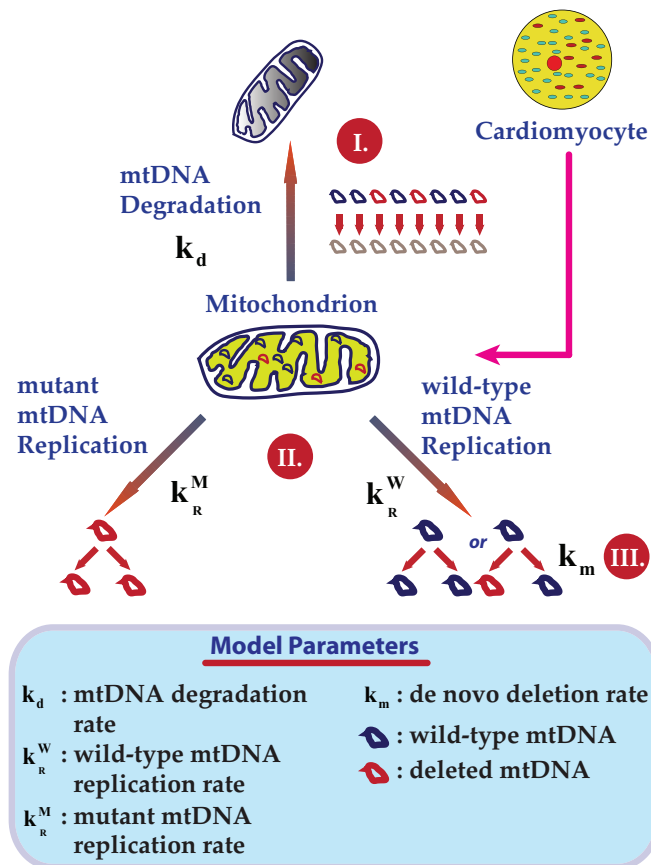
can be represented as:

$$\begin{aligned} \frac{\partial P(W, M; t)}{\partial t} = & k_d \cdot \sum_{x=0}^{10} \left\{ \frac{\binom{W+x}{x} \binom{M+(10-x)}{10-x}}{\binom{W+M+10}{10}} \right\} \cdot (W+x, M+(10-x)) \\ & \cdot P(W+x, M-(10-x); t) \\ & - k_d \cdot \sum_{x=0}^{10} \left\{ \frac{\binom{W}{x} \binom{M}{10-x}}{\binom{W+M}{10}} \right\} \cdot (W+M) \cdot P(W, M; t) \\ & + (1-k_m) \cdot k_R \cdot P(W-1, M; t) - (1-k_m) \cdot k_R \cdot P(W, M; t) \\ & + k_m \cdot k_R \cdot P(W, M-1; t) - k_m \cdot k_R \cdot P(W, M; t) \\ & + \beta \cdot k_R \cdot P(W, M-1; t) - \beta \cdot k_R \cdot P(W, M; t) \end{aligned} \quad (4.2)$$

As discussed in the earlier chapters, the probability density function $P(W, M; t)$ denotes the probability of a cell in a tissue to contain W and M copy numbers of wild-type and mutated mtDNA, respectively, given an initial condition of the mtDNA population in the cell (not explicitly stated here for brevity). The parameters k_d , k_m and k_R are the specific probability rate constants of mtDNA degradation, *de novo* mutations and replication rates, respectively (Table 4.1). The terms in the curly braces of Equation 4.2 represent the hypergeometric sampling of mtDNA from the mitochondrial population. The first two terms in Equation 4.2 represents degradation of mtDNA in a single mitochondrion. The third line in the Equation 4.2 represents the replication of wild-type mtDNA template in

the absence of *de novo* mutations. Forth term in Equation 4.2 represents the *de novo* mutations generated during the replication of a wild-type mtDNA molecule. The last term in Equation 4.2 corresponds to replication of mutant mtDNA. The solution of CME was obtained numerically using a Monte Carlo approach following modified Stochastic Simulation Algorithm (SSA) developed in the earlier chapter on the stochastic dynamics of mtDNA point mutations in mouse heart tissue (Chapter 3; [269]).

Figure 4.1: **Stochastic model of mtDNA turnover dynamics in a mouse cardiomyocytes.** Stochastic drift in mtDNA mutation dynamics results from following random processes. (I) The mitochondrion that undergoes a turnover event is randomly selected from the population. The autophagy of mitochondrion is simulated by removing all the mtDNA molecules associated with the mitochondrion. (II) Replication of a single mtDNA molecule occurs by random selection of mtDNA from the mitochondrial DNA population. (III) During the wild-type mtDNA replication, there exists a finite probability equal to the *de novo* mutation rate (k_m) for the replication process to give a mutant mtDNA.



4.2.2 Tissue-level modeling details

The number of mtDNA copy number in the embryonic cells during the mouse embryogenesis is roughly 1000 mtDNA copies [291]. During the 21 days of mouse gestation, binary cell division of mouse embryonic cell occurs at roughly every 15.5 hours [236, 243]. The developmental simulation not only captured the rapid increase in cell number and the associated increase in the mtDNA copy number in the developmental embryonic cells during this time, but also accounted for the normal turnover of mtDNA and cell death and

replacement of respiratory deficient cells generated during the gestation period (mutant burden; $(M/(W + M) = F_M) > 90\%$). After birth, the mtDNA population in cardiomyocytes is simulated to undergo hypertrophic growth. The mtDNA population in the mouse cardiomyocytes increases from 1000 molecules to about 3500 copies per cell in a span of 6 months [230, 292, 293]. After reaching the nominal count of mtDNA in adult cardiomyocytes [230], the mtDNA copy number of cardiomyocytes was held at constant level (Table 4.1), by relaxed replication [107, 183, 232]. The other implementation details related to the mtDNA deletion model, for example simulation of developmental binary cell divisions and mtDNA relaxed replication in the somatic non-dividing cardiomyocytes, is the same as that in the point mutation model in Chapter 3 (see also [269]).

Simulations were performed using an IBM high performance computing cluster with 112 Intel 1.6 GHz processors. The Monte Carlo algorithms were implemented in C++ on a Linux Platform (CentOS; GNU C++ compiler (v4.1.1)). On average a complete simulation of a heart tissue (~25 million cells) from the development to the end of 3 years of mouse's life span required approximately 7 hours.

4.2.3 Model parameters

In the present work, the mtDNA point mutation model was modified to include the mitochondrial biogenesis in the presence of retrograde signaling, to model the functional effects of *de novo* mutations. The abundance of functionally deficient mutant mtDNA is a crucial factor in determining the existence of a biochemical defect [122]. Due to the abundance of wild-type mtDNA in a cell and due to efficient complementation arising from frequent mitochondrial fusion-fission [140], mtDNA mutations are functionally recessive and usually $> 60\%$ of mtDNA molecules must contain mtDNA deletions before respiratory defect is manifested [140, 294]. Cybrid cell studies with varying fractions of pathogenic deletions (Δ mtDNA4696), have indicated that the functionality of cybrids maintained at almost

100% until the abundance of deletions reached to about 60% [140]. Beyond this level, there was a sharp decrease of the COX activity with increasing proportions of deletions (resembling a switch behaviour). Therefore in this modeling framework, a Hill-type cooperative kinetics [245] was adopted to simulate retrograde signaling process, insofar as the mitogenesis increased in response to a drop in the proportion of wild-type mtDNA. The parameters of the Hill-type kinetics, K_H and n_H can be used as set point levels of mtDNA copy number and the sharpness of the switch response, respectively [245]. Consistent with the experimental observations [205], a Hill constant of $K_H = 85\%$ and a Hill coefficient of $n_H = 2$ was used in this model.

It's a well known fact that the mitochondrial metabolites are efficiently complemented in the mitochondrial network, due to presence of mitochondrial fusion-fission process. The metabolites synthesized by both the wild-type and mutant mtDNA are equally sequestered into the mitochondrial network due to the efficient complementation resulting from mitochondrial fusion-fission process [140]. The data shown in this work hints of no preferential protein expression capacity for or against the mutant mtDNA molecules. COX deficient cells having a large fraction of mutant mtDNA population [26] will also produce significant fraction of dysfunctional respiratory proteins in the cells. With the abundance of the respiratory dysfunctional proteins in the mitochondrial mass, protein content available for the OXPHOS machinery will be mostly dysfunctional. Thus, it is most likely possible that the deficiency status in the cell is better characterized by the abundance of mutant mtDNA in the mtDNA population, which is also explicitly shown in one of the experimental data [26, 140]. Thus, the respiratory deficiency of the cells in the model developed in this work is tracked using the abundance of mtDNA mutation or the mtDNA mutation burden in the cell (F_M).

In simulations, cardiomyocytes were considered COX deficient when the mutation burden in the cell is $0.6 < F_M \leq 0.9$ and cardiomyocytes were considered completely

respiratory deficient when the mutation burden is $F_M > 0.9$. In other words, the cells with mtDNA mutation burden $> 60\%$ were characterized into two classes: COX deficient cells for those having $0.6 \leq F_M < 0.9$ and completely respiratory deficient cells for those having $F_M \geq 0.9$. The latter cells are treated as dead cells in the simulations.

The propensity of replication process comprises of: a.) maximum replication rate ν_{max} and, b.) negative Hill term capturing the retrograde response of the nucleus.

$$k_R = \nu_{max} \cdot \left(1 - \frac{F_W^{n_H}}{K_H^{n_H} + F_W^{n_H}} \right) \quad (4.3)$$

where ν_{max} is computed based on the basal level of mtDNA turnover rate in cells, considering no mtDNA mutations ($M = 0$; Table 4.1), and F_W is the ratio $W / (W + M)$.

Different experimental evidence hints for the presence of replicative advantage of different types of pathogenic mutant mtDNA involved in the clonal expansion process. In the case of mtDNA deletions, the smaller size of the mtDNA deletions have been suggested to result in a replicative advantage of deleted mtDNA [25, 295, 296]. Repopulation dynamics of deletions in heteroplasmic population of mtDNA, showed that the mtDNA with large deletions had the ability to repopulate the organelles much faster than the smaller length deletions or full length wild-type genome, within the same cell, under the relaxed replication regime [296]. However, in the case of mtDNA point mutations, the modification of DNA nucleoside, resulting from the point substitution has been suggested to alter specific control sequences (e.g. destruction of cis-inhibitory sequences) and rendering the mutated mtDNA template with replicative advantage [211]. Similarly, mtDNA with partial duplication have also been found to accumulate with age. In this case, it has been proposed that partial duplications have additional replication origin, which might enhance the replicative propensity associated with the partially duplicated mtDNA [297]. Considering these data, a replicative gain factor (β) is used to calculate the replication propensity of the mu-

tant mtDNA as:

$$k_R = \beta \cdot \nu_{max} \cdot \left(1 - \frac{F_W^{n_H}}{K_H^{n_H} + F_W^{n_H}} \right) \quad (4.4)$$

In stochastic modeling of mitochondrial DNA turnover in rat cardiomyocytes there are two uncertain parameters (i) the experimental data on the estimate of *de novo* mtDNA deletion rate (k_m) spans an order magnitude [215], and (ii) the value of replicative gain factor (β), for which the experimental data is unavailable. As discussed in previous paragraph, the replicative advantage of the mutations such as partial deletions/duplication or point mutation are attributed to some characteristics that arise from the modification of the mtDNA templates resulting from these mutations. Also, considering that mtDNA are generally localized in the form of nucleoids, and depending on the availability of localized replication factors, the quanta of mtDNA mutant replication arising due to short length of the mtDNA deletion, would be much higher than that of wild-type in some cells, and this would eventually lead to clonal expansion process. Since the present model only considers a conservative assumption of well mixed mtDNA population, an average replicative gain factor arising from such replication "burst" has been assumed in this work. The mechanism of clonal expansion of one or few mutant mtDNA in the cells is still at best hypothesized. The simulations presented in this chapter and in the next chapter strongly indicates the necessity for the presence of replicative advantage of mutants in explaining the mutation accumulation data in mouse heart tissue. Due to this uncertainty, we have instead rely on the data on mutation burden to infer the advantage factor using experimentally determined *de novo* mutation rate between 5×10^{-8} to 5×10^{-7} per replication event [215].

In a recent work [296], Fukui and Moraes quantified the relative abundance of mtDNA deletions in mice tissue at different ages. Figure 5 in their work describes the abundance of different mtDNA deletions [296]. The mtDNA deletion content was normalized with respect to the total mtDNA content (F_M). In order to understand at what values of β such

fold difference appear, ODE simulations were performed assuming no *de novo* mutations.

The following equations were used:

$$\begin{aligned}\frac{dW}{dt} &= k_R \cdot F_W - k_d \cdot W \\ \frac{dM}{dt} &= \beta \cdot k_R \cdot F_M - k_d \cdot M\end{aligned}\quad (4.5)$$

where, W and M stands for wild-type and mutant mtDNA, respectively. k_R , k_d represents the mtDNA replication and degradation rates, respectively, $F_W = W/(W + M)$ and $F_M = M/(W + M)$. Using this model, the normalized values of mtDNA mutations was obtained for different values of β and for different initial burden of mutant mtDNA. As shown in [Figure 4.2](#), boost factor values of greater than or equal to 2 is required for explaining the experimental data. Given the uncertainty in the mechanism of replicative advantage of mutant mtDNA and since the determination of this parameter is significantly challenging, different replicative gain factor ranging from ($\beta \approx 2 - 5$) was considered for the simulations.

The influence of different selections of β and k_m on the model predictions are illustrated in [Figure 4.3](#). Based on the number of COX deficient cells observed in experimental data [28], optimization indicates values of: $\beta = 2.25$ and $k_m = 3.5 \times 10^{-7} \text{ rep}^{-1}$ ([Table 4.1](#)). For all the simulations performed in [Chapter 4](#) and [Chapter 5](#), a replicative gain factor of $\beta = 2.25$ is chosen. For the analysis considering the influence of different replicative gain factor on the mtDNA mutation dynamics, β range from 2-5 is chosen based on the results obtained in [Figure 4.2](#). All the other model parameters are adapted from the [Chapter 3](#) (also see [269]) and are listed in [Table 4.1](#).

4.2.4 Modeling details of different hypotheses

Several hypotheses have been proposed over past decades in explaining the mechanism for clonal expansion process. Commonly proposed hypotheses are broadly classified into

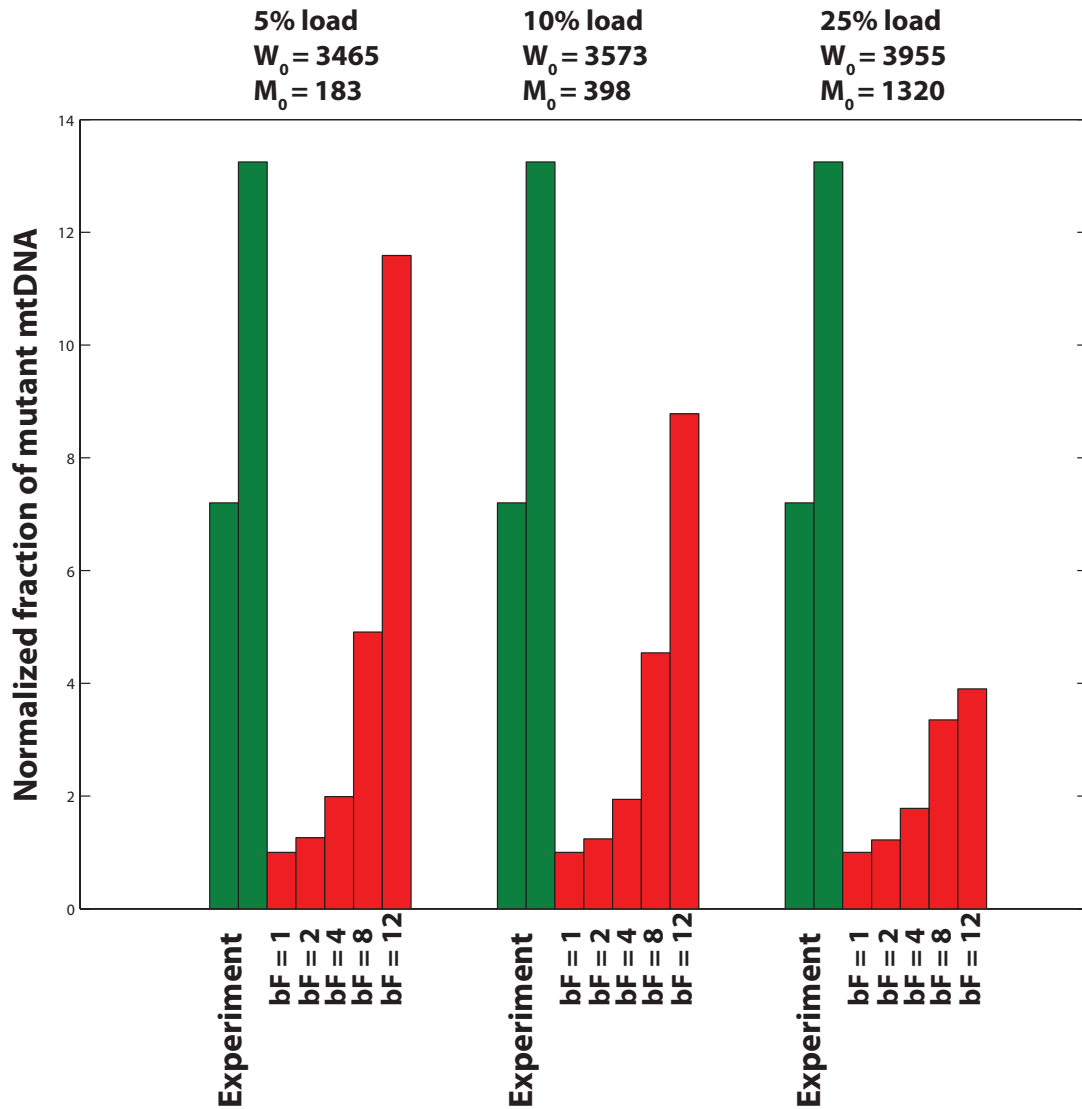


Figure 4.2: Simulation results of the fold increase at different replicative advantage factors and at different initial load. Experimental data were obtained from [296].

following categories:

Hypothesis A: Structural hypothesis

Earlier hypotheses based on purely structural accounts [25] proposed that clonal expansion of mtDNA mutations result from the replicative advantage of any mtDNA template alone. However, this did not explain the mitochondrial hyperproliferation seen in COX deficient regions of muscle fiber [26, 140]. The primary effect of functional mutations is to cause a nuclear metabolic feedback, which initiates mtDNA proliferation. A great deal of evidence hint for the existence of such mitochondrial proliferation in response to oxidative

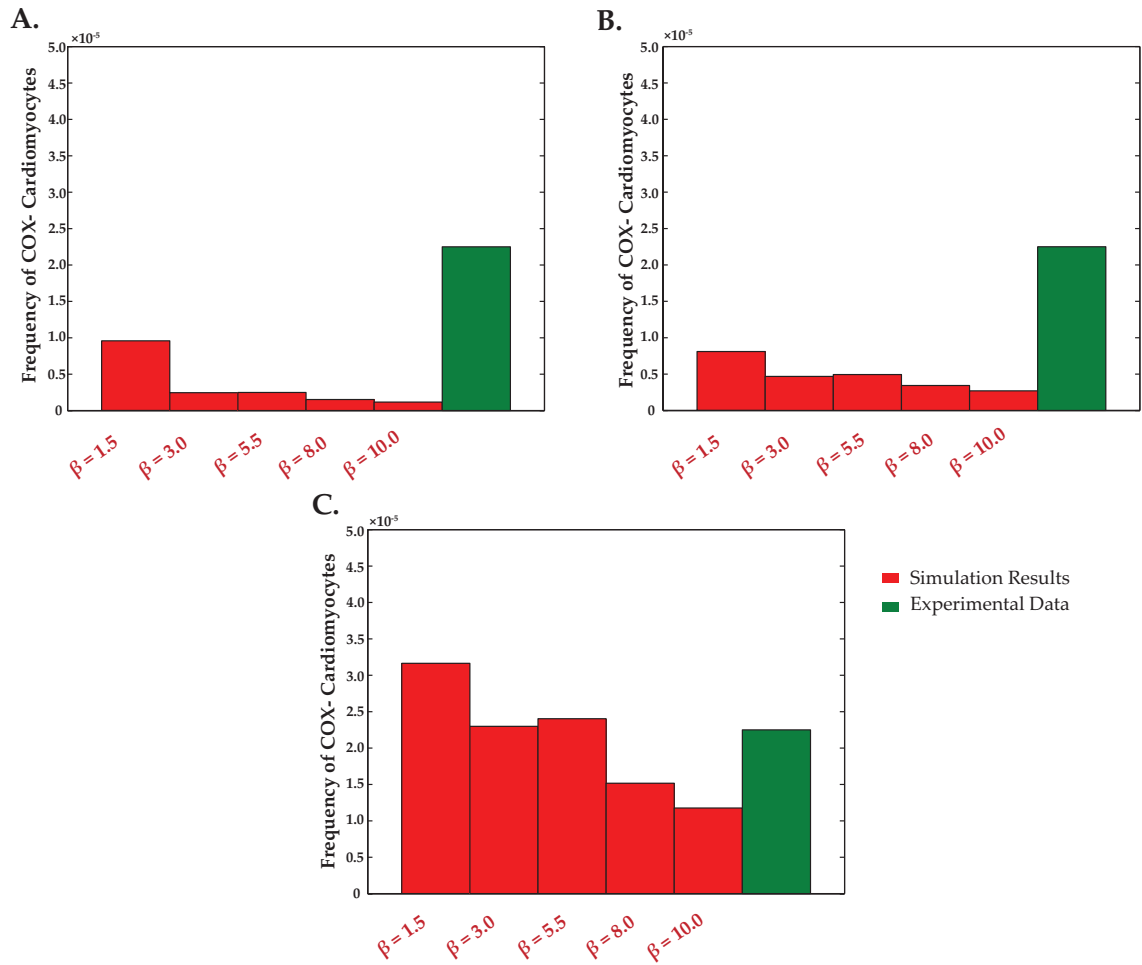


Figure 4.3: **Parametric perturbation analysis on mouse heart ($n = 3$)**. Plots indicated a comparison of simulation results at three different values of k_m for different values of mtDNA replicative gain factors (β), with the experimental data [28]. The simulation trials were performed to obtain the optimum value of these parameters. A.) simulation results with $k_m = 5 \times 10^{-8} \text{ rep}^{-1}$ with varying β , B.) simulation results with $k_m = 1 \times 10^{-7} \text{ rep}^{-1}$ with varying β , and C.) simulation results with $k_m = 5 \times 10^{-7} \text{ rep}^{-1}$ with varying β . Based on the optimization and consistent with the experimental data on the conservative estimates of the mtDNA replicative gain factor, the following parameter values were chosen for the simulation: $\beta = 2.25$ and $k_m = 3.5 \times 10^{-7} \text{ rep}^{-1}$.

demand [140, 298]. In the structural hypothesis, the mutant mtDNA molecules are proposed to have certain replicative advantage due to changes in the mitochondrial genetic structure. As discussed above, different structural modification to the mitochondrial DNA such as, shorter replication length of deleted mtDNA [25, 295, 296] and structural changes caused due to mtDNA point mutations [211], have been proposed to give replicative advantage to mutated mtDNA. This hypothesis however makes no reference to functional consequence of proteins that are generated by mutated mtDNA (e.g. influence of altered proteins on the mitochondrial respiratory process). In modeling the structural hypothesis,

only the replication propensity of the wild-type and mutant mitochondrial genome were modified as:

$$k_R = \nu_{max} \quad (4.6)$$

$$k_R = \beta \cdot \nu_{max}, \quad (4.7)$$

respectively (i.e. no nuclear retrograde signaling). Consistent with our point mutation model [269], the value of ν_{max} was computed based on the basal level of mtDNA turnover rate in cells (Table 4.1). The other modeling details and parameters related to the stochastic turnover process of mtDNA and the tissue level modeling details, including the developmental binary cell lineage, used for this hypothesis is consistent with the modeling details used for the present work.

Hypothesis B: Random drift hypothesis; "Basic" model

As detailed in Chapter 2, this hypothesis proposes that random drift in mitochondrial genetics, resulting from either mitotic segregation in the proliferative tissue [219, 299], or due to constant mitochondrial turnover in non proliferating tissues [210], could eventually result in its clonal expansion. In the Basic model, Chinnery and Samuels groups proposed random intra-cellular drift resulting from relaxed replication of mtDNA alone could cause clonal expansion of mtDNA mutations in human post-mitotic cells [209, 210]. The model was based on neutral alleles, causing no functional deficiency to cellular respiration. The original works by Chinnery and Samuel's group did not specifically model for the origin and accumulation dynamics of mtDNA mutations during the animal development [209, 210]. Consistent with these works, relaxed replication of mtDNA population undergoing nominal population turnover in a post-mitotic cells was simulated [210], where the replication propensity of mtDNA (both W and M) is described by Equation 4.6 and the value of mitochondrial degradation, k_d was based on the turnover rate of 10 days [218]. The value of ν_{max} was computed based on the basal level of mtDNA turnover rate in cells

and is computed to be 69.31 mtDNA molecules per day. In accordance with the "Basic model" of relaxed mtDNA replication [210], the *de novo* mutation rate of $k_m = 1 \times 10^{-5} \text{ rep}^{-1}$ was used for simulations.

Hypothesis C: Random drift hypothesis; "Disease" model

In the "Disease model", Chinnery and Samuels proposed [209] that in the presence of pathogenic mutant mitochondrial allele, which causes functional deficiency in the OXPHOS process [209] and subsequent nuclear retrograde response results in clonal expansion of mutant mtDNA. For simulating this model, same parameter values used for the Basic model were used. However for this hypothesis, replication propensity of the mtDNA population was modeled as a piece-wise linear function [209]:

$$k_R = \nu_{max} \cdot \left[\alpha + (1 - \alpha) \cdot \left(\frac{W}{W_{opt}} \right) \right] \quad (4.8)$$

where, $\alpha = 5$ is a growth factor for the clonal expansion process [209] and W_{opt} is the optimal level wild type mtDNA templates in a cell; Table 4.1.

Hypothesis E: Functional hypothesis

Another commonly hypothesized mechanism for explaining the mutant mtDNA clonal expansion process is the "Survival of the slowest" (SOS) [202]. Several recent experimental data have challenged the appropriateness of this hypothesis, as it indicates that the selective mitophagy by the autophagosomes, specifically targets the dysfunctional mitochondria rather than the functional members [251, 255, 300]. It is also proposed that the mitochondrial membrane potential which is modulated by the OXPHOS status of mitochondria is a key factor in controlling mitophagy [251]. In light of these evidences, the feasibility of SOS hypothesis to explain the mtDNA mutation clonal expansion mechanism was not tested in this work.

Table 4.2, summarizes different assumptions that have been used to explain the clonal expansion dynamics of mutant mtDNA.

Table 4.2: Details of different hypotheses proposed to explain the clonal expansion of mtDNA mutations.

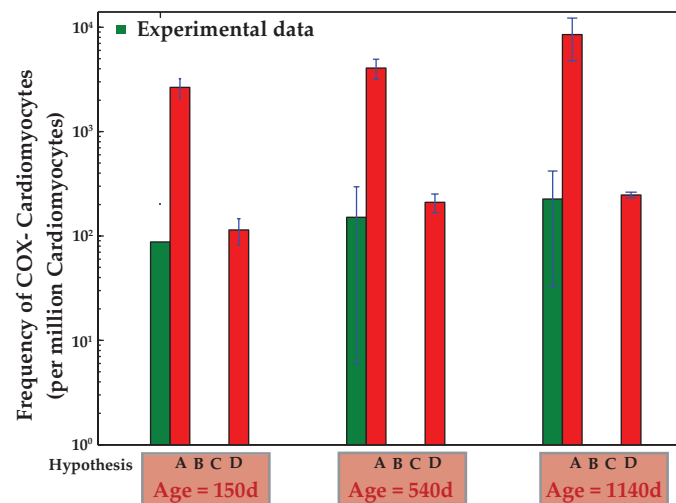
Hypothesis	Associated mitochondrial genome processes			
	Replicative advantage of genome	Nuclear retrograde response	Random intra-cellular drift of mtDNA template	Degradation advantage of mutant mtDNA
A Structural hypothesis (Survival of Tiny)	yes	no	no	no
B Random drift of mtDNA mutations; "Basic" model	no	no	yes	no
C Random drift of mtDNA mutations; "Disease" model	no	yes	yes	no
D Present Model	yes	yes	yes	no
E Functional Hypothesis (Survival of Slowest)	no	yes	no	yes

4.3 Results & Discussion

4.3.1 Clonal expansion mechanism

Figure 4.4, illustrates the mtDNA mutation burden in rat cardiomyocytes modeled using different clonal expansion hypotheses presented in Table 4.2. Simulation results of the

Figure 4.4: Comparison of the results from different hypothesis testing on the mechanisms of mtDNA mutant clonal expansion process. Comparison of temporal dynamics of COX deficient cells accumulating in heart tissue with age, using different clonal expansion hypotheses (Table 4.2). Clearly, both replicative advantage of mtDNA mutations with mitochondrial hyper-proliferation is required to match the experimental data [28].



frequency of COX deficient cardiomyocytes in mouse heart tissue indicated that both the nuclear feedback and replicative advantage of mtDNA mutations was required to explain

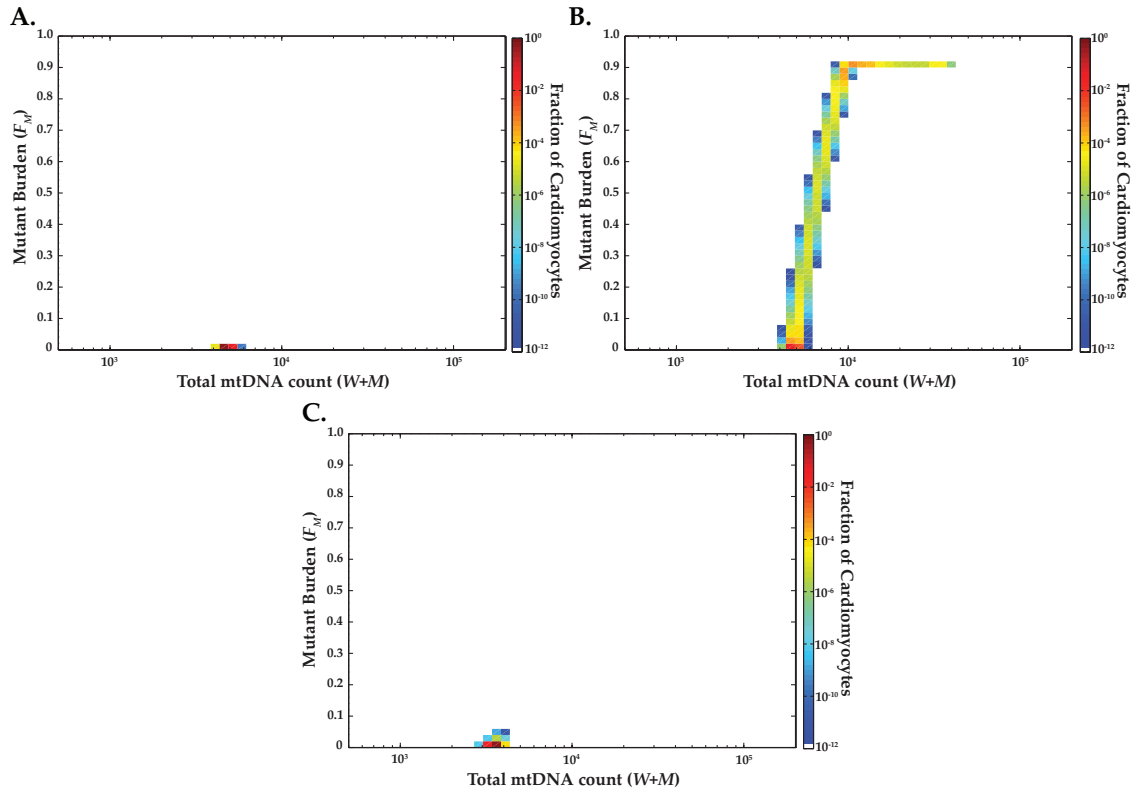


Figure 4.5: **Parametric perturbation analysis on mouse heart ($n = 3$), considering model used for hypothesis D.** A.) State space probability density function of Total mtDNA count ($W + M$) and mutation burden (F_M) in mouse cardiomyocytes of the complete heart tissue in the absence of both nuclear retrograde signaling for mtDNA biogenesis and replicative advantage of mutated mtDNA (parametric sensitivity analysis of the model used in the present work). Evidently, the results indicated that the mtDNA copy number did not deviate significantly from the nominal mtDNA count in the cells. Similarly the mutation burden in all the cardiomyocytes remained at the minimal level and all the cardiomyocytes were COX sufficient at the end of life. B.) State space probability density function of Total mtDNA count ($W + M$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue in the absence of only the nuclear retrograde signaling for mtDNA biogenesis (i.e. in the presence of mutant mtDNA replicative advantage; parametric sensitivity analysis of the model used in the present work). There was a large fluctuation of copy number distribution and the COX deficient cells were found to have total mtDNA copy number of $\sim 10,000$, which on an average was an order of magnitude lower than the experimental data [28] and with the simulation results obtained using the present model. C.) State space probability density function of Total mtDNA count ($W + M$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue in the absence of only the mitochondrial replicative advantage (i.e. in the presence of nucleo-mitochondrial retrograde feedback; parametric sensitivity analysis of the model used in the present work).

the mouse cardiomyocytes data (Figure 4.4 and 4.5).

All the colored state space plots presented in this chapter, like the ones in Figure 4.5, represents the distribution of total mtDNA copy numbers (both W and M) and the mtDNA mutation fraction (F_M) in all the cells that have at least a single copy of mutations ($M > 0$). Figure 4.6A indicates that at the end of development, the majority of cells have low fractions of mtDNA mutations. Due to mitochondrial hyperproliferation and

replicative advantage, a minor fraction of cells drifted to higher copy numbers (both W and M) and some of them became completely respiratory deficient. Mostly, the mutated cells in development formed seeds for clonal expansion of COX deficient cells. By neglecting developmental stage, **Figure 4.6C** shows that all cells remain COX positive (mtDNA mutation load $F_M \leq 60\%$) at the end of mouse life span. Consistent with the modeling results obtained in the earlier work on point mutation dynamics in mouse cardiomyocytes (**Chapter 3**), **Figure 4.6C** clearly highlights the importance of developmental contribution and the important role of *de novo* mutagenesis during development, in mtDNA mutation clonal expansion process.

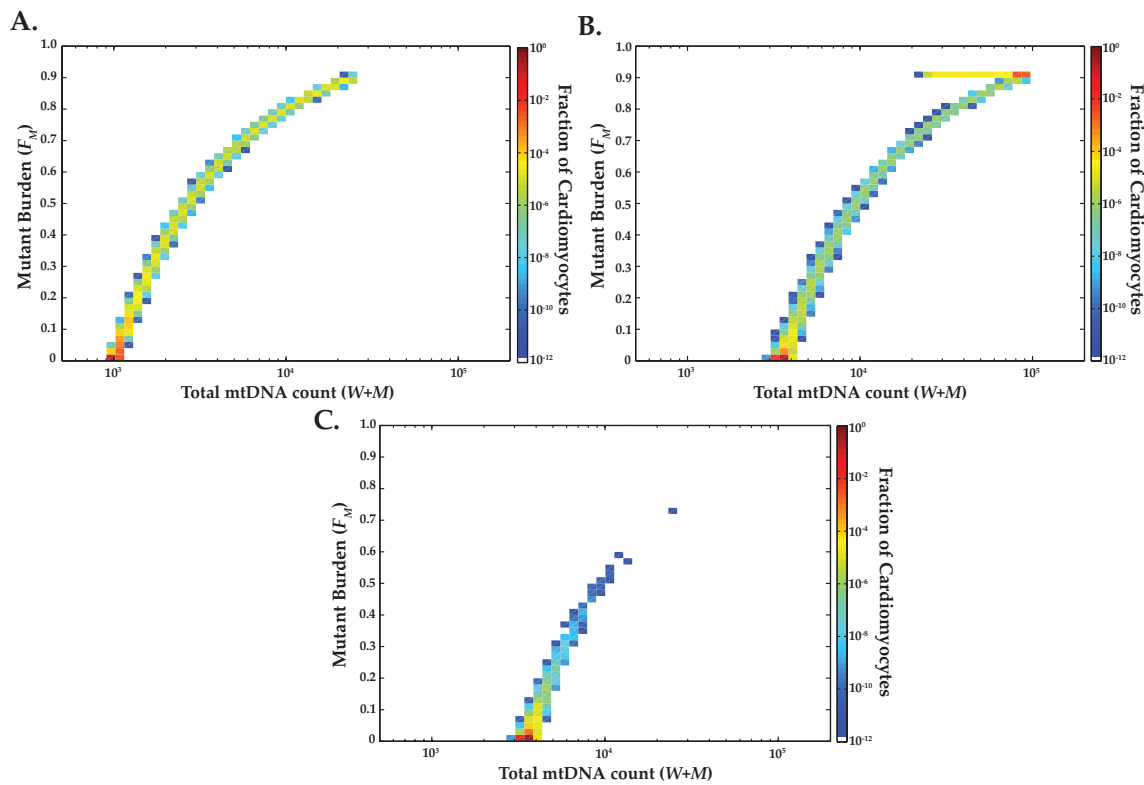


Figure 4.6: Mitochondrial DNA deletion accumulation in mouse heart tissue ($n = 3$), Influence of animal development in the mtDNA mutation heterogeneity observed in the mouse heart tissue. Simulations were performed using the model and parameters of hypothesis D. A.) State space probability density function of Total mtDNA count ($(W + M)$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue at birth. B.) State space probability density function of Total mtDNA count ($(W + M)$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue at the end of 38 months. C.) State space probability density function of Total mtDNA count ($(W + M)$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue at the end of 38 months, with the exclusion of development.

Figure 4.6B indicates that at the end of mouse's life span, there are two major factions

of cardiomyocytes, one containing the cells that are COX positive and the other being completely respiratory deficient. Although the transition between the two groups appears to be instantaneous (low frequency of cells having intermediate mutation burden), the entire process of clonal expansion, from the initial *de novo* mutation to becoming respiratory deficient, takes a long time (starts early in the developmental stage; see also [270]). Consistent with experimental data [217, 301], at birth approximately 0.3% cells harbored mutations. However at the end of 38 months, approximately 0.5% of cardiomyocytes have mtDNA mutations (i.e. $F_M > 0$).

Figure 4.7 and 4.8 compares the simulation results of the spectrum of mtDNA mutation burden in mouse cardiomyocytes, considering different clonal expansion hypotheses (Table 4.2). Clearly, the pure intra-cellular genetic drift model, without the replicative advantage of mtDNA mutations was insufficient in explaining the mtDNA deletion dynamics data observed in mouse heart cells (Figure 4.5, 4.7 and 4.8). Both the random intra-cellular drift models ("Basic" and "Disease models") produced no COX deficient cardiomyocytes during the mouse life span (Figure 4.5 and 4.7). Also, it can be seen that the Random intra-cellular drift model produced relatively narrow spectrum of mtDNA states (Figure 4.8A and Figure 4.8C), unlike what is generally observed with *in vivo* studies related to mouse and human tissues [26, 296]. Most of the cells in these simulations have a very low mutation load (F_M is usually $< 10\%$), and a very negligible fraction of cells have mutation fraction $F_M > 10\%$. In the case of Disease model of random intra-cellular mtDNA drift, the spread of the mtDNA states was very similar to that of the "Basic model" (Figure 4.7 and 4.8C).

Another interesting aspect related to the Random intra-cellular genetic drift model (hypothesis B) was the spread of distribution of the mtDNA states. Comparing Figure 4.8A and 4.8B, indicates that for tissue having lower population or count of mtDNA (tissue with lower oxidative capacity; Figure 4.8B), the fraction of cells having higher mutation burden

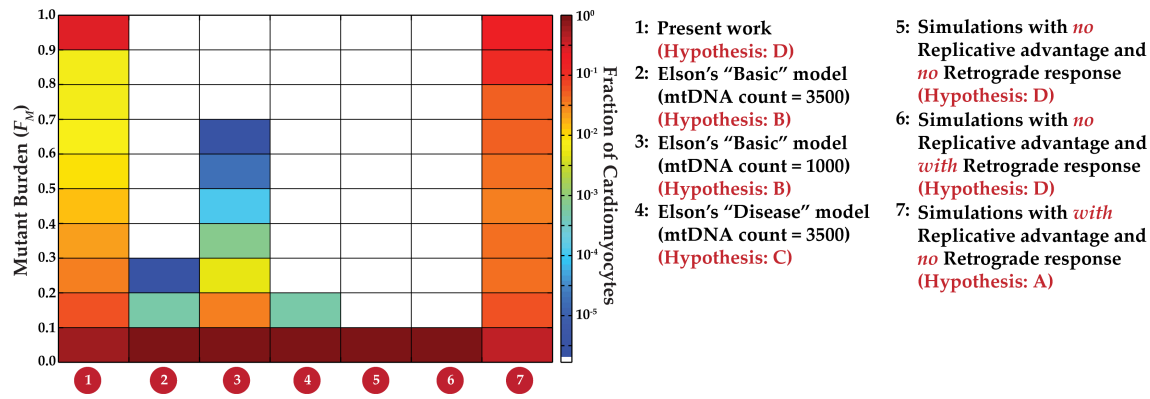


Figure 4.7: Comparison of simulation outcome of using different clonal expansion hypothesis for capturing the dynamics of mtDNA mutation accumulation in mouse heart ($n=3$). Color plots illustrating the results of modeling hypothesis D (1, 5 and 6 in the above plot), represents different parametric sensitivity analyses related to the model used in this work.

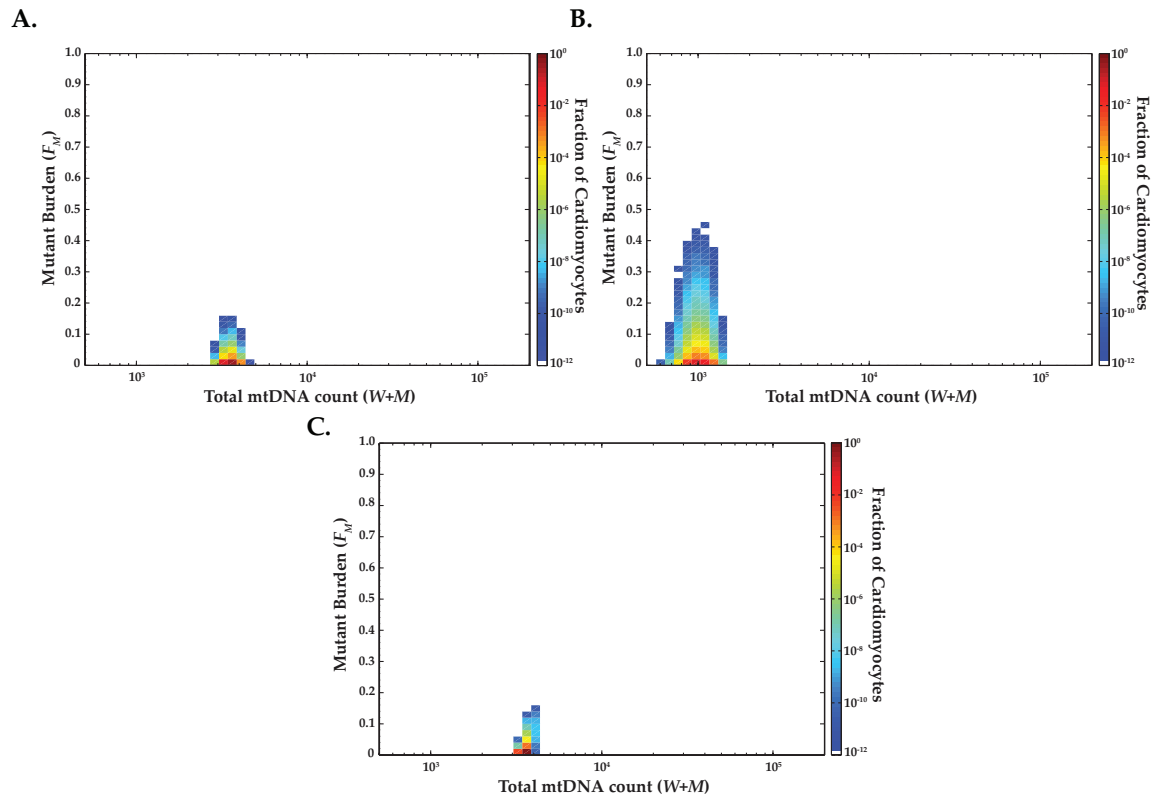


Figure 4.8: Comparison of simulation outcome on the state space distribution of mtDNA, obtained using different clonal expansion hypothesis for capturing the dynamics of mtDNA mutation accumulation in mouse heart ($n=3$). A.) State space probability density function of total mtDNA count ($W + M$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue at the end of 38 months, obtained using Random intra-cellular drift "Basic" model (hypothesis B) including the animal development. B.) State space probability density function of Total mtDNA count ($W + M$) and mutation burden F_M in mouse cardiomyocytes containing mtDNA count of 1000, at the end of 38 months, obtained using Random intra-cellular drift "Basic" model (hypothesis B) including the animal development. C.) State space probability density function of Total mtDNA count ($W + M$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue at the end of 38 months, obtained using Random intra-cellular drift "Disease" model (hypothesis C).

is higher compared to tissue having larger mtDNA copy number (tissue with higher oxidative capacity; [Figure 4.8B](#)). This demonstrated a high sensitivity of random intra-cellular drift model to the initial population of mtDNA in cells. However, it is generally seen that metabolically active tissues with higher oxidative capacity (mitochondria-rich) have larger mutation burdens [139], which is consistent with the predictions obtained from the present stochastic model (hypothesis D; [Figure 4.4](#)). Furthermore, as observed in [Figure 4.8C](#), the results of mtDNA mutation dynamics with only the presence of nuclear retrograde response is similar to the "Disease model" ([Figure 4.5C](#)).

As predicted in the earlier chapter ([Chapter 3](#); [269]), the most sensitive parameter in modeling stochastic mtDNA mutation dynamics is the turnover rate ([Figure 3.8](#) and [4.9](#)). Chinnery and Samuels groups used a turnover rate of ~ 10 days [209] for modeling mtDNA turnover dynamics (hypothesis B and C). However, studies on mtDNA replication in post-mitotic cells have indicated that the replication frequency in post mitotic tissues like brain [302], skeletal muscle [240] and heart [240, 274] may be fairly low. A low frequency of mtDNA turnover with random intra-cellular drift model gave at most no COX deficient cardiomyocytes. However, even in the case of higher mtDNA turnover rate, simulation results did not produce enough clonal expansion of mtDNA mutations in mouse cardiomyocytes ([Figure 4.4](#), [4.7](#) and [4.8](#)). While random intra-cellular drift can give clonal expansion in some cells, it fails to explain how a substantial fraction of cells ($\sim 50\%$) in substantia nigra region of human brain have clonally expanded mtDNA deletions [27, 154]. Therefore a critical challenge to accurately identify the frequency of mtDNA turnover in the post mitotic cells is warranted for better understanding of mtDNA mutation clonal expansion mechanism.

4.3.2 Influence of mtDNA replicative advantage

in vivo analyses have indicated a marked replicative advantage of deleted mtDNA [296].

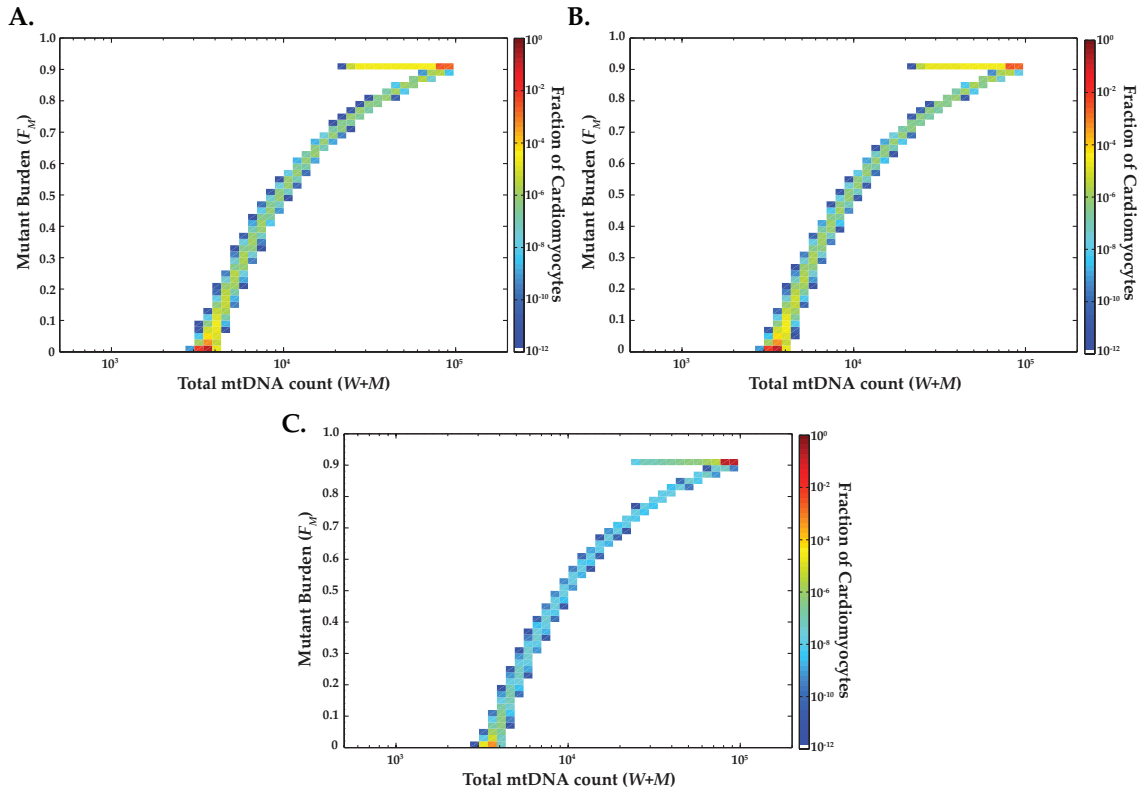


Figure 4.9: Parametric perturbation analysis on mouse heart ($n = 3$). State space probability density function of Total mtDNA count ($W + M$) and mutation burden (F_M) in mouse cardiomyocytes of the complete heart tissue simulations performed using the model and parameters of hypothesis D. A.) results of present modeling work, with parameters presented in Table 4.1, B.) with *de novo* mutation rate $10\times$ the normal rate. All the other parameters were kept at the same level as used for the present work. C.) with mtDNA turnover rate of 10 days [218]. All the other parameters were kept at the same level as used for the present work. Results representing the effect of increased *de novo* mutation rate indicated a low sensitivity of changing this rate on copy number distribution of mtDNA. Due to increased quanta of replication, the mtDNA turnover has a significant effect on mtDNA copy number distribution. With increasing turnover many of the cells with mtDNA deletion at birth becomes COX deficient at the end of life span.

As discussed earlier, this phenomenon has an important implication on the mechanisms of mtDNA deletions accumulation in aging and disease processes. For instance, deletions formed in transgenic mouse neurons due to a double strand break formed by the action of *PstI*, generated a wide range of mtDNA deletions in the neuron cells [296]. Using this mouse model, age-dependent accumulation of different deletion lengths was monitored. Mitochondrial DNA sequence analysis of deletion products identified that majority of clonally expanded deletions spanned $> 5\text{kb}$, indicating that larger deletions have greater potential to clonally expand [296]. This observation is also consistent with other analysis considering variety of normal aging tissues from mice, rats, rhesus monkeys and

human [152]. To analyze the influence of replicative advantage of mutated mtDNA on the clonal expansion dynamics, simulation of mtDNA mutation accumulation with different replicative gain factor (β) were performed, assuming the basal frequency of different *de novo* mtDNA mutation to be the same (considering the model developed in this work; hypothesis D).

Two different replicative gain factors (β) were simulated in mouse cardiac tissues, starting with development until the end of mouse's life span. State space distribution of mtDNA in rat cardiomyocytes harboring mtDNA mutations of replicative gain factor of $\beta = 1.5$, indicates a very minor fraction of cells that are COX deficient (Figure 4.10A). However, at lower level of replicative gain factor, the simulation results indicated no cells becoming completely respiratory deficient ($F_M > 0.9$). However, in the case of mutations having a larger replicative gain factor ($\beta = 4.5$), the clonal expansion dynamics was much swifter (Figure 4.10B). Comparatively, a large fraction of cells became completely COX deficient at such high replicative gain factors (comparing Figure 4.6B and Figure 4.10B). In accordance with the experimental data [296], the above analysis also indicated a phenomenal propensity of mutations with higher replicative gain factor to clonally expand.

Another interesting aspect is the amount of COX deficient cells formed during post-natal life, considering deletions with different β . The clonal expansion simulations without developmental seeds (i.e., both the *de novo* mutation and the subsequent clonal expansion occurs in the post-natal stage) was a rare phenomenon in post-mitotic cells with mutations having lower replicative gain factor (two cell were found to clonally expand in heart tissue having mutations with $\beta = 2.25$). However, a higher fraction of cardiomyocytes having no developmental contribution were estimated to become completely COX deficient ($F_M > 0.9$) with mutations having higher values of $\beta = 4.5$. For instance, $0.104 \pm 0.0012\%$ of cardiomyocytes having mutations with $\beta = 4.5$ were estimated to become clonally ex-

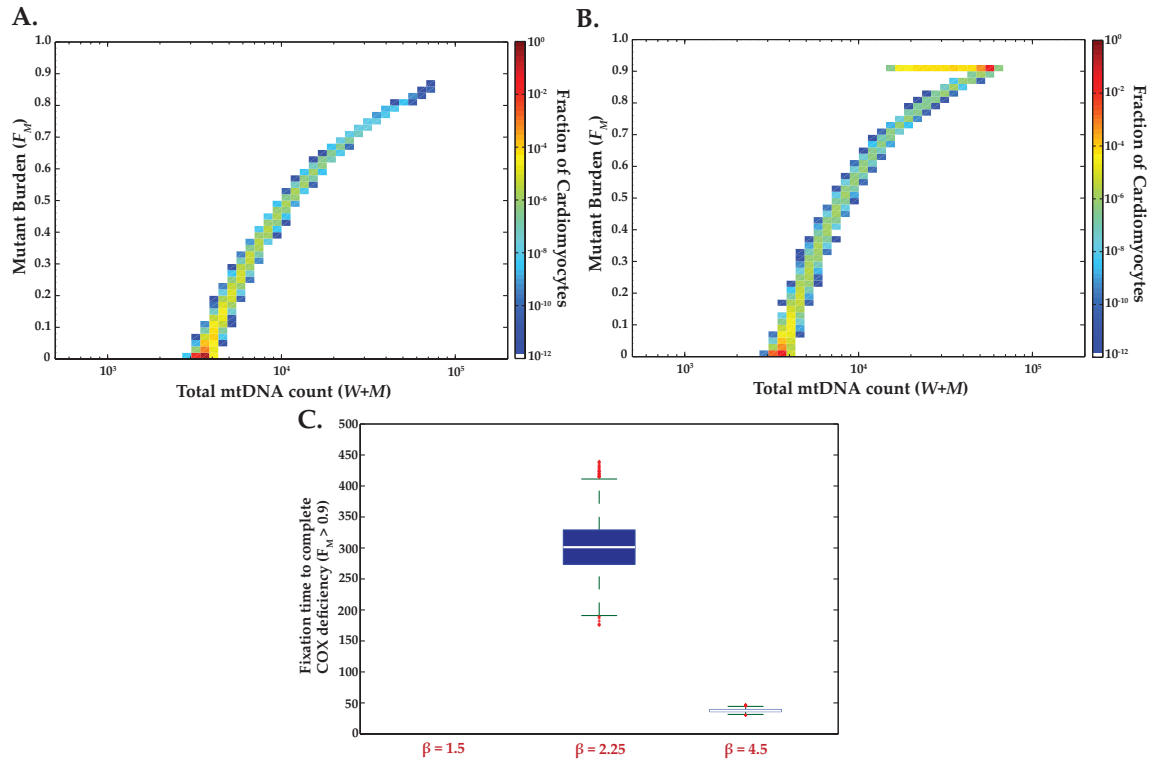


Figure 4.10: Influence of mutant mtDNA replicative advantage on mtDNA mutation clonal expansion dynamics. Simulations performed using the model and parameters of hypothesis D. State space probability density function of Total mtDNA count ($W + M$) and mutation burden (F_M) in mouse cardiomyocytes of the complete heart tissue at the end of 38 months, obtained using different replicative gain factors A.) $\beta = 1.5$ and B.) $\beta = 4.5$, respectively. C.) Distribution of time required for mtDNA template fixation. Temporal distributions were obtained by tracking time taken by post-natal cardiomyocytes to become completely respiratory deficient ($F_M > 90\%$). In each of these simulations, the time taken by heart cells containing mtDNA mutation load between 20 to 25% at the end of development, to become completely respiratory deficient at the end of 38 months was tracked.

panded in the absence of developmental contribution (against $0.138 \pm 0.0121\%$, in case of normal simulations). In accordance with the experimental data [296], the above analysis also indicated a phenomenal propensity of mutations with larger replicative advantage, in clonally expanding.

Considering the swiftness of clonal expansion dynamics of mtDNA mutations with larger replicative gain factor, subsequent simulation were performed to estimate the distribution function of time at which cells having fixed proportion of mutated mtDNA at the end of development with a fixed value of β , became completely COX deficient ($F_M > 90\%$) at the end of life span. Figure 4.10C indicates that template fixation time of mtDNA (time required for cells to become completely COX deficient) decreases with increasing β value

of the mtDNA mutations and the variability in fixation time decreases exponentially with increasing β value.

4.3.3 Cluster size distribution of clonally expanded cells

Several works quantifying mtDNA somatic mutations at individual cell level have revealed the presence of mutant cells/skeletal fiber clusters having same single clonally expanded mutations [303, 304]. Such cluster formations of mutant cells, have been observed in human brain [305], heart [288] and murine liver [115].

Mutant cloning is an attractive hypothesis, which explains the formation of mutant clusters. In this interpretation, the clonal mutations in a cluster are proposed to have originated from a single mutational event in a precursor cell, which generates the mutant cells belonging to that cluster in the subsequent cell lineage. Interestingly, the time of such mutational events play a significant role in determining the size of mutant cluster cells. For example, a precursor cell of late origin, like satellite cells of skeletal muscle or colonic crypt cells [29] or bronchial epithelium turnover units [299], is likely to produce local clusters. In contrast, early mutant precursors generated during developmental cell lineage is likely to have enormous cluster size, spanning from a part to an entire tissue, depending on how early in development the mutation was acquired [270]. At multi-cellular level several studies have reported mutant cluster sizes varying from few tens to hundreds of cells [306, 307]. Consistent with the experimental evidence, simulation results in Figure 4.11, considering the model developed in this work (hypothesis D), indicates that the distribution of clonally expanded mutant cluster cells in the normal aging heart tissue of rats ranges from a single cell to few hundreds of cells.

One of the most intriguing questions pertains to the time at which clonally expanding mutations get seeded. Clonal expansion may take a long time to get phenotypically relevant [270]. This implies that mutation events that seeded clonal expansion might have

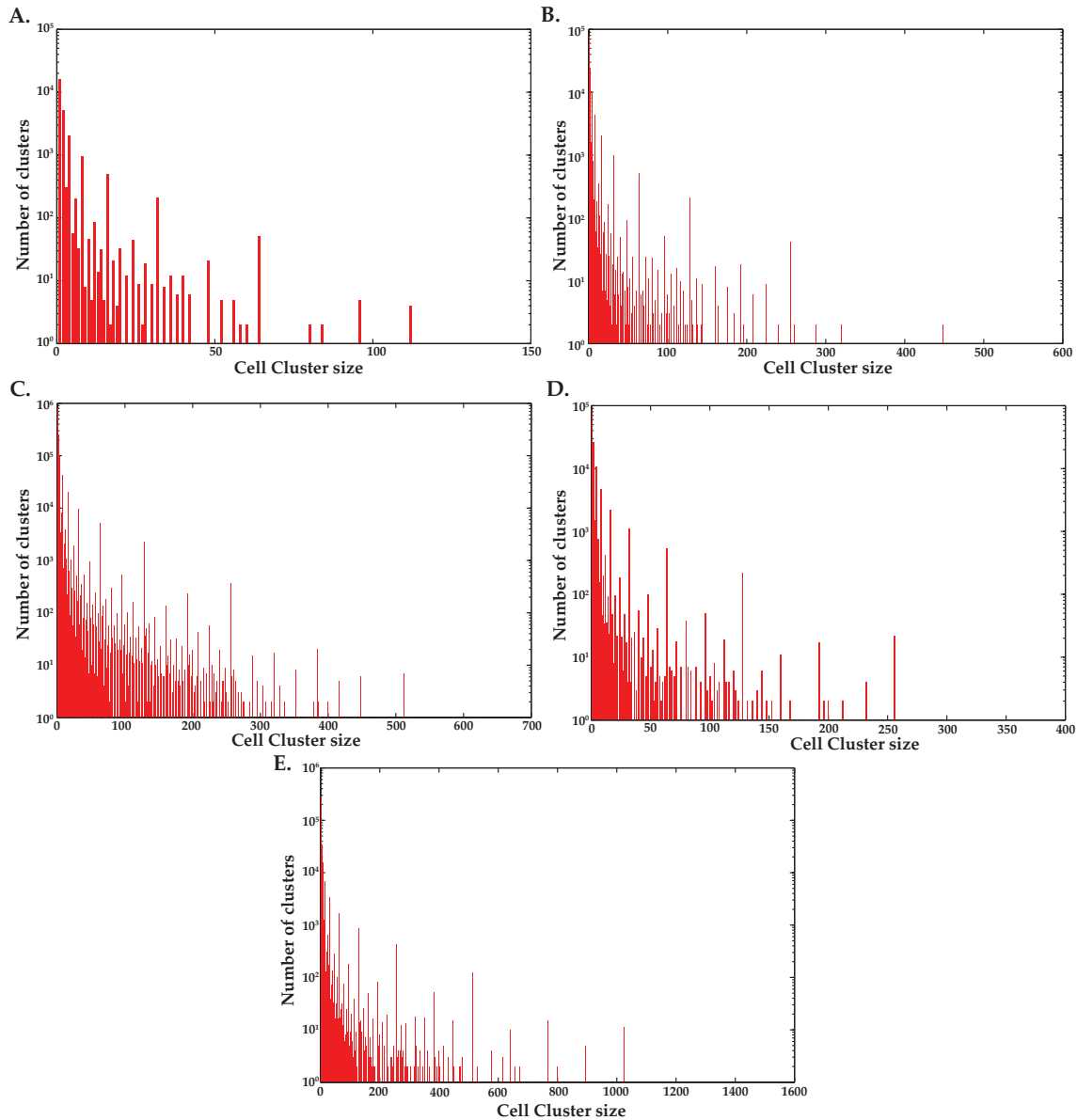


Figure 4.11: *in silico* perturbation analysis to understand the influence of different parametric variation on the origin and dynamics of mtDNA mutation clusters in the murine tissue. Simulations were performed using the model and parameters of hypothesis D. Influence of parametric variation of mtDNA turnover process on the cluster size distribution of heart cells harboring clonal mtDNA mutations. Progenitor cells with different mtDNA count refer to the mtDNA copy number in the early progenitor cells at the start of development. Influence of different mtDNA turnover rate on the cluster size distribution was also studied. A.) simulation results with progenitor cells having initial mtDNA copy number of 200 ($W_0|^{dev} = 200$), B.) simulation results with progenitor cells having initial mtDNA copy number of 1000 ($W_0|^{dev} = 1000$), C.) simulation results with progenitor cells having initial mtDNA copy number of 1000 ($W_0|^{dev} = 1000$), simulated with a *de novo* mtDNA mutation rate of $k_m = 10\times$ the rate used in the present work (Table 4.1), D.) simulation results with progenitor cells having initial mtDNA copy number of 1000 ($W_0|^{dev} = 1000$), simulated with an mtDNA turnover rate of 10 days [218], and E.) simulation results with progenitor cells having initial mtDNA copy number of 3500 ($W_0|^{dev} = 3500$).

happened long before the expansion is complete, potentially during development. Massive replication bursts of mtDNA during embryonic development can potentially provide

excellent opportunity for *de novo* mutations. To characterize the underlying distribution of the embryonic cell division stages, at which the earliest *de novo* mutations arise, 10^4 independent simulations of binary cell divisions during development were performed using the model developed in this work (hypothesis D). The simulation here was slightly modified such that the cells having only single *de novo* mutations that end up becoming clonally expanded in the post natal stage was considered. The cells that harbor more than single *de novo* mutation during the animal life was excluded from the analysis, as this forms a very insignificant portion of the tissue (as can be seen from the panel B of Figure 4.12, the fraction of such cells only form less than 10^{-4} %). In each simulation, the earliest stage at which the first *de novo* mutation that resulted in clonal expansion was tracked. Figure 4.12A illustrates the underlying distribution of the first occurrence of the *de novo* corresponding to the respective developmental stages. Figure 4.12A, illustrates that the density function has a long tailed distribution and seeding mutation can arise as early as the first stages of the embryonic division. It also indicated that most of the first *de novo* mutations occurred between the 9th and the 13th stage of cell division (first quartile).

To understand whether the clonal expansion was a result of multiple mtDNA mutations or a consequence of single clonal expansion, the number of *de novo* mutations arising per cell was tracked. Figure 4.12B, illustrates that most of the cell contains at most single *de novo* mutations per cell at the end of life. A minor fraction cells contained more than a single *de novo* mutations being generated during mouse's life span ($< 1\%$ of the simulated cells). This implies that the repeated occurrence of a particular mutation in a given cell is not due to the independence in generation of the mutation (like in hotspots, mutations that frequently occur), but rather clonal expansion of one/few seeding mutations. If the mutational events in mtDNA population were triggered by the ongoing oxidative assault, as postulated by the mFRTA, one would expect a large spectrum of mtDNA mutations in individual cells. On the contrary, biochemical analyses of mutated cells have indicated a

predominance of a single type of mutations [270, 289, 290]. This is further confirmed by simulations results in Figure 4.13, where the mutation rate (k_m) was set three orders of magnitude higher than the normal value, to simulate mutation hotspot (mutations that frequently occur). These mutations were simulated such that they caused functional deficiency in the OXPHOS process (sets in the nuclear-retrograde response), but did not possess any replicative advantage. Results indicated an elevation in overall deletion load in the tissue. However, even at such high mutation rate no cells in the tissue ever became COX deficient.

Since the mtDNA *de novo* mutation rate is very low ($k_m 5 \times 10^{-7}$ per replication), the number of *de novo* mutation arising per cell is close to one, in majority of cell population, immaterial of the mtDNA turnover rate. For example considering a higher mtDNA turnover rate of 10 days (Figure 4.12B), the spectrum of number of *de novo* mutation arising per cell, is very similar to results obtained considering the long turnover rate of one

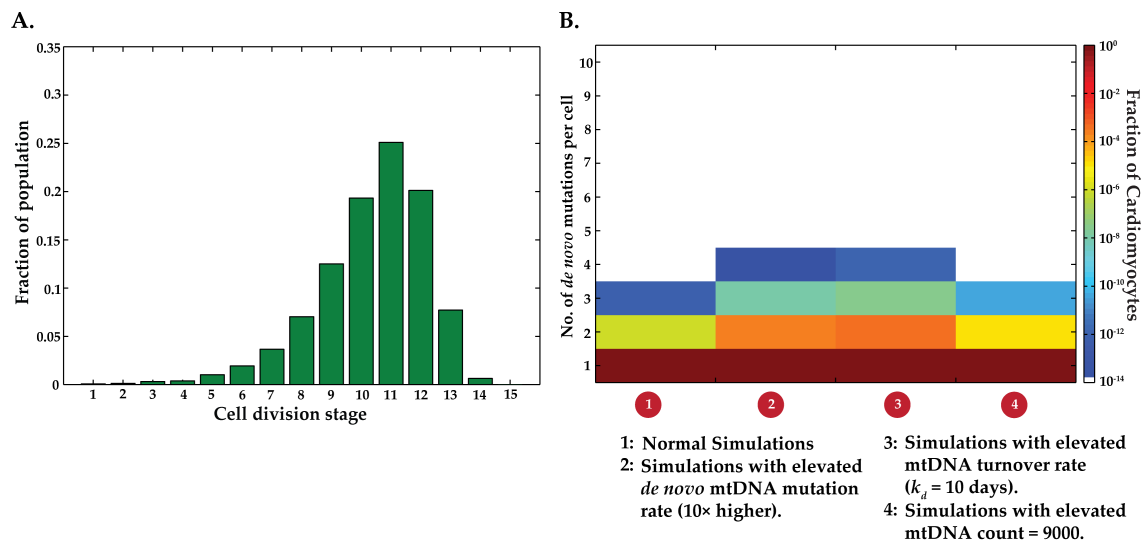


Figure 4.12: Results on the origin and accumulation of *de novo* mutation in clonally expanded mutant cells during mouse life span. Simulations were performed using the model and parameters of hypothesis D. A.) Distribution of developmental cell division stages at which the earliest *de novo* mutations, fated to clonally expand. B.) Influence of parameteric variation of mtDNA turnover process on the number of *de novo* mutations arising in single cells of mouse heart tissue at the end of 38 months of age. Simulation results presented in 4. represents post mitotic tissue simulation with the post natal cardiomyocytes having mtDNA copy number of 9000. In this simulation, it was assumed that the developmental progenitor cells start with 1000 wild-type mtDNA templates ($W_0 = 1000; M_0 = 0$) and following the birth, the mtDNA population in the skeletal progenitor cells (stem or satellite cells) underwent hypertrophic growth such that the mtDNA count in the cells increased to 9000 [292, 293].

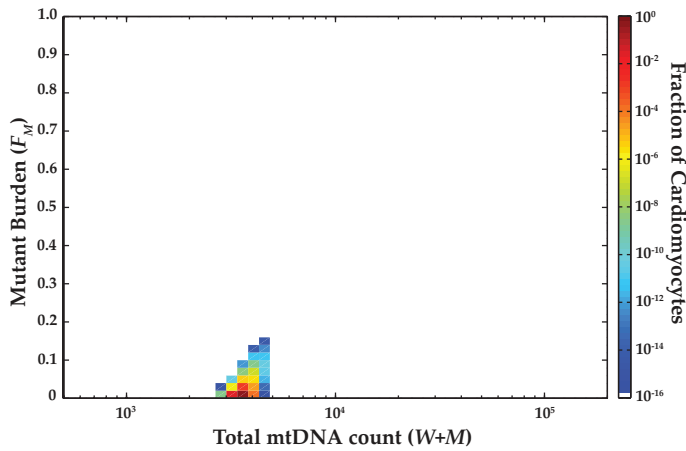


Figure 4.13: **Parametric perturbation analysis on mouse heart ($n = 3$). Simulations were performed using the model and parameters of hypothesis D.** State space probability density function of Total mtDNA count ($W + M$) and mutation burden F_M in mouse cardiomyocytes of the complete heart tissue in the absence of only the mitochondrial replicative advantage (i.e. in the presence of nucleo-mitochondrial retrograde feedback). Simulation results obtained with *de novo* mutation rate (k_m) $1000\times$ the normal mutation rate [Table 4.1](#). All the other parameters are kept at the same level as used for the present work.

year used the present dissertation. This highlights the importance of developmental stage in the mtDNA mutation accumulation dynamics in somatic tissue. This can be further understood by considering the quanta of replication during the animal developmental and post natal stages. The total number of replications during development is comparable to that during the entire adult life. In mice, the heart tissue develops in about 20 days [244]. Considering the degradation rate described in [Table 4.1](#) and the mouse heart to contain 2.5×10^7 cardiomyocytes [230, 248] arising from 22 cell divisions (6 progenitor cells), the total number of mtDNA replications needed to maintain homeostatic value of mtDNA ([Table 4.1](#)) [230] per cell should exceed 9×10^{10} times during the development. On the other hand, from the degradation rate of mtDNA in postnatal stages ([Table 4.1](#)) [240], the number of mtDNA replications events over the three years lifespan of mice is about 1.3×10^{11} . Thus, it is evident that the developmental period may carry comparable contributions in *de novo* mtDNA mutations as does the entire adult life.

4.4 Conclusions

Mitochondrial DNA mutations, especially deletions, are commonly associated with cellular aging and mitochondria-associated degenerative diseases. Recent progress in the de-

velopment of protocols for estimating mutation burden at single cell-level has highlighted the influence of random mitochondrial genetic drift at the cellular level on tissue/organ - level morbidity. Biochemical analyses of respiratory deficient cells have indicated abundance of single type functional mtDNA mutations, strongly hinting of the existence of a replicative mechanism favoring one type of mtDNA template. However, the understanding of functionality of such mechanism is still incomplete. Several hypotheses have been proposed to explain the underlying clonal expansion mechanism. Considering the discreteness of mtDNA mutation dynamics (mutation mosaicity) in post mitotic tissues, a stochastic model of mtDNA turnover dynamics in mouse cardiomyocytes have been developed in this work based on CME, to test different clonal expansion hypotheses and to elucidate underlying mechanisms which best explains the experimental data.

Simulation results indicated that the mosaicity seen in cellular mtDNA mutation burden in post-mitotic tissue arises due to stochasticity in mitochondrial turnover process, predominantly during developmental cell lineage. However, for explaining the mtDNA mutation expansion observed in post-mitotic tissue, both the replicative advantage of mutant mtDNA and the nuclear retrograde response, which is modulated by the cellular energy deficit, were required. Replicative advantage of mtDNA mutations, which is expressed as a replicative gain factor (β) in this work, was found to play a significant role in the dynamics and severity of clonal expansion process. Consistent with experimental observation, simulation results indicated that, mtDNA mutations with larger replicative gain factor (β) has a higher propensity to clonally expand. An accurate understanding of the underlying mechanisms which render these mutations, the replicative advantage compared to the wild-type mtDNA and other types of mtDNA mutations, and the quantification of these replicative gain factors (β) is an important area, which requires further investigation.

Accumulations of mtDNA mutations in the post-mitotic cells arise either due to mutant

cloning or because of persistent *de novo* mutant generation. Model predications indicated that the mutant cloning due to replicative advantage of mutant mtDNA is the most likely mechanism for explaining the clonal expansion process observed in non-replicative, post mitotic tissues. Also, *de novo* mtDNA mutations during developmental cell division are critical in determining the mutant cell cluster sizes. Simulation results of mtDNA mutation accumulation in normal aging tissue, indicated that the size and distribution of mtDNA mutant cell clusters in normal aging tissue are consistent with the late origin of the seeding precursor cells during development or in early post-natal life. The stochastic model of mtDNA turnover dynamics developed in this work (model developed considering hypothesis D) is based on minimal number of conservative assumptions, and closely resembles the actual biological process. Due to which, the present model can be used as a work-bench to analyze the influence of different parametric perturbations on the origin and severity of clonal expansion process, thereby providing useful insight into the possible interventions to retard its severity and possibly provide a hypothesis-driven design of experiments to gain a better understanding of the role of mtDNA mutations in aging or mitochondrial related degenerative diseases.

Sarcopenia *in silico* — a multiscale stochastic-deterministic modeling approach

5.1 Introduction

In this chapter, different aspects of mitochondrial genetics like the stochastic origin of mtDNA mutagenesis and its subsequent clonal expansion are integrated into a tissue-level multiscale modeling framework to study the origin and accumulation of mtDNA mutations and its consequence in the progression of an important age related condition known as sarcopenia. Sarcopenia refers to the progressive loss of muscle mass and strength with age, which contributes to frailty in aged individuals. In 2000, the health care cost related to sarcopenia in the United States alone was estimated to be US\$18.5 billion, or roughly 1.5% of the total healthcare expenditures [308]. The muscle loss causes a negative effect on muscle strength, metabolic rate and respiratory function; in short, the mobility of elderly subjects is severely compromised. As the proportion of elderlies is increasing world-wide, their health problems and medical costs associated with old age, such as sarcopenia, are also expected to increase warranting further epidemiological and clinical research is warranted.

In general, there are two factors at play in the progression of sarcopenia: the initial muscle mass and its rate of decline with age. Although there are significant changes in body composition with age, the muscle mass is relatively stable between the age of 25- and 55-year in men. However, there is an average decline of 25% between 50 and 75-year old age [309]. In active individuals, exercises can delay the onset of sarcopenic symptoms, but once started, the rates of decline are similar to their sedentary peers [310]. This loss is

generally attributed to the reduced cross section of the muscle fiber, subsequently leading to loss of muscle fiber or fiber atrophy [311, 312]. Similarly, in rat quadriceps muscles, fiber number decreased by 58% between 18 and 36 months [151].

The aetiology of sarcopenia is not clearly understood, but several mechanisms have been proposed in explaining sarcopenia including: (i) irreversible fiber damage or permanent denervation resulting from a loss of contact between nerve and muscle fiber [313, 314], (ii) Contraction-induced injuries to the muscle fiber [315], (iii) changes in tissue response to nutrients and/or malnutrition on protein metabolism [316, 317], (iv) deficiency arising due to the alterations in protein turnover [316], (v) alterations in the endocrine system (for example, variations in the growth hormone and testosterone, increase in cortisol and cytokines) or due to impaired tissue response to hormonal stimuli [318, 319], (vi) changes in satellite cell recruitment [320], (vii) apoptosis, and (viii) increase of free radicals mediated oxidative stress [151, 321] and age associated accumulation of mitochondrial abnormalities (e.g., mtDNA deletion mutations and other ETC abnormalities) [322, 323]. This work relates to the last hypothesis in which death of muscle fibers occur progressively due to the gradual dysfunction of mitochondria.

Evidence for mitochondrial aspect in development of muscular sarcopenia is observed in several histochemical tissue analyses showing focal accumulation of mtDNA deletion mutations in human and rat skeletal and cardiac tissue (e.g. [150, 216, 217, 324, 325]). In these studies, a defined number of muscle fibers were analyzed for the presence of mtDNA deletions. These analyses indicated that the deletion amplification product decreased with increase in the quantity of fiber analyzed [326]. Thus, these trials demonstrated that mtDNA deletions are not distributed evenly throughout muscle, but rather focally accumulated to high levels in a subset of fibers.

The accumulation of mtDNA mutations in these cells, likely involves stochastic factors such as the inherent random nature of mtDNA mutation accumulation. Stochasticity

of somatic mutation loads manifest in the form of large heterogeneity amongst individuals of same population [185, 269]. Deletions have also been found to be distributed non-uniformly between other tissues [153] and between individual cells of the same tissue [185]. Particularly a fraction of neuron cells in brain (substantia nigra) harbor a few orders of magnitude higher deletions than others [27, 154].

Devising treatment methods for sarcopenia thus requires a critical understanding of these stochastic factors and dynamics of mtDNA mutation accumulation in the muscle fibers. Limited data are available on the stochastic origin and dynamics of sarcopenia and the possible interventions to reduce its progression in vertebrate animal models [327]. This is due to significant challenges in obtaining such data for long-lived mammalian animal model, such as high cost and invasive nature of harvesting tissue sample for clinical trials. To characterize the influence of stochastic dynamics of mtDNA mutation accumulation on the mosaicity observed in the muscle fabric, a hybrid multiscale mathematical models is developed which describes: (1) the expansion of mtDNA mutation in a myo-nuclear region, (2) the spread of mtDNA mutations across a muscle fiber, and (3) the associated progressive loss of muscle strength. The hybrid model developed in this work uses both deterministic and stochastic algorithms and can be further developed as a work-bench to analyze the influence of different parametric perturbations on the development and progression of sarcopenia. This approach can further provide useful insights into the possible interventions strategies to retard its severity.

5.2 Methods

Skeletal muscle is made up of several individual components called as the muscle fiber. The muscle fiber is long, cylindrical in shape consisting of multinucleated fused cells. The muscle fiber is composed of repeated units of actin and myosin and mitochondria reside in the region between these repeating units. Each unit is referred to as sarcomere, which

forms the basic functional unit of muscle fiber and is responsible for striated appearance of the skeletal muscle fiber and is the basic machinery necessary for muscle contraction. The muscle fibers are bundled together in a plasma membrane called as sarcolemma.

In mice, it has been found that during development, skeletal muscle mass increase by 50-fold and this causes almost 50% increase in total mass of the organism at birth [328]. This arises due to the action of a population of muscle stem cells known as satellite cells, which come from the embryonic somatic progenitor cells or the muscle myoblast cells [328, 329]. During myogenesis, the primitive cells are arranged in clusters and they further undergo fusion to form small multinucleated cells known as the myotubes. Elongation of the myotubes occurs by end-to-end fusion of the adjacent tubes [330, 331]. Finally, these earliest differentiated myotubes are bundled together forming the muscle fibers. The multiscale model, developed in this work was simulated to mimic the basic shape and attributes of the muscle fiber and included the multiscale nature of the reaction-diffusion processes of mtDNA in the fused myonuclear regions. At this level of modeling, a hybrid stochastic-differential formalism was used. Following the earlier models (Chapter 3 and 4; [269]), the sarcopenia model comprised of tracking mtDNA mutations in mouse skeletal muscle tissue (Vastus lateralis) during two stages of life: developmental and post-natal stage (Figure 5.1). Sections below provide details of the hybrid stochastic-differential framework developed in this work.

5.2.1 Stochastic model

During development, the simulation of the origin and accumulation of mtDNA mutations were modeled, in the muscle progenitor cells or the myoblasts, considering these cells as independent entities. Consistent with the stochastic model described in the previous chapter (Chapter 4), the mtDNA turnover process (Figure 5.2), in each of these cells was modeled using stochastic chemical kinetics based on CME and solved using the SSA

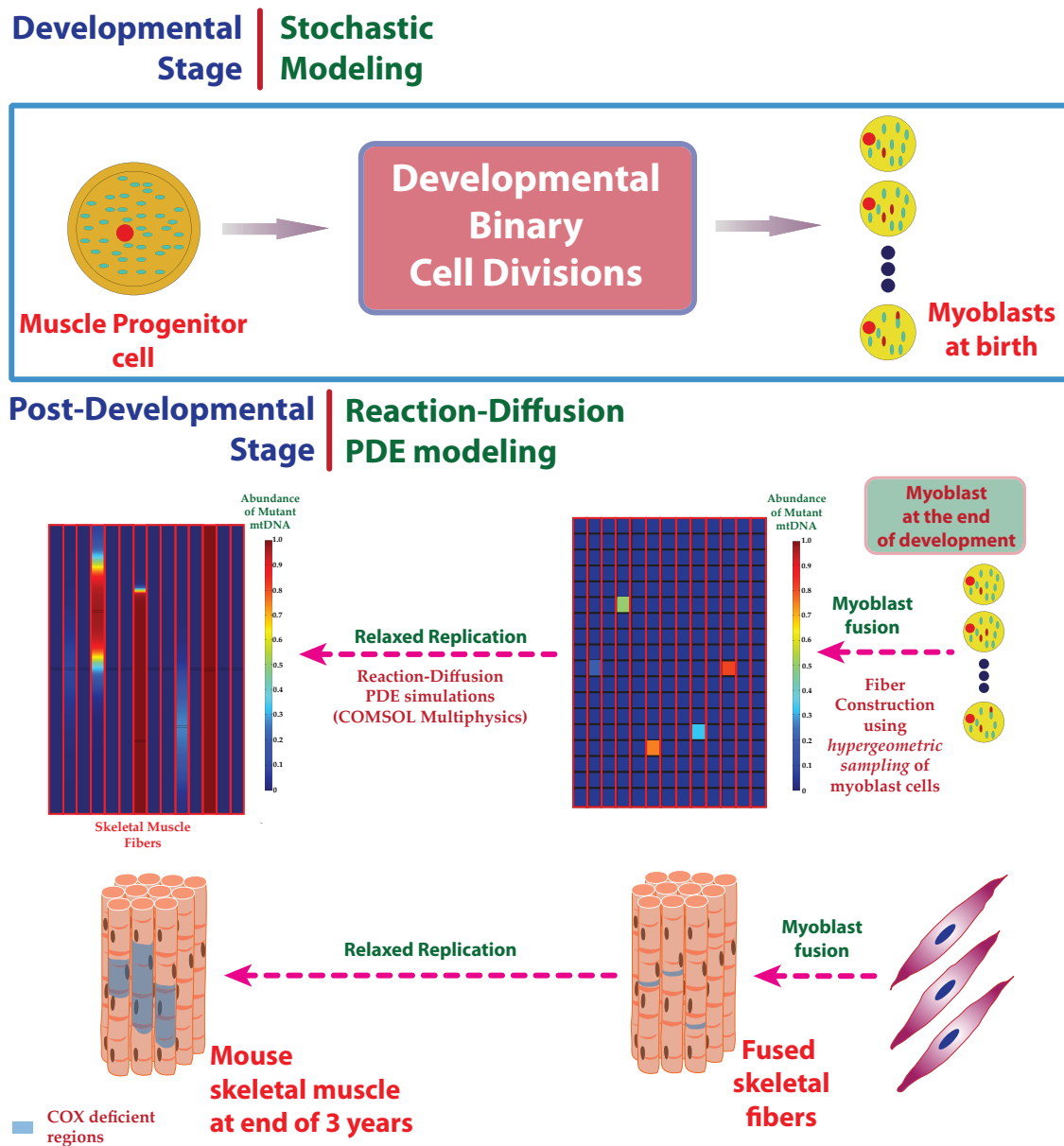


Figure 5.1: Details of the hybrid (Stochastic-Deterministic) model of mtDNA turnover process in mouse skeletal muscle tissue. During the mouse development the mtDNA turnover was simulated in the individual muscle progenitor cells or the myoblast cells as discrete stochastic process as illustrated in Figure 5.2. The stochastic *in silico* model tracked for the mtDNA mutation burden in the muscle myoblast cells during the skeletal tissue development. The modeling details of the stochastic mtDNA turnover process are consistent with earlier chapters (Chapter 3 and 4; [269]). In the Post-natal stage, the process of fusion of myoblast cells to form skeletal muscle fiber was simulated by fusing the muscle progenitor cells obtained at the end of development and the sampling of the muscle progenitor cells was done using the hypergeometric sampling method (discussed in Methods section). The dynamics of mtDNA mutation burden in the post natal skeletal muscle fiber was simulated using a deterministic reaction-diffusion process.

(Chapter 4; [269]). The model details related to the stochastic mtDNA turnover process is detailed in Chapter 4.

5.2.2 Reaction-diffusion model

In post-development simulations, the individual developmental myoblast cells were selected randomly and subjected to fusion to form the muscle fiber in Figure 5.1. The probability of choosing a cells containing mtDNA mutation is obtained by hypergeometric distribution function [233],

$$f(x) = \frac{\binom{n_{\Delta mtDNA}}{x} \binom{N_{tot}}{n_{Fib} - x}}{\binom{N_{tot}}{n_{Fib}}} \quad (5.1)$$

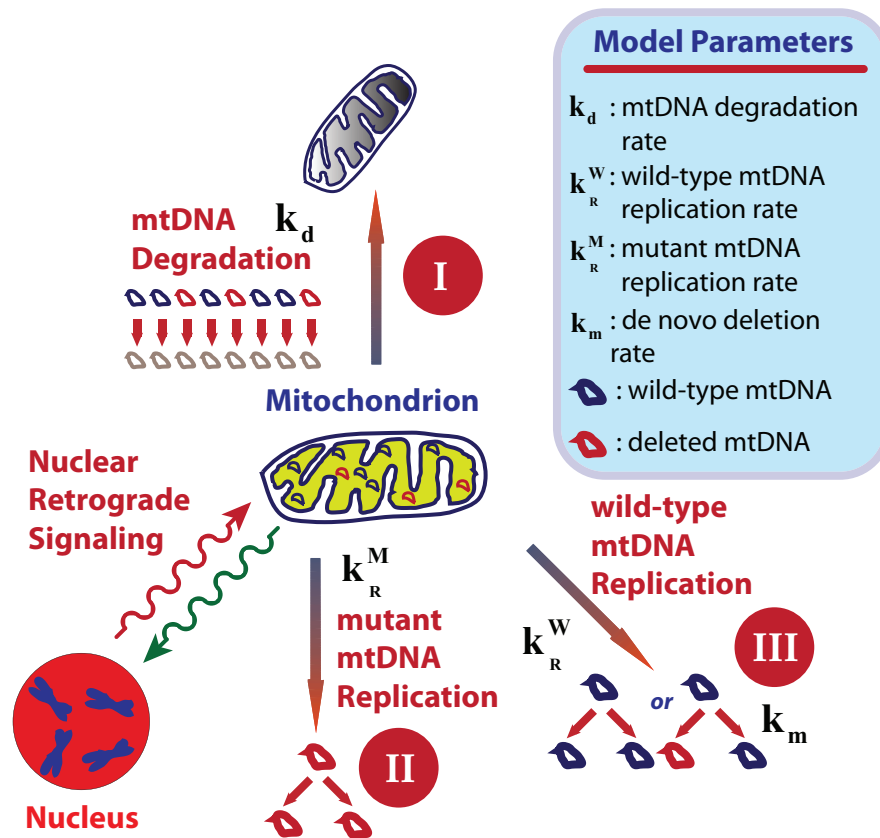


Figure 5.2: **Stochastic model of mtDNA turnover dynamics in a mouse skeletal muscle progenitor cells (myoblast cells).** The stochastic drift in mtDNA mutation dynamics results from following random processes. (I) The mitochondrion that undergoes a turnover event is randomly selected from the population. The autophagy of mitochondrion was simulated by removing all the mtDNA molecules associated with the selected mitochondrion. (II) Replication of a single mtDNA molecule simulated by random selection of mtDNA from the mtDNA population. (III) During the wild-type mtDNA replication, a finite probability equal to the *de novo* mutation rate (k_m) is assigned for the replication process to give rise to a mutant mtDNA. Since the mutations modeled in this work was assumed to have functional consequence to the cellular OXPHOS activity, the presence of mtDNA mutant was assumed to initiate the nucleo-mitochondrial retrograde response [231], which, similar to the previous chapter (Chapter 4), is modeled using Hill-type kinetics.

where $n_{\Delta mtDNA}$ is the number of cells with mtDNA mutations in them, N_{tot} is the total number of satellite/stem cells at the end of development and N_{Fib} is the number of satellite/stem cells fused to form a single muscle fiber.

A morphometric analysis of murine skeletal muscle has been performed to understand the changes that occur in the population dynamics of muscle cell nuclei and satellite cell nuclei in rat skeletal muscle fiber [332]. In this analysis, an empirical relation was deduced to obtain the average number of nuclei per muscle fiber:

$$N_{nuc} = \frac{f_{CS}^{nuc} \cdot L_{fib}}{d} \quad (5.2)$$

where, f_{CS}^{nuc} is the frequency of average nuclei per fiber cross section [332], L_{fib} is the average length of the muscle fiber, and d is the average length of the myonuclear regions [332]. In this work, this empirical relation described in this equation was used to determine the average number of myoblast nuclei required to create a given length of muscle fiber. The corresponding muscle characteristics and the muscle fiber dimensions used for the simulations are described in Table 5.1.

The multiscale fiber model developed in this work tracked the mtDNA population (W and M), in the myonuclear domains along the length of muscle fibers with age. The governing mass transport equation that describes the dynamics of reaction-diffusion of

Table 5.1: Morphometric details of the *Vastus lateralis* murine skeletal muscle tissue, used for hybrid stochastic-deterministic modeling.

Muscle Type	<i>Vastus lateralis</i>	
Muscle Characteristics		
Average total fiber count N_{tot}	10870	[217, 325, 333]
Average length of muscle fibers L_{fib}	4 cm	[217, 325, 333]
Frequency of average nuclei per fiber cross section f_{CS}^{nuc}	0.64	[332]
Average length of the myonuclear region d	12.75 μm	[332]
Muscle dimensions		
Average length of single myo-nuclear domain	25 μm	[332]
Average diameter of the fiber	63 μm	[217]
Number of myo-nuclear domain per fiber	1929	

mtDNA in the fiber is:

$$\frac{\partial}{\partial t} c_i(z, t) + \nabla \cdot (-D_i \nabla c_i(z, t) + c_i(z, t) \mathbf{u}) = R_i \quad (5.3)$$

where $i = W, M$, z is the spatial coordinates along the length of a fiber. Due to the symmetry of fiber, axi-symmetry spatial coordinated system was used for modeling, $c_i(z, t)$ denotes the concentration of mtDNA (W and M) and D is the effective diffusivity constant of mtDNA. The reaction term R_i represents the mtDNA transformation arising due to the mitochondrial turnover process. Assuming the mtDNA turnover dynamics to be a pure diffusion process (negligible convection), Equation 5.3 becomes:

$$\frac{\partial}{\partial t} c_i(z, t) = D_i \nabla^2 c_i(z, t) + R_i \quad (5.4)$$

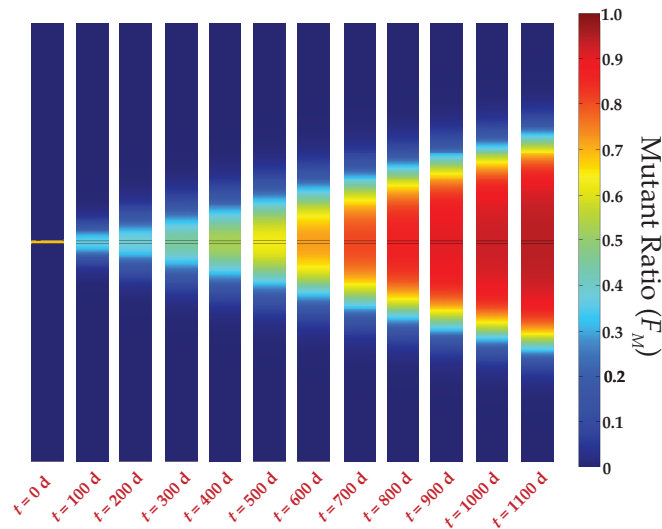
Mitochondria move in a complex way in a living cell, transported by the actin/myosin and dynein/kinesin based motor along the microfilament and microtubule networks [334]. Experimental data indicated that mtDNA nucleoid foci are tethered through mitochondrial membrane to the microtubule networks. Recently, an analysis using fluorescent dyes was performed to understand the movement of mtDNA in living human fibroblast cells. A mean square displacement (MSD) analysis and subsequent autocorrelation analysis between the displacement distance and the time span difference between these displacements indicated that the movement pattern of mtDNA nucleoids typified a random walk [204]. Furthermore, based on the MSD observations, an apparent diffusion constant for the mtDNA foci displacements was estimated to be $1.1 \times 10^{-3} \mu\text{m}^2/\text{s}$, which was also comparable to the mtDNA foci diffusion observed in *Drosophila* cells [335]. For the reaction-diffusion model developed in this work this apparent diffusion constant of the mtDNA was used. In accordance with the reactions in Equation 4.1, the evolution of mtDNA states was updated (R_i in Equation 5.4) by tracking the rate of change mtDNA concentration as:

$$R_W = k_R \cdot F_W \cdot (1 - k_m) - k_d \cdot W \quad (5.5a)$$

$$R_M = \beta \cdot k_R \cdot F_M + k_m \cdot k_R \cdot F_W - k_d \cdot M \quad (5.5b)$$

The parameters used for the simulation, have their usual interpretation and the rate constants used for the ODE simulations above are the same as the parameters used for the stochastic model (see Table 4.1). One instance of the multiscale simulation result related to the progression of sarcopenia in a single fiber is illustrated in Figure 5.3. All simulations were performed using an IBM high performance computing cluster with 110 Intel 1.6 GHz processors. The stochastic code was compiled using GNU C++ compiler G++ (v4.1.1) and run on CentOS (RHEL) Linux platform. The PDE model was solved using COMSOL Multiphysics 3.5a. The model report of the PDE formulation used in COMSOL is presented in Appendix B.

Figure 5.3: Progression of the COX deficiency in the skeletal muscle fiber with age, from the onset to the complete dysfunction of muscle fiber. The COX deficiency is represented by the mutant burden F_M in the muscle fiber. The simulations were performed with a fiber segment having a mutant progenitor with initial condition of $W_0 = 1717$ and $M_0 = 3800$.



5.3 Results & Discussion

5.3.1 Simulation & validation Results

The role of mtDNA mutation accumulation in the phenotypic manifestation of sarcopenia in mammalian skeletal muscles is ascribed by the following mechanisms: a.) *de novo*

mutations occur in a segment of muscle fiber due to some, as yet unknown mechanism, b.) these mutations then begin to expand and the expansion generally occurs along the length of the skeletal muscle fiber, c.) the accumulation of mtDNA mutations in the muscle segment causes the disruption of mitochondrial encoded ETS complexes. The loss of OXPHOS activity in the presence of wild-type genome suggest that the dysfunctional transcripts generated from the deleted mtDNA genome may compete with the functional transcripts generated from normal wild-type mtDNA, resulting in the eventual breakdown of the OXPHOS system; d.) the cellular respiratory deficit caused due to the accumulation of dysfunctional mitochondrial transcripts sets in nuclear retrograde signaling, which enhances mitogenesis. e.) The enhanced mitogenesis results in clonal expansion of mtDNA mutation, presumably due to replicative advantage of mtDNA mutations. This increased mitochondrial biogenesis and the associated increase in the mitochondrial content in the affected area gives the muscle fiber a red ragged appearance (red ragged fiber, RRF). f.) Both the energy deficiency and the oxidative damage would continue to accrue and this would eventually result in fiber atrophy and breakage [333].

The *in situ* hybridization techniques harness the effect of mtDNA mutations on the ETS respiratory complexes to indicate the respiratory deficiency in the cells [207, 336]. Using this technique, absence of partial mitochondrially encoded cytochrome c oxidase activities (COX) is commonly measured to indicate the presence of ETS phenotypes [333]. This is normally accompanied with an increase in the expression of entirely nuclear encoded succinate dehydrogenase (SDH) activity, which is an outcome of the elevated mitochondrial biogenesis. The combination of reduction in COX activity and elevation of the SDH phenotypes in the skeletal muscle represents the red ragged phenotype observed in the Gomori trichrome staining of the skeletal muscle [151]. In these histochemical analyses the complete muscle biopsy samples are prepared in a section length of 8-10 μm and 100-200 such sections are dissected and subjected to the COX-SDH analysis [333]. These histochemical

analyses are used to estimate different attributes of sarcopenic skeletal muscle tissue such as: the total number of fiber lost with age, the number of COX deficient regions in the skeletal fiber and the statistics of length of the COX deficiency in individual muscle fibers. In the present study, different experimental data related to the skeletal muscle histochemical analyses were used for the purpose of model validation (Figure 5.4).

Recently, by performing a longitudinal single fiber analysis, it was shown that the ETS abnormality due to mtDNA deletion accumulation resulted in fiber atrophy and breakage. This fiber breakage forms an important contributor to the muscular sarcopenia [26]. Based on this information, the number of fibers lost in the whole muscle tissue was estimated using the complete fiber simulations. The fiber count having mutant load $F_M > 0.9$ were tracked for estimating the number of fibers lost during the time of consideration. Figure 5.4A illustrates the number of muscle fibers lost in two different time points during the life span of rats and indicates a very good agreement with the experimental data. The progression of the muscle fiber loss in single heart tissue, with age is illustrated in the Figure 5.5 and it is interesting to see that the rate of fiber loss decreases with age.

Serial longitudinal cross section analysis of muscle fibers using histochemical analysis has indicated an average length of COX deficiency ranging from 450 μm to about 760 μm [324, 325, 333]. To determine the average length, complete muscle simulations were performed, and for each of the fiber being simulated the length of COX deficient regions having mutation burden of $0.6 < F_M \leq 0.9$ were tracked at the end of life span ($t = 1140$ d). Average of this data is presented in Figure 5.4B, indicating a good agreement with the experimental data.

In another histochemical analysis, the number of COX deficient regions in the complete cross section of the skeletal muscle was estimated for a fiber length of 2 mm [324, 333]. Simulations of this data were performed by hypergeometric samples of equivalent number of myocytes that constitute a single fiber of 2 mm muscle fiber (obtained using Equation

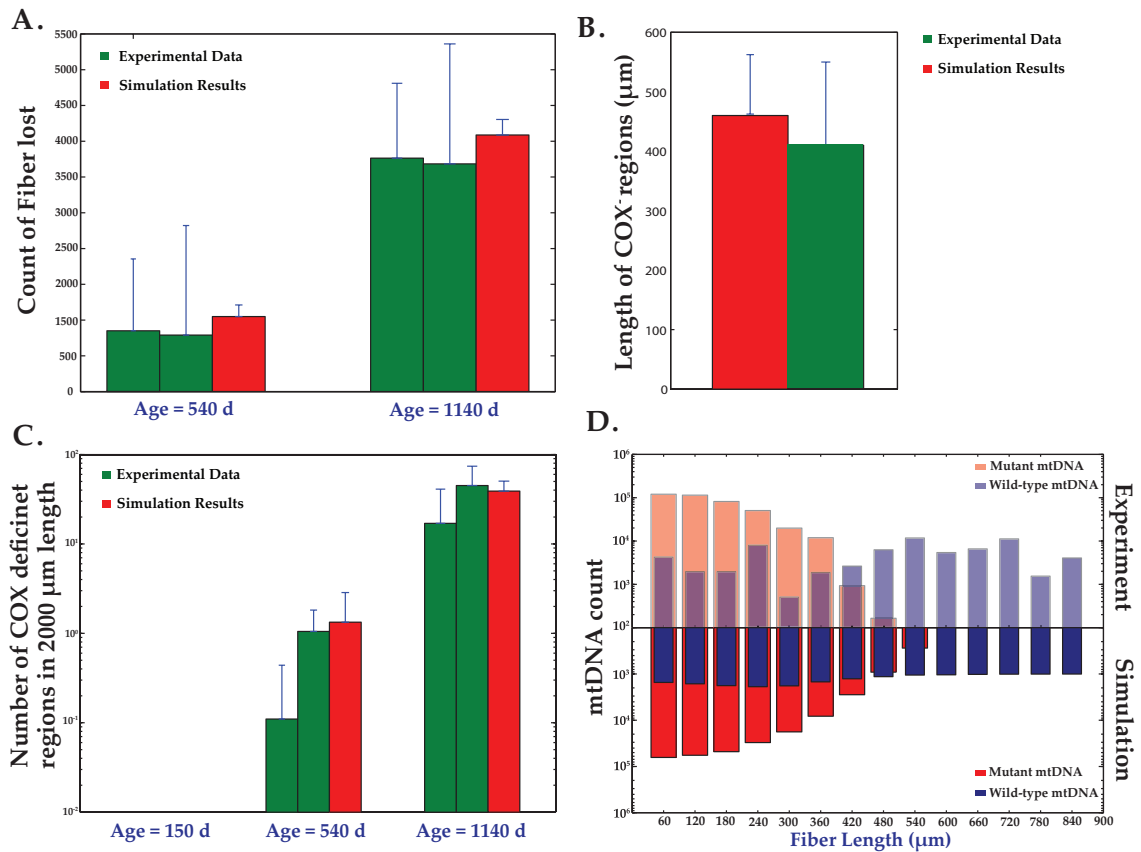


Figure 5.4: Validation of the simulation results with the data from the histochemical studies on rat skeletal muscle tissue (simulation statistics is based on five mouse tissue ($n = 5$)). A.) Comparison of the simulation results of number of fibers lost as a function of age, with the experimental data [151, 333], respectively. B.) Comparison of the simulation results of average length of the COX deficient regions, with the experimental data [217, 325, 333]. In each of these trials, the average statistics of the length of COX deficient regions in the skeletal muscle fiber was tracked until the age of 1140 days. C.) Comparison of the simulation results of the number of COX deficient region in a 2 mm fiber section of the complete muscle tissue with different experimental data [151, 217, 324], respectively. As shown in the figure, the analysis was carried out for three different age length. D.) Comparison of the simulation results of the variation of mtDNA count (both W and M) as a function of the fiber length with the experimental data [26]. The simulation results are illustrated in the bottom frame of the figure and the experimental data is illustrated in the top frame.

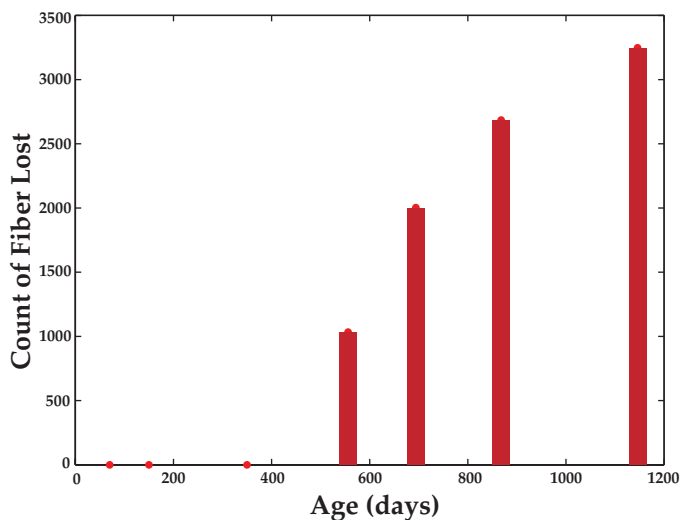


Figure 5.5: Simulation results on the number of fibers lost as a function of age. Fiber loss is tracked in a single mouse heart tissue as a function of age.

5.2) from the myoblast pool at the end of development and the number of red-ragged phenotypic regions ($0.6 < F_M \leq 0.9$) was estimated by simulating the reaction-diffusion process for a desired length of time. This was repeated for total fiber count in the tissue in Table 5.1. Figure 5.4C indicates that simulated COX deficient regions in the 2mm section of muscle is again in excellent agreement with the experimental data.

The absolute quantification of the abundance of deleted and wild-type mtDNA has been performed in the COX deficient and normal regions of fiber using breakpoint-specific quantitative PCR assays in laser-captured cell sections along the length of the muscle fibers [26]. The COX deficient regions in the analyzed fibers were flanked by ETS normal regions that contained detectable levels of deletion, albeit at lower levels. In fiber regions distant from the COX deficient regions, only wild-type mtDNA were present and mtDNA deletions were undetectable (top half of Figure 5.4D). The SDH hyperreactive phenotype mostly occurred in regions that contained greater than 90% of mtDNA deletions, and the wild-type mtDNA was consistently detected in both the ETS normal and abnormal regions [26]. This indicates that the accumulation of mtDNA deletion and not the selective depletion of wild type mtDNA is responsible for the mitochondrial COX deficiency [26]. Such characteristic has also been demonstrated in the earlier *in situ* hybridization studies from patients with mitochondrial myopathies [211] as well as aged muscle fibers [26, 337]. Simulation results also closely matched the profile of mtDNA states along the length of the fiber, with the experimental data. In accordance with the experimental observations [26], the level of wild-type mtDNA count was maintained in the COX deficient region at a similar level as in the ETS normal region.

5.3.2 Parametric perturbation analysis

Effect of mtDNA nucleoid diffusivity

Mitochondria are highly dynamic organelle and are able to adapt different shapes, localization and motility depending on the cell-type and the extent of metabolic demand of the cells. Strenuous physical activity like endurance/resistance exercise significantly alters the energy demand in skeletal muscle and is therefore expected to influence the mitochondrial dynamics. A recent experimental evidence indicated that the fusion-fission proteins like Mfn-1/2 and Fis 1, were significantly upregulated in human vastus lateralis muscle 24 hours post acute bout of exercise [338]. These findings indicate that mitochondrial fusion-fission protein expression could be rapidly altered in response to changing energy demand of the tissue. In mitochondrial fusion, it has been postulated that the elongation of mitochondrion enables rapid transmission of membrane potential and other mitochondrial metabolites across a great distance [339]. Inter-mitochondrial fusion and intra-mitochondrial mobility of endogenous nucleoids and mitochondrial respiratory complexes ensures functional complementation. Further, observations indicated that mtDNA nucleoid dynamics was closely tied to the mitochondrial fusion-fission process [95, 204]. After mitochondrial fusion, the mtDNA nucleoids have been found to be motile and interact with each other for genetic complementation [340]. Trials involving fusion of ρ^+ (cybrid cells with mitochondrial population) and ρ^0 (cybrid cells lacking mitochondrial population) cells have demonstrated that the mtDNA nucleoids had swift diffusion into the ρ^0 cells.

Based on these evidences, simulations were performed to understand the role of mtDNA diffusion on the dynamics and progression of sarcopenia in murine skeletal muscle. Simulations were performed using different effective mtDNA diffusion coefficient (D_{eff}) [204]. Results indicated that, faster diffusion of mtDNA, though intuitively not desirable as it enhances the spread of mtDNA mutation load along the skeletal muscle fiber, it is in fact ben-

eficial, as the overall mutation burden in the fiber is lower compared to the fibers having lower mtDNA diffusivity (Figure 5.6A). The fibers having faster mtDNA diffusivity, thus remain functional in mitochondrial respiration. In experiments, higher counts of respiratory functional muscle fibers have been seen in mitochondrial myopathic mice subjected to extensive endurance exercise [341]. Also, enhancement of mitochondrial nucleoid diffusion might explain the general beneficial effects of exercise observed in normal mammalian aging phenotypes [310, 341]. To estimate the number of fiber lost at the end of 38 months, complete fiber simulations were done using two different effective mtDNA diffusion coefficients. In this case, runs having higher mtDNA diffusivity resulted in lower number of atrophied fibers compared to the simulations considering lower mtDNA diffusivity (Figure 5.6B).

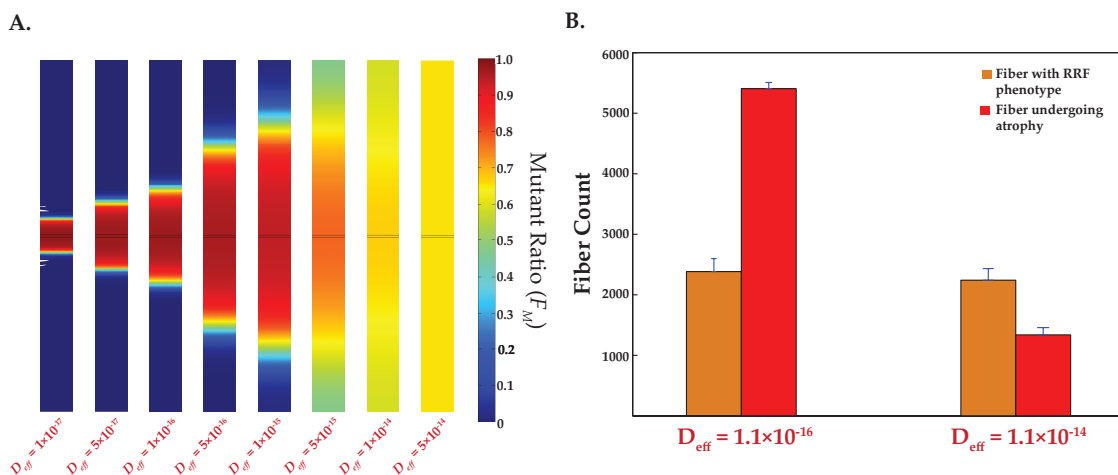


Figure 5.6: Effect of changes in effective mitochondrial diffusivity on the development and progression of COX deficient regions in muscle fiber (simulation statistics is based on five mouse tissue(n = 5)). A.) Single fiber simulation results illustrating the mutant burden (F_M) in fibers simulated with different effective mtDNA nucleoid diffusion constants (D_{eff}). The simulations for this part were performed with fiber segments having a mutant progenitor cell with initial condition of $W_0 = 1717$ and $M_0 = 3800$ and were performed for a time span of 1140 days. B.) Statistics (simulation statistics is based on five mouse tissue(n = 5)) illustrating the complete skeletal muscle simulation results of the number of fibers getting atrophied and the number of fibers having red ragged segments at the end of 1140 days, using different effective mtDNA diffusion constants (D_{eff}). The figure illustrates the number of fibers having COX deficient phenotypic regions in them (orange bar), and number of fibers that undergoes atrophy (red bars).

Significance of animal development

In [Chapter 3](#) and [Chapter 4](#), it was shown that organism's development play an important role in determining the stochasticity of mtDNA mutation burden between cells and tissues [269]. During the developmental binary cell divisions, mtDNA of embryonic cells undergo tremendous quanta of replication. Since the mutations are generally believed to result from errors arising during mtDNA replication [58], there is a high propensity for *de novo* mutations during development. Further, the *de novo* mutations arising during the developmental cell divisions may contribute in forming the seeds for the post-natal clonal expansion process. In order to elucidate the role of *de novo* mtDNA mutations acquired during the organism's development in progression of sarcopenia, single fiber simulation were performed using progenitor cells having only one mtDNA mutant molecule. Simulation results indicated that at such low burden in the early progenitor cells, the mutation burden in the post-natal fiber remained low at the end of life span ([Figure 5.7](#)).

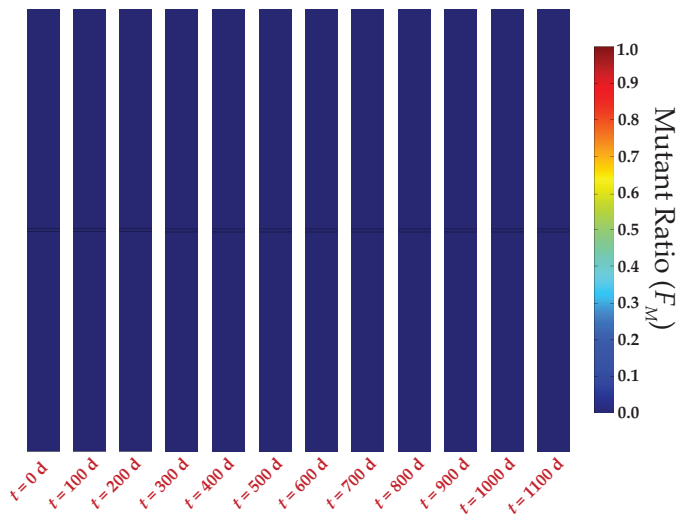


Figure 5.7: Progression of the COX deficiency in the skeletal muscle fiber with age, from the onset to the complete dysfunction of muscle fiber — Significance of animal development. The COX deficiency is represented by the mutant burden F_M in the muscle fiber. The simulations were performed with a fiber segment having a mutant progenitor with initial condition of $W_0 = 989$ and $M_0 = 1$.

Since the mtDNA *de novo* mutation rate is very low ($k_m 5 \times 10^{-7}$ per replication), the number of *de novo* mutation arising per cell is close to one, in majority of cell population, immaterial of the mtDNA turnover rate ([Figure 4.12B](#)). [Figure 5.7](#) highlights the importance of developmental stage in the mtDNA mutation accumulation dynamics in somatic tissue. This can be further understood by considering the quanta of replication during

the animal developmental and post natal stages. The total number of replications during development is comparable to that during the entire adult life. In mice, the heart tissue develops in about 20 days [244]. Considering the degradation rate described in Table 4.1 and the mouse heart to contain 2.5×10^7 cardiomyocytes [230, 248] arising from 22 cell divisions (6 progenitor cells), the total number of mtDNA replications needed to maintain homeostatic value of mtDNA (Table 4.1) [230] per cell should exceed 9×10^{10} times during the development. On the other hand, from the degradation rate of mtDNA in postnatal stages (Table 4.1) [240], the number of mtDNA replications events over the three years lifespan of mice is about 1.3×10^{11} . Thus, it is evident that the developmental period may carry comparable contributions in *de novo* mtDNA mutations as does the entire adult life.

Effect of mtDNA copy number regulation in cells

Embryonic cells committed for the somatic cell lineage demonstrate a large variation in the mtDNA content prior to 7.5-dpc [291]. The typical value of the mtDNA count at this stage ranges from 57–3345 mtDNA copies per cell. However, from 8.5 dpc onwards, both mtDNA copy number and variation of mtDNA content was observed to decrease sharply [291]. Mitochondrial DNA content was found to be vastly down-regulated in somatic cell lineage from the 7.5 dpc stage onwards. To understand if the down-regulation of mtDNA copy number in the embryonic cells committed to somatic cell lineage has any influence on the resulting mtDNA mutation burden in the post-natal somatic tissues, simulations were performed with different value of mtDNA copy number in the embryonic cells during developmental stage. In this work, it was hypothesized that the down-regulation of mtDNA copy number during the development may exist to moderate the overall mtDNA mutation load resulting in the post mitotic tissue. Since cells with higher mtDNA content has a greater quanta of replication, there is an associated increase in mtDNA *de novo* mutations, during development. All the other parameters associated with mouse heart tissue simulation, including the *de novo* mutation rate, were maintained at the original values. A

comparison of tissues formed from embryonic cells having an initial mtDNA copy number of $W_0 = 200$; $M_0 = 0$ copies per cell (Figure 5.8II) and $W_0 = 1000$; $M_0 = 0$ copies per cell (Figure 5.8I), indicates a 49% decrease in the number of fibers undergoing atrophy.

Mitochondrial DNA of post mitotic cells range from a few hundreds (e.g., in the typical epithelial cells) to several thousands (e.g., cardiomyocytes) per cell. The mtDNA copy number per cell is an indicator of the oxidative capacity of the tissue [342]. In other words, the higher the respiratory requirement of a tissue, the higher mtDNA copy number is likely to be associated with cells of that tissue. *In vivo* dynamics of mtDNA mutations in somatic cells have only been studied sparsely. Earlier studies on morphometric analysis of the mitochondrial density in different tissues from several mammalian species indicated a significant negative correlation between the mitochondrial density and the maximum longevity in mammalian species [342]. In contrast to the common belief that cells containing larger amount of mtDNA has lesser mutation burden associated with them [210], the morphometric analysis indicated that tissues with higher mtDNA content have higher mutation burden associated with them [342]. This is also consistent with the observation that metabolically active tissues with higher oxidative capacity (mitochondria-rich) have larger mutation burdens [139]. To conform this, simulations were performed with different hypertrophic growth of mtDNA in the skeletal muscle fiber, subsequent to development, keeping all the other parameter at the same level. Comparison of Figure 5.8I and 5.8III indicate that, at the same mtDNA turnover rate, a higher mitochondrial DNA contents correspond to higher fraction of fiber atrophy.

Effect of mtDNA mutant replicative advantage

ex vivo analyses of mtDNA deletions have indicated a marked replicative advantage of mutant mtDNA [99]. This phenomenon has important implications on the mechanisms of mtDNA deletions accumulation in aging and disease processes. For instance, deletions formed in transgenic mouse neurons due to a double strand break formed by the

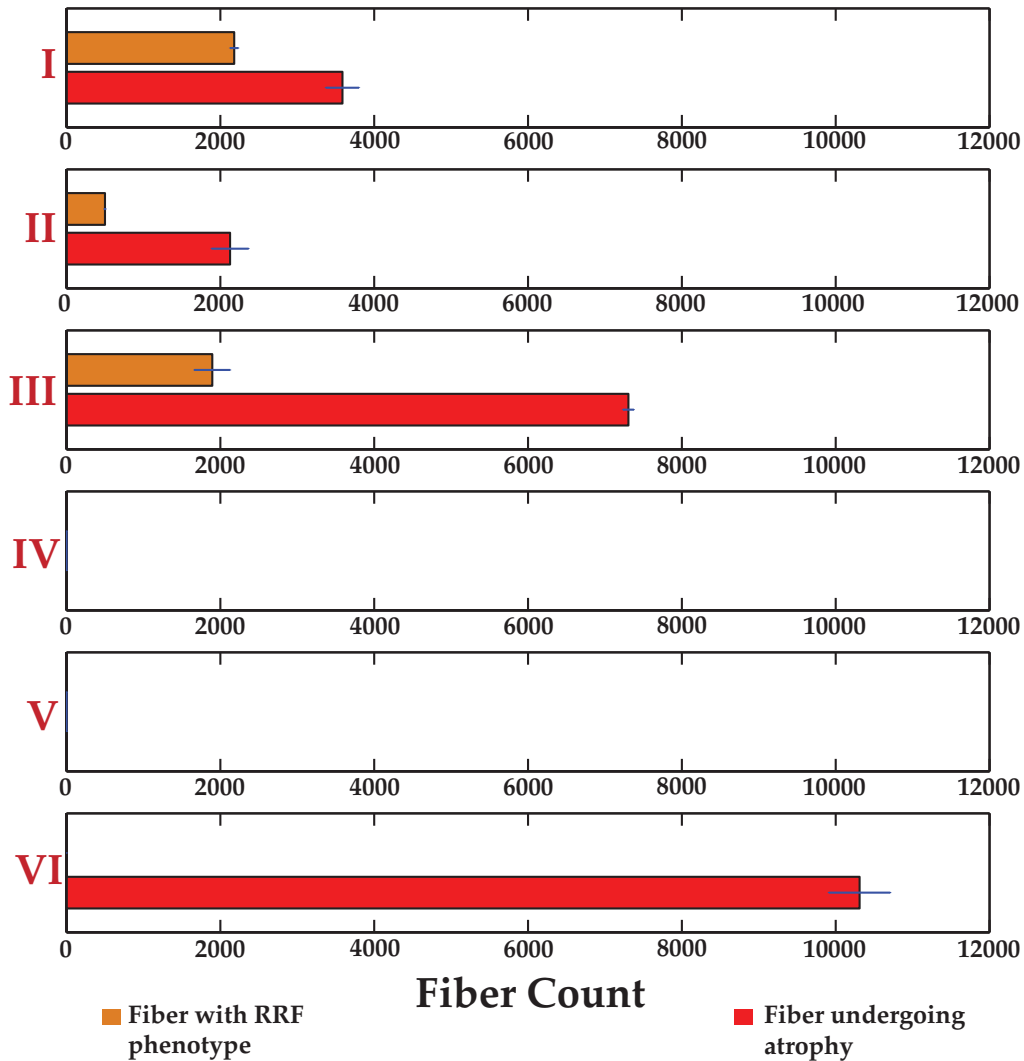


Figure 5.8: Results of different parametric perturbation analysis on COX deficient attributes of skeletal muscle (simulation statistics is based on five mouse tissue ($n = 5$)). Figure illustrates the simulation results of the complete skeletal muscle tissue simulation results in terms of the statistics of the number of fibers having COX deficient phenotypic regions (orange bar), and number of fibers that undergo atrophy (red bars) at the end of 1140 days. (I-III) Influence of mtDNA copy number variations on the overall mtDNA mutation burden and fiber statistics in the skeletal muscle tissue. Several hypothetical cases have been considered for the simulations: I.) results of the present simulation, II.) in this trial, it was assumed that the developmental progenitor cells have 200 wild-type mtDNA templates ($W_0 = 200; M_0 = 0$) and following the birth, the mtDNA population in the skeletal progenitor cells (stem or satellite cells) underwent hypertrophic growth such that the mtDNA count in the cells increased to 1000 [292, 293], and III.) in this trial, it was assumed that the developmental progenitor cells have normal count of mtDNA templates ($W_0 = 1000; M_0 = 0$) and following birth, the mtDNA population in the skeletal progenitor cells (stem or satellite cells) underwent hypertrophic growth such that the mtDNA count in the cells increased to 3500, and subsequently the elevated mtDNA count was used for simulating the post-natal reaction-diffusion dynamics of mtDNA in skeletal muscle fibers. IV-VI.) Influence of mtDNA mutation replicative gain factor (β) on mtDNA mutation dynamics and progression of COX deficient regions in muscle fibers. Plots indicate the simulation results obtained by considering different gain factors: IV.) $\beta = 0$, V.) $\beta = 1.5$, and VI.) $\beta = 4.5$.

action of *PstI*, generated a wide range of mtDNA deletions in neuron cells [296]. Using this mouse model, age-dependent accumulation of different deletion lengths was studied. Mitochondrial DNA sequence analysis of deletion products identified that majority of clonally expanded deletion products spanned > 5kb, and further investigation using deletions of different length indicated that larger deletions had greater potential to clonally expand [296]. This observation is also consistent with other analyses considering variety of normal aging tissues from mice, rats, rhesus monkeys and human [152]. To analyze the influence of replicative advantage of mutant mtDNA on the clonal expansion dynamics in sarcopenic muscles, simulations were performed to determine mtDNA mutation accumulation dynamics with different replicative gain factor (β), assuming the basal frequency of different *de novo* mtDNA mutation to be the same.

In these trials, the hybrid stochastic model were simulated using different replicative gain factors. Complete fiber simulations indicated that at a replicative gain factor of $\beta = 1.5$, there was no fiber which either underwent atrophy or demonstrates the RRF phenotypes (Figure 5.8V). Thus, mutations with lower value of β were found to confer minimal replicative advantage, compared to the wild-type mtDNA genome (Figure 5.8IV, V). However, in the case of mutations having larger replicative gain factor ($\beta = 4.5$), the clonal expansion dynamics of such mutations was much swifter (Figure 5.8VI). In accordance with the experimental data [296], and consistent with the predictions from Chapter 4, the above analysis indicated a phenomenal propensity of mutations with higher replicative gain factor to clonally expand.

5.3.3 Perturbation analysis related to Caloric Restriction

Several research works have indicated that caloric restriction is the most effective method of extending the median and maximum life span and further have a positive effect on health in different organisms [343]. Observational studies also indicate that the CR nu-

tritional regimen have beneficial effect on human longevity [343]. Studies of Okinawa centenarians supported the idea that low-caloric food regime can increase the prospects of good health and longevity in humans [344]. The mechanism by which CR enhances the maximum life span is not completely understood. Several hypotheses have been proposed to explain the cellular mechanisms responsible for the anti-aging action of the caloric restriction, such as the growth and developmental retardation, reduced body fat content and reduced metabolic rate, delayed neuroendocrine or immunological changes, increased DNA repair capacities, reduced body temperatures [345]. However, most of these hypotheses still lack sufficient proof. Some of the parametric sensitivity analyses related to the different mechanisms, commonly proffered in explaining the beneficial effects of CR are discussed below.

Reduced mutation rate

One of the most commonly proposed hypotheses relates to the reduction of total amount of oxidative stress in tissues [345–347]. Mitochondria harvested from the brain and kidney of mice undergoing CR indicated reduced levels of hydrogen peroxide and superoxide radicals [346]. Further support to this hypothesis include the increase in transcript levels of many genes involved in ROS scavenging function (antioxidant enzymes) in CR fed animals [348, 349]. It was further seen that the mitochondrial free radical production was lowered by calorie restriction as measured from the biomarkers such as protein and lipid peroxidation products, in number of tissues such as liver, heart, brain and muscle [345]. Despite the wide prevalence of the evidence of reduced mutation burden resulting from reduced ROS production in animals subjected to CR dietary regime, a recent investigation involving rat hepatocytes subjected to CR have shown no significant changes in the cellular ROS production levels or with the expression of proteins involved in the synthesis of ROS scavengers. Similarly they found no changes in the cellular bioenergetics level of the hepatocytes subjected to CR regime [350]. Thus, the influence of reduced mutation

rate on progression of sarcopenia was evaluated by simulating lower mtDNA mutation rate (k_m). Result of the complete fiber simulations in the post-natal skeletal muscle tissue is presented in [Figure 5.9I](#). Comparison of [Figure 5.9I](#) and [Figure 5.9II](#) indicates that the post-natal reduction in the mutation rate caused due to the CR intervention, might not be a major contributor for the reduction in mutation burden in the sarcopenic skeletal muscle fibers, since most of the damage is already occurred during developmental cell lineage. The mutation burden results indicated by the total number of fibers atrophied and the number of fiber having RRF phenotypic regions are of similar magnitude compared to the normal simulations.

Enhanced mtDNA biogenesis

Another commonly proposed mechanism by which CR is able to extend the lifespan of various species, has been linked to increased mitochondrial biogenesis [[351–353](#)]. Resveratrol, a sirtuin activator, has also been proposed to result in an increase of mitochondrial biogenesis and thus result in the extension of life span. Analysis of human tissue subjected to caloric restriction indicated an elevation in the gene expression involved in mitochondrial biogenesis. Similarly, eNOS gene expression, which enhances the mitochondrial biogenesis in mammalian tissues, was also found to be significantly elevated in different mouse tissues [[353](#)]. A rather intriguing question that arises from these observations is; why would cells synthesize larger quanta of mitochondrial metabolites when the nutrient availability is relatively sparse? In a recent re-evaluation study on the effects of caloric restriction on the mitochondrial biogenesis in 30% calorie restricted rats, indicated that none of the proteins/mRNAs involved with mitochondrial biogenesis showed any elevation in expression level, in several rat tissues like heart, brain, liver and adipose tissue [[354](#)].

There is also an indication in yeast and mammal cells, that increased mitochondrial biogenesis might simply be a mechanism to compensate for damaged mitochondrial genome [[355](#)]. Further, there is also evidence that uncontrolled mitochondrial biogenesis, for exam-

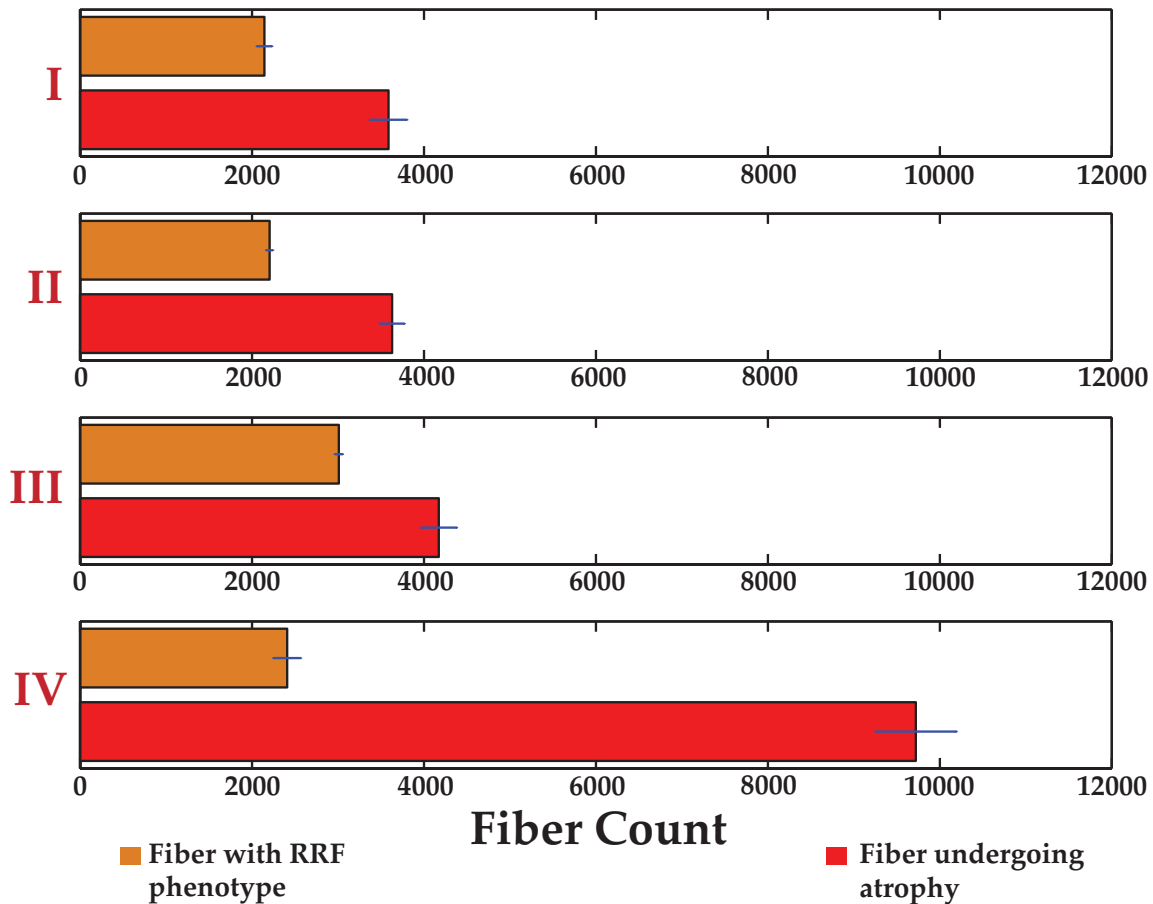


Figure 5.9: Results of different parametric analysis related to beneficial effects of caloric restriction (CR) (simulation statistics is based on five mouse tissue ($n = 5$)). The effects of different commonly hypothesized mtDNA processes that are set-in due to the induction of caloric restriction are tested and the result of the statistics related to COX deficiency in skeletal muscle fiber is compared. The figure illustrates the statistics of the number of fibers having COX deficient phenotypic regions (orange bar), and the total number of fibers that undergo atrophy (red bars) at the end of 1140 days. I.) Simulation results of the present work. II.) Simulation results of the hypothesis relating reduced mutation burden arising from the lower levels of mutation rate in caloric restricted tissues. In this trial, the basal mutation rate (k_m) for the post-natal mtDNA turnover process was set $10\times$ lower than the value used in the present work (Table 4.1). III. Simulation results of the hypothesis relating the lower mutation burden in tissue arising from the enhanced mitogenesis of mtDNA. In these simulations, the basal value of maximum post-natal mtDNA replication rate (ν_{max}) was chosen to be $1.5\times$ the value of ν_{max} used in the present modeling framework (Table 4.1). IV. Simulation results of the hypothesis relating the lower mutation burden in tissue arising from the enhanced mtDNA turnover process. In these simulations, the basal value of the mtDNA turnover rate of $k_d = 10$ days [218] was used.

ple through cardiac-specific over-expression of PGC-1 α in transgenic mice, resulted in increased mutation burden and subsequent cardiomyopathy and death [356] or can induce senescence in human fibroblast [357]. Overall, it is clear that mitochondrial biogenesis plays an important role in the aging process, with the extent of damage in the pre-existing mitochondria likely to be an important factor. Proliferation of mitochondria from dam-

aged precursors is likely to result in aggravation of damaged mitochondrial population in the cells. Similarly, increase in mitochondrial biogenesis will probably result in elevated mutation burden as *de novo* mtDNA mutagenesis is tied to mtDNA replication.

In order to evaluate if the enhanced mitochondrial biogenesis affects the mutation burden observed in the sarcopenic skeletal tissue, the post-natal complete fiber simulation were performed using higher mtDNA replication rate (ν_{max}). Results indicate that the total number of fibers undergoing atrophy is about 16% higher compared to normal simulations (Figure 5.9I and Figure 5.9III). Consistent with the Hancock *et al.*'s work, the simulation results indicated that the enhanced mitogenesis was not beneficial for reducing mtDNA mutation burden in skeletal muscle tissue. In fact, the simulation results predicted the reverse scenario, that the enhanced mitogenesis caused an increase in the mtDNA mutation burden. Thus, the enhanced mitochondrial biogenesis also might not be useful in explaining the beneficial effects of CR.

Enhanced mtDNA turnover

On similar accounts, increased mitochondrial turnover process has been hypothesized to lower the mtDNA mutation burden in post-mitotic tissues like heart and skeletal muscles [273]. The post-natal complete fiber simulations at higher mtDNA turnover rate [218], is presented in Figure 5.9IV. Similar to the enhanced mitochondrial biogenesis hypothesis, simulation results at increased mitochondrial turnover indicated a 171% increase in total number of muscle fibers undergoing atrophy, compared to the normal simulations (Figure 5.9I). The simulation results further indicated that the enhanced mitochondrial turnover might also not be a valid hypothesis for explaining the beneficial effects of CR, in the absence of any preferential degradation of mutated mtDNA.

Enhanced stress resistance

In caloric restriction regime, the availability of metabolic resources are limited and it has been observed that the animals subjected to CR regimen appeared to reset their energy utilization machinery for higher efficiency and channel these resources for more critical cellular processes [358]. It has further been observed that the animals subjected to CR undergoes certain allometric adjustments like: reduced growth, retarded reproduction, reduced thermal tolerance, reduced proliferative activity and delayed wound healing [359]. CR dietary regime enhances the ability of cells to be more stress tolerant [359, 360]. The anti-aging effect of dietary restriction provides a "nutritional stress", which results in stimulatory metabolic response for organism's survivability.

Based on these observations, in this work it was hypothesized that the somatic cells under caloric restrained condition becomes more tolerant to higher mitochondrial DNA mutation load. In other words, under CR regime, the phenotypic threshold limits associated with the mtDNA mutation burden is enhanced and correspondingly, the mitochondrial biogenesis resulting from the nuclear retrograde response is reduced. In order to simulate this condition, the Hill constant (K_H), which modulates the set point of the nuclear retrograde signaling in response to the mutation burden was adjusted to tolerate a higher mutation burden. A closer look at the Equation 4.4 reveal that this adjustment causes the effective mtDNA replication rate to drop and correspondingly increase the effective mtDNA autophagy rate, which is consistent with the experimental observations [361].

Post-natal complete fiber simulation results indicated that both the total number of fiber lost due to the fiber atrophy and the number of RRF phenotypic regions in 2 mm section of the complete muscle fiber, were in excellent agreement with the experimental data (Figure 5.10). Thus, we propose that the nutritional stress induced by CR dietary regime, causing retarded mitochondrial proliferation, to be an important process involved in the reduction of cellular mtDNA mutation burden and the consequential beneficial effects of

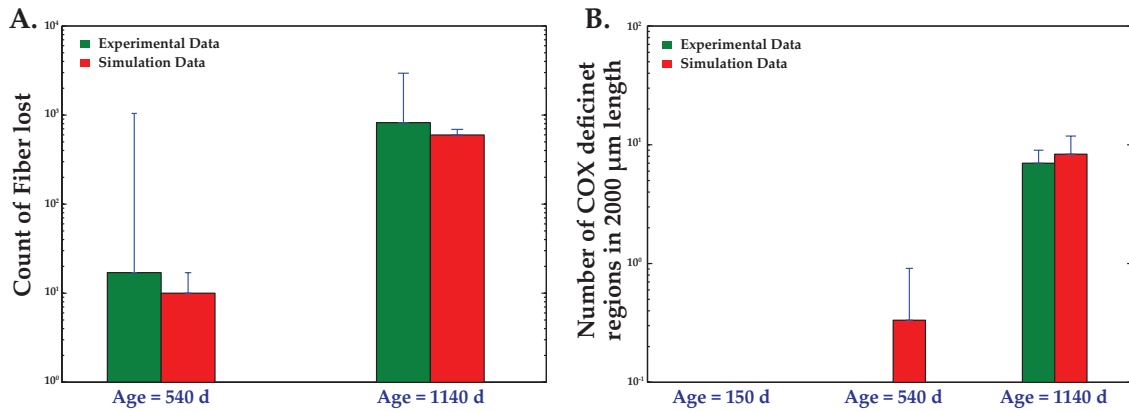


Figure 5.10: **Simulation and validation results of muscle fiber subjected to caloric restriction (simulation statistics is based on five mouse tissue (n = 5)).** The validation is done with experimental results obtained from the histochemical studies on rat skeletal muscle tissue [217, 325]. These simulation were performed with a constant of Hill-type kinetics (K_H) chosen to correspond to 93% of mutant burden, compared to 85% mutant burden in the case of normal simulations (Table 4.1). A.) Comparison of the number of fibers lost as a function of age, with the experimental data [325]. As shown in the figure, the analysis was carried out for two different age length. B.) Comparison of the simulation results of the number of COX deficient region in a 2 mm fiber section of the complete muscle tissue with experimental data [217]. As shown in the figure, the analysis was carried out for three different age length.

CR dietary regime in the life span extension.

Effects of partial CR intervention

Further simulations were performed to understand the effects of partial CR intervention on the progression and severity of sarcopenia in skeletal muscles. The duration and level of CR have important effect on the anti-aging benefits [362]. In rodents, CR regime was found to be highly effective when initiated in young ages and progressively deteriorated with age [362]. In this experimental trial, the effect of feeding regimen transfer between the *ad libitum* (AL) and CR groups were studied. Experimental data indicated that switching of old mice under AL (26 mo) to a CR regimen (AL→CR) tended to decrease their probability of survival relative to AL maintained cohorts. Similarly, transfer of CR mice to the AL regimen (CR→AL) had little or no effect on the survival when compared to the CR regimen cohorts. Thus, these trials indicated that the life span extension by CR was neither quickly reversible nor inducible, but instead required a long-term reduction in calorie intake.

Simulation results of such dietary regimen transfer are illustrated in Figure 5.11. Simi-

lar to the experimental data, the post-natal complete fiber simulation results also indicated that a short term dietary regimen transfer did not cause any significant change in the total number of fibers atrophied (which could be considered as an important indicator of tissue dysfunction). However, it is interesting to note that there was significant difference in the number of fibers having RRF phenotypic regions between different dietary regimen transfers. Thus, similar to the experimental observations, simulation data also indicated that the beneficial effect of CR accrues gradually and that they are not rapidly inducible/reversible upon dietary regimen transfer.

5.4 Conclusions

Sarcopenia refers to an age related condition in which decline in skeletal muscle mass and strength occurs with age and causes frailty in aged organisms. Due to the increasing aging population, and the associated increase in the health care expenses related to sarcopenia, the need for better understanding of the molecular etiology and progression of sarcopenia is warranted. Multiple molecular hypotheses for the etiology of sarcopenia have been proposed, and one of the mechanisms that has recently gained much interest is the mitochondrial etiology of sarcopenia, which proposed that the age-associated focal accumulation of mitochondrial genome abnormality is the main cause of skeletal muscle sarcopenia. But, the molecular mechanisms of the origin and accumulation of mtDNA mutations are not completely understood. Similarly, limited knowledge exists on the possible genetic/pharmacological interventions to reduce or retard the progression of sarcopenia. In this work, we have developed a hybrid multiscale model of mitochondrial DNA turnover process in rat skeletal muscle fibers to understand the role of stochastic origin and dynamics of pathogenic mtDNA mutations in skeletal muscle sarcopenia. Specifically, the model simulates: a.) the origin and expansion dynamics of mtDNA mutations in single cell muscle progenitor cells (myoblasts) during the animal development, using a Chemical

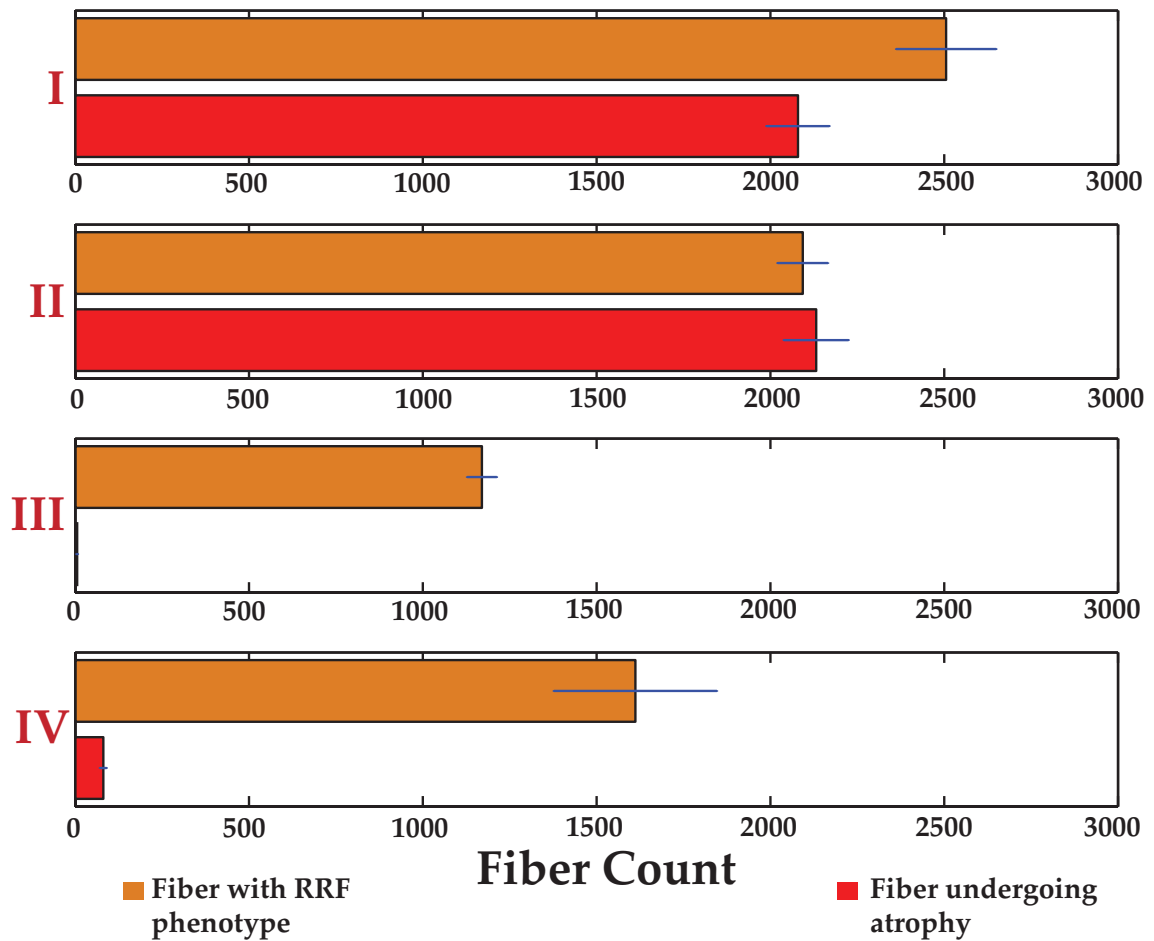


Figure 5.11: Results of tissue mutant burden arising from different short-term dietary regimen transfer (simulation statistics is based on five mouse tissue($n = 5$)). The effects of different dietary regimen transfers were simulated and the results of different statistics related to COX deficiency in skeletal muscle fiber was compared. The dietary regimen transfer was carried out such that the simulations were carried at the prevailing dietary condition for an age length of 26 months and subsequently the dietary regime was transferred for a short span of 11 weeks. The figure illustrates the statistics of the number of fibers having COX deficient phenotypic regions (orange bar), and the total number of fibers that undergo atrophy (red bars) at the end of 1140 days. I.) Simulation results related to the mutant burden in skeletal tissue of rats subjected to only *ad libitum* dietary condition for the entire simulation period. II.) Simulation results related to the mutant burden in skeletal tissue of rats subjected to *ad libitum* dietary condition for the initial period of 26 months and subsequently subjected to caloric restriction for time span of 11 weeks. III. Simulation results related to the mutant burden in skeletal tissue of rats subjected to caloric restriction dietary regime for the entire simulation period. IV. Simulation results related to the mutant burden in skeletal tissue of rats subjected to caloric restriction dietary regime for the initial period of 26 months and subsequently subjected to AL dietary condition for time span of 11 weeks.

Master Equation (CME) model and, b.) the expansion of mtDNA mutations in post-mitotic skeletal muscle fiber using deterministic reaction-diffusion formulation.

In rat muscle fiber simulations, different statistics, e.g., the average count of fibers lost during murine lifespan, the total number of red-ragged phenotypic regions and the average length of the spread of the COX deficient regions in a given section of muscle fiber, are

estimated and validated against the experimental data on histo-chemical staining and the related *in situ* hybridization analysis. The simulation results indicated an excellent agreement with the experimental data. Consistent with works in [Chapter 3](#) and [4](#), the simulation results of skeletal muscle tissue also indicated that the mosaicity observed in cellular mtDNA mutation burden in post-mitotic tissue arises due to stochasticity in mitochondrial turnover process, predominantly during developmental cell lineage. The myoblast cells having significant mutation burden at birth, form the seed of COX deficient regions in a sarcopenic muscle fiber. The muscle progenitors having lower mutation burden at birth or the skeletal muscle fiber acquiring *de novo* mutations during the early post-natal stages rarely expand to form COX deficient regions in skeletal muscle fiber.

Strenuous physical activity like the different exercise regimes have been shown to significantly affect the mitochondrial motility and the associated mtDNA nucleoid dynamics. Simulation results indicate that the faster diffusion of mtDNA in the skeletal muscle fiber is beneficial, as it enhances the spread of mtDNA mutation load along the fiber length and thus reducing the overall local mutation burden in the fiber. Further, this resulted in reduction of nuclear retrograde response, which modulates mitogenesis in response to the cellular respiratory demands. The enhanced mtDNA diffusivity manifested by the exercise regime is thus beneficial in reducing the mutation burden associated with sarcopenic muscle fibers.

A different parametric perturbation analysis encompassing some of the commonly proposed mechanisms to explain the beneficial effects of CR, was performed using skeletal muscle model. Different hypotheses such as reduced mutation rate, enhanced mitochondrial biogenesis/turnover were tested. In this work, it was further hypothesized that the beneficial effects of caloric restriction dietary regime, results from the enhanced stress resistance of the post-mitotic cells to withstand higher mtDNA mutation burden. Simulation results indicated that only the stress resistance hypothesis produced an excellent agree-

ment with the histochemical analysis data associated with the rat skeletal tissue. The duration of CR has been found to have important implication on the anti-aging effect of CR. Consistent with the experimental evidence, simulation results indicated that a short-term feeding regimen transfer does not lead to any significant change in the mutation burden in skeletal muscle tissue. Further, the results indicated that the beneficial effects of CR accrue gradually and it is neither rapidly inducible nor reversible upon the dietary regime transfer.

The present model can be further developed as a work-bench to analyze the influence of different parametric perturbations associated with the mitochondrial DNA turnover process, on the origin and progression of sarcopenia with age, thereby providing important insight into the possible interventions to retard its severity and possibly provide a hypothesis-driven design of experiments to gain a better understanding of the role of mtDNA mutations in aging or mitochondrial related degenerative diseases.

Maximizing signal-to-noise ratio in the Random Mutation Capture assay

6.1 Introduction

As discussed in the previous chapter ([Chapter 3](#)), an accurate estimation of tissue level DNA mutation abundance is challenging, but critical for understanding the impact of these mutations on cellular and tissue physiology [[141](#), [363](#), [364](#)]. The ability to sensitively detect the frequency of random spontaneous mutations of DNA is necessary to define critical parameters like the mutation rate, which is essential for understanding the role of genomic mutagenesis in human diseases and aging. In [Chapter 3](#), it was shown that the quantification assays used for measuring low-levels of mtDNA mutation burden requires high end point dilution to single molecular content. Such high dilution factors makes the measurement estimates obtained from these protocols highly noisy and thus causing significant uncertainty in predicting the age-dependent mtDNA mutation accumulation dynamics. In this chapter, statistical modeling techniques are used to investigate the impact of high dilution in RMC assay [[116](#), [141](#)] and to propose an alternate, statistically-optimal assay sampling procedure to reduce the extent of measurement variability and to increase the amount of information returned per sampling trial.

Many procedures have been developed previously based on quantification of point mutation frequency by PCR amplification of short fragments of DNA. However these methods have recently come under criticism, since mis-incorporation rate of spurious mutations during PCR, arising due to the intrinsic error rate of polymerase enzymes, can exceed the actual mutation frequencies, especially when measuring low-level DNA mu-

tation level ($\sim 10^{-6}$ bp⁻¹). Upon cloning these mutations become indistinguishable from the true mutations and may provide an overestimation of the mutation load.

In 2005, Bielas and Loeb developed a novel assay for quantifying low levels of spontaneous random mutations in DNA [141]. The RMC assay is based on PCR detection of mutations that render the mutation target sequence resistant to restriction enzyme digestion [141, 363]. Mutations are subsequently quantified by real-time PCR amplification after dilution of the restriction enzyme digested template to single-molecule level [226], avoiding artifactual mutations due to polymerase error that has confounded previous methods [226, 363]. Application of the RMC assay in determining the mutation frequency of mouse mitochondrial DNA (mtDNA) [226] gave values about two orders of magnitude lower than the previous estimates [130, 131].

In the RMC assay, mutation burdens are quantified by real-time PCR after dilution to single-molecule level (Figure 6.1), such that each PCR-amplified well contains mostly single mutant molecule and the likelihood of finding more than a single mutant in a well is negligible [116, 141, 226]. While at the population level, low sample size has been shown to negatively affect the conclusion drawn from DNA mutation studies [365], in this chapter it is shown that the statistical aspects related to protocols necessitating such high dilution (down to single molecule level) can also introduce significant noise in the mutation frequency estimates. Unfortunately, the statistics associated with the discrete molecular nature of such assays can become non-trivial and was overlooked in the original protocol and subsequent applications.

6.2 Methods

6.2.1 Statistical optimization of the RMC assay

The conventional RMC assay [116, 141, 226] is based on PCR amplification of a single mutant molecule (restriction enzyme digest resistant) in the presence of excess copy numbers

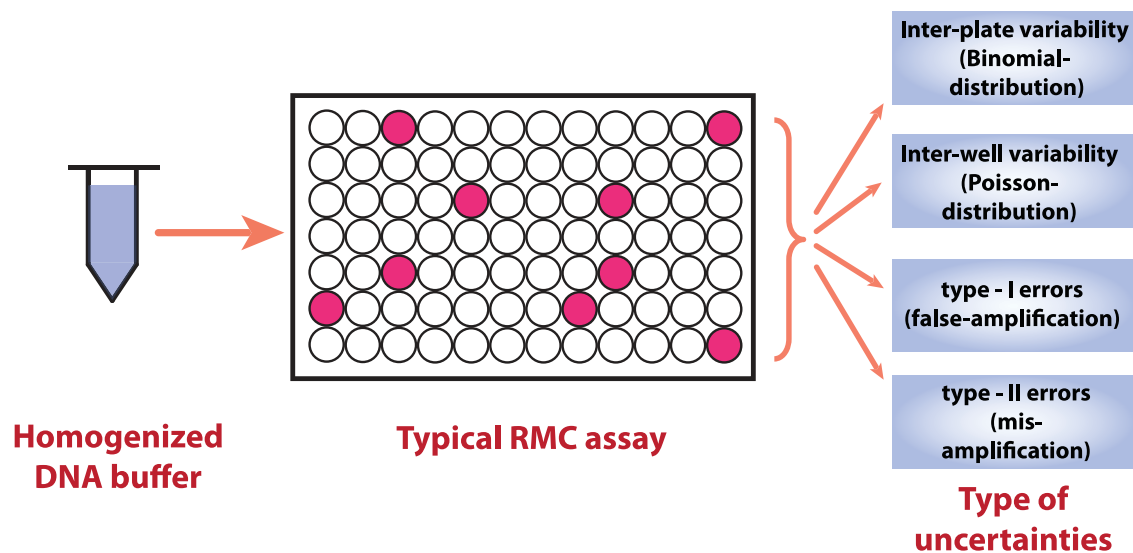


Figure 6.1: **Statistical aspects associated with the RMC assay.** Different sources of variability in the RMC assay. There are primarily four sources of variability that are important for the sensitive estimation of DNA mutation load using the RMC assay: the number of PCR (un)amplified wells among independent RMC trials (inter-plate variability), assuming that each PCR well receives same number of mtDNA templates, the molecular count of mutant DNA in different wells of a single plate (inter-well variability), i.e., randomness in mtDNA copy number during the sampling of mtDNA samples into PCR wells, false amplification (type-I error) and mis-amplification (type-II errors).

of wild-type sequences (digest sensitive). Here, the mutation frequency is calculated by dividing the fraction of PCR wells that are amplified (in a 96-well plate) by the amount of DNA per well [116, 141, 226]. In the RMC assay, intrinsic variability arises due to the well-to-well and plate-to-plate variability that can arise from non-uniform sampling of DNA content to a single PCR well (Poisson distribution) and random number of amplified/non-amplified wells in a plate (Binomial distribution), respectively (Figure 6.1). The other sources of error (bias) in this protocol are associated with (1) the assumption that each amplified well only contains a single mutant molecule, (2) false PCR amplification due to contamination or incomplete digestion (Type I), and (3) failed PCR amplification (Type II). Each of these factors contributes in a non-trivial manner to the statistics of the mutation frequency data, and has not been carefully addressed previously.

To arrive at the optimal protocol, two types of statistical analyses were performed. The first involved a linearized variance propagation of functions of random variables, while the second used a Monte Carlo approach that simulated the protocol a large number of

times ($n = 10,000$). The optimization was then performed to minimize the coefficient of variation, by changing the amount of DNA molecules that are sampled into each well. The details of each analysis are outlined below.

First order uncertainty propagation analysis

Assuming that the DNA homogenate for which mutation frequency is to be determined is well mixed, the probability of sampling x number of mutant molecules into a PCR well can be described using the Poisson distribution [233]:

$$P_{Poisson}(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!} \quad (6.1)$$

where λ is the mean number of mutant molecules per well. The parameter λ is of interest in this assay as the mutation frequency (per bp) can be calculated by dividing λ with the total amount of DNA present in the well. Using the Poisson distribution, the probability of sampling zero mutant template into an PCR well, denoted by p_0 , is given by:

$$p_0 = P_{Poisson}(x = 0) = e^{-\lambda} \quad (6.2)$$

Furthermore, assuming that p_0 is constant and ignoring Type I and II errors as described above, the probability of observing n_0 wells that are not PCR-amplified among a total of n_{wells} wells is described by a Binomial distribution:

$$P(n_0) = \binom{n_{wells}}{n_0} \cdot p_0^{n_0} \cdot (1 - p_0)^{n_{wells} - n_0} \quad (6.3)$$

While the expected value of n_0 is equal to $n_{wells} \cdot p_0$, the outcome in a random experiment may not be so. Thus, one can only obtain an estimate of p_0 , which is calculated as \hat{p}_0 . Consequently, an estimate of λ can be calculated from \hat{p}_0 by rearranging Equation 6.2, such that:

$$\hat{\lambda} = -\log \hat{p}_0 \quad (6.4)$$

Since \hat{p}_0 is an unbiased estimator of p_0 (i.e. the expected value of \hat{p}_0 is equal to p_0), the variance of $\hat{\lambda}$ can be approximated by using a Taylor series expansion of Equation 6.4. The uncertainty analysis of $\hat{\lambda}$ relies on a first order linear approximation, in which the variance of any function of random variables, $y = g(x)$ is approximated from a Taylor Series Expansion (TSE) about the estimator for the mean of the random variables \mathbf{x} , given by:

$$V(y) \approx \sum_{\mathbf{x}} \left(\left. \frac{\partial g}{\partial x} \right|_{x=\hat{x}} \right)^2 \cdot V(x) \quad (6.5)$$

where \hat{x} is an unbiased estimator of the mean of x . Since \hat{p}_0 is an unbiased estimate for p_0 , the uncertainty in the estimate $\hat{\lambda}$ or $V(\hat{\lambda})$ can be obtained as a function of the uncertainty in \hat{p}_0 according to:

$$V(\hat{\lambda}) \approx \left(\left. \frac{\partial (-\log p_0)}{\partial p_0} \right|_{p_0=\hat{p}_0} \right)^2 \cdot \frac{\hat{p}_0 \cdot (1 - \hat{p}_0)}{n_{wells}} \quad (6.6a)$$

$$V(\hat{\lambda}) \approx \left(\frac{1}{\hat{p}_0} \right)^2 \cdot \frac{\hat{p}_0 \cdot (1 - \hat{p}_0)}{n_{wells}} \quad (6.6b)$$

$$V(\hat{\lambda}) \approx \frac{1 - \hat{p}_0}{n_{wells} \cdot \hat{p}_0} \quad (6.6c)$$

Consequently, the coefficient of variation (CV) of $\hat{\lambda}$ is approximately:

$$CV(\hat{\lambda}) \approx \frac{1}{-\log(\hat{p}_0)} \cdot \sqrt{\frac{1 - \hat{p}_0}{n_{wells} \cdot \hat{p}_0}} \quad (6.7)$$

The variation of the above CV as a function of different DNA template dilution (λ) was approximated using the above Equation 6.7 and the minimum CV was obtained at $\hat{p}_0^* \approx 0.2$.

Monte Carlo analysis

The analytical treatment above did not consider false positives and false negatives (Type-I or Type-II errors) as described in Figure 6.1. To better understand the influence of these

errors on the mutation frequency estimates, Monte Carlo (MC) simulations of the conventional RMC assay and the proposed alternative protocol (i.e. counting non-amplified PCR wells) with and without Type-I and -II errors, were also performed. Briefly, a single MC simulation emulates the statistical random sampling of DNA material into each PCR well on a plate and the PCR amplification experiment. The lower bounds of the frequencies of false positive and false negative errors were estimated from experimental data described in [Table 6.1](#). False amplification can be avoided by sequencing in the conventional RMC assay and thus was not taken into account in the corresponding MC simulations. Such simulation was then performed for a large number of times to obtain the statistics of the mutation frequency estimates, such as mean and CV. The pseudo-codes for the MC algorithms implemented are given in [Table 6.2](#) and [Table 6.3](#) for original and optimized RMC protocols, respectively.

The Monte Carlo algorithms were implemented in C++ on a Linux Platform (CentOS; GNU C++ compiler (v4.1.1)), using a combination of a long period random number generator [[366](#)] and a multiple independent streams generator [[367](#)]. In addition, non-uniform random numbers were generated using algorithms described elsewhere [[368](#), [369](#)].

6.2.2 Validation of RMC optimization

The performance of the proposed optimized protocol was validated by a mock RMC assay. Briefly, a plasmid containing the target sequence and a matching TaqMan probe/primer set based on the TaqMan MGB Probe with 5'-6FAM and 3'-MGBNFQ dye technology was prepared (Applied Biosystems, Calif.). The plasmids were serially diluted to approximately either 0.1 or 1.6 molecules per μl . The former corresponds to the dilution used in the single molecule amplification of the original RMC assay [[141](#)], while the latter represents the suggested dilution of the optimized assay ($\hat{\lambda}^* = -\log \hat{p}_0^* = 1.6$). A $1\mu\text{l}$ sample of either the high or the low copy number dilutions were added to a $19\mu\text{l}$ of standard Taq-

Table 6.1: Estimation of the false positives and false negatives frequencies. Frequencies of false positives and false negatives encountered during the RMC trials. Data are based on the quality control experiments of the RMC assay conducted using the no template control (NTC) and mtDNA templates. Trials used for obtaining the false positive error frequencies were conducted with PCR amplification of the wells having pure buffer (NTC's). Whereas, the trails used for determining the false negative error frequencies were based on amplification of the PCR wells with an average of 10 mtDNA templates in each of the PCR wells.

Stage	Repetitions	Temperature	Time	Ramp Rate
1	1	50°C	2:00	100
2	1	95°C	10:00	100
3	70	95°C	0:45	100
		55°C	0:45	100
		72°C	1:30	100
Standard 7500 Mode				
Data Collection: Stage 3 & Step 3				
PCR Volume: 20 μ l				
Type - I error Data				
Number of wells used for the trial				100
Number of wells amplified				6
Error frequency				6%
Type - II error Data				
Number of wells used for the trial				100
Number of wells amplified				96
Error frequency				4%

Man UDP master-mix (TaqMan Universal PCR Master Mix, Applied Biosystems, Calif.) to each well of a 96 well plate PCR plate (Applied Biosystems). The wells were subsequently subjected to PCR amplification and detected using an Applied Biosystems 7500 real-time PCR system running 60 cycles of PCR of a modified universal amplification profile (50°C for 2 min, 95°C for 10 min followed by 60 cycles of: 95°C for 45 s, 55°C for 45 s and 72°C for 90s). Amplification was monitored at 520nm (FAM Dye) and under these conditions single molecules were found to be reliably amplified (data not shown).

6.3 Results & Discussion

Following the derivation in the Method section, the optimal protocol involves (1) preparation of DNA sample such that about $1 - \hat{p}_0^* \approx 0.8$ fraction of the wells will be amplified

Table 6.2: Pseudo-code for simulating the sampling protocol used in the original RMC assay.

Steps	Algorithm details
1	Initialization
	set the number of PCR wells n_{wells} .
	set frequencies of false positive (α) and false negatives (β) errors.
	seed random number generators (both uniform and non-uniform).
	set template copy number n_{mtDNA} .
	set the iteration counter to $k \Rightarrow 0$.
2	Iterations
	while $k \leq 10000$ (I_{tot}).
	initialize the counter for amplified wells $i_A \Rightarrow 0$.
	set counter for number of wells per PCR plate $n \Rightarrow 1$.
	while $n \leq n_{wells}$.
	generate a Uniform random number $U(0, 1) : r$.
	generate a Poisson random number $P(\lambda) : R$.
	if $R \neq 0$, then:
	if $r \geq \beta$, then:
	$i_A = i_A + 1$
	else
	if $r < \alpha$, then:
	$i_A = i_A + 1$
	increment $n + 1$.
	calculate the estimate of mutation frequency $\hat{\theta} = i_A / (n_d \cdot n_{wells})$.
	increment $k + 1$.

by adjusting the dilution factor and (2) the mutation frequency estimate $\hat{\theta}$ (mutations per base pairs) can be computed according to:

$$\hat{\theta} = \frac{\hat{\lambda}}{n_{bp} \cdot n_{DNA}} \quad (6.8)$$

where n_{DNA} is the total number of DNA molecules in a single PCR well and n_{bp} is the length restriction site ($n_{bp} = 4$ for *TaqI* recognition site [116, 226]). Figure 6.2 shows that the results of statistical optimization using the linearized analysis and MC approach, and from the RMC validation experiments. In particular, Figure 6.2A and Figure 6.3 suggests that the minimum CV can be achieved by diluting DNA sample such that $1 - \hat{p}_0^*$ of the wells are PCR amplified, which corresponds to an average of $\hat{\lambda} = 1.6$ mutant molecules per well. This is in contrast to the original RMC assay that prescribes end-point dilution down to $\hat{\lambda} = 0.1$ mutant molecules per well.

Table 6.3: Pseudo-code for simulating the sampling protocol used in the optimized RMC assay, developed in this work.

Steps	Algorithm details
1	Initialization
	set the number of PCR wells n_{wells} .
	set frequencies of false positive (α) and false negatives (β) errors.
	seed random number generators (both uniform and non-uniform).
	set mean fraction of unamplified well (p_0).
	set the iteration counter to $k \Rightarrow 0$.
2	Iterations
	while $k \leq 10000$ (I_{tot}).
	set the counter for unamplified wells $i_{UA} \Rightarrow 0$.
	set counter for number of wells per PCR plate $n \Rightarrow 1$.
	while $n \leq n_{wells}$.
	generate a Uniform random number $U(0, 1) : r$.
	generate a Poisson random number $P(\lambda) : R$.
	if $R = 0$, then:
	if $r \geq \alpha$, then:
	$i_{UA} = i_{UA} + 1$
	else
	if $r < \beta$, then:
	$i_{UA} = i_{UA} + 1$
	increment $n + 1$.
	calculate the fraction of unamplified wells $\hat{p}_0 = i_{UA}/n_{wells}$.
	calculate the estimate of mutation frequency $\hat{\theta}$.
	increment $k + 1$.

Importantly, the optimized protocol offers a significant improvement of the accuracy in mutation estimates compared to the conventional RMC assay. For example, Monte Carlo (MC) simulations of RMC assays ($n = 10,000$) on 48 PCR wells ($n_{wells} = 48$) predicted a 53% reduction in coefficient of variation (CV) (from 40% to 19% CV) by using the optimal protocol in place of the original (Figure 6.2B). Furthermore, this protocol provides a more robust estimate of λ than the traditional RMC assay, where the CV is approximately constant for the values of p_0 between 0.1 and 0.4, even in the presence of experimental artifacts (false positive and false negative errors) (Figure 6.2, Figure 6.3 and Figure 6.4).

The theoretical predictions were tested experimentally by conducting a mock RMC assay using plasmid DNA (see Section 6.2.2). As expected, at the optimal dilution factor of $p_0^* = 0.2$, this protocol provides a substantial reduction of variability with an actual CV

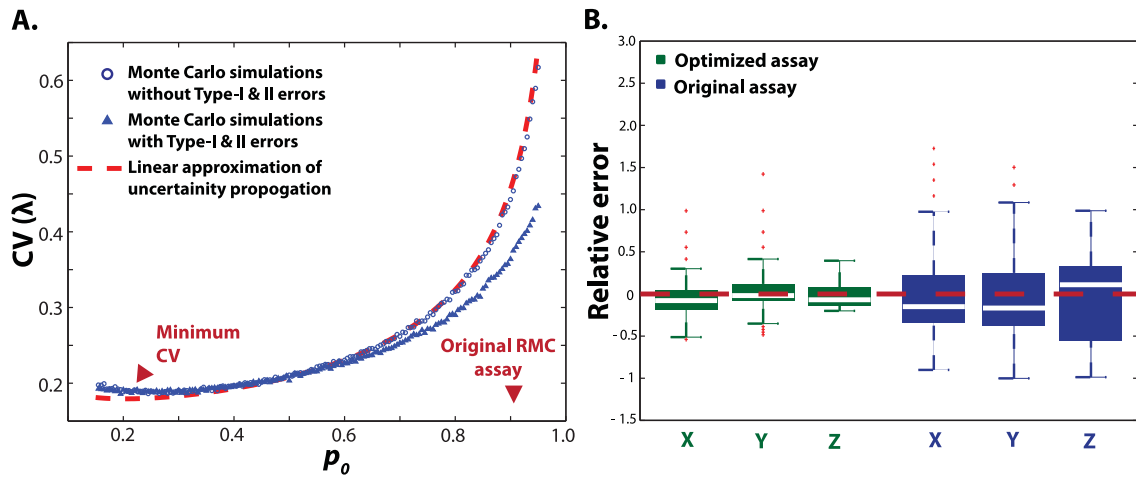


Figure 6.2: Results of Optimized RMC assay considering the approximation of sampling process to the Poisson process. A.) Coefficient of variation (CV) of DNA mutant frequency as a function of p_0 (the mean fraction of unamplified wells) as predicted by MC simulations and statistical analysis using assay with 48 wells (n_{wells}). The parameter p_0 relates to the DNA template dilution factor, where higher dilution increases p_0 (less DNA templates per PCR well). The CV has a minimum value around $p_0 = 0.20$, as opposed to the conventional RMC assay of $p_0 \approx 0.91$. The lower CV in simulations with type-I and type-II errors comes at the cost of lower accuracy. B. Comparison of the relative errors between optimized and conventional RMC assays in (i and ii) simulation and (iii) experiment; ($n_{wells} = 48$; Table 6.4). MC simulations ($n = 10,000$ realizations) were performed (i) with and (ii) without type-I and type-II errors. The type-I and type-II errors were determined based mock RMC repeats ($n = 100$ PCR wells, Table 6.1). Type-I error in the original RMC assay is set to zero, since existence of mutant DNA in amplified wells is confirmed by sequencing. Both MC simulations and experimental data confirmed the superiority of the optimized protocol in terms of CV reduction and accuracy improvement.

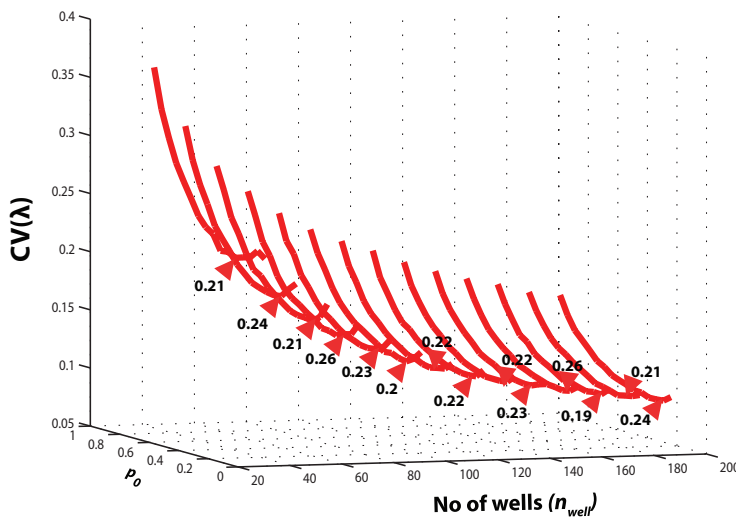


Figure 6.3: Variation of the Coefficient of variation (CV) obtained using Monte Carlo analysis. Data obtained using Poisson statistics as a function of the fractions of unamplified wells (p_0) and the total number of PCR wells (n_{wells}). These contours are determined using Monte Carlo simulations. The arrows indicate the minimum value of CVs. The optimal value of p_0 mostly ranged between 0.2 and 0.3.

reduction of 68% (from 57% to 18% CV) (Figure 6.2B and Table 6.4). In other words, the optimal RMC protocol offers the same information as the original protocol using only 1/10 the number of wells. These findings highlighted that in the quantification of low-level DNA mutations, the manner in which samples are prepared for analysis can contribute non-trivially to the statistics of the measurement data.

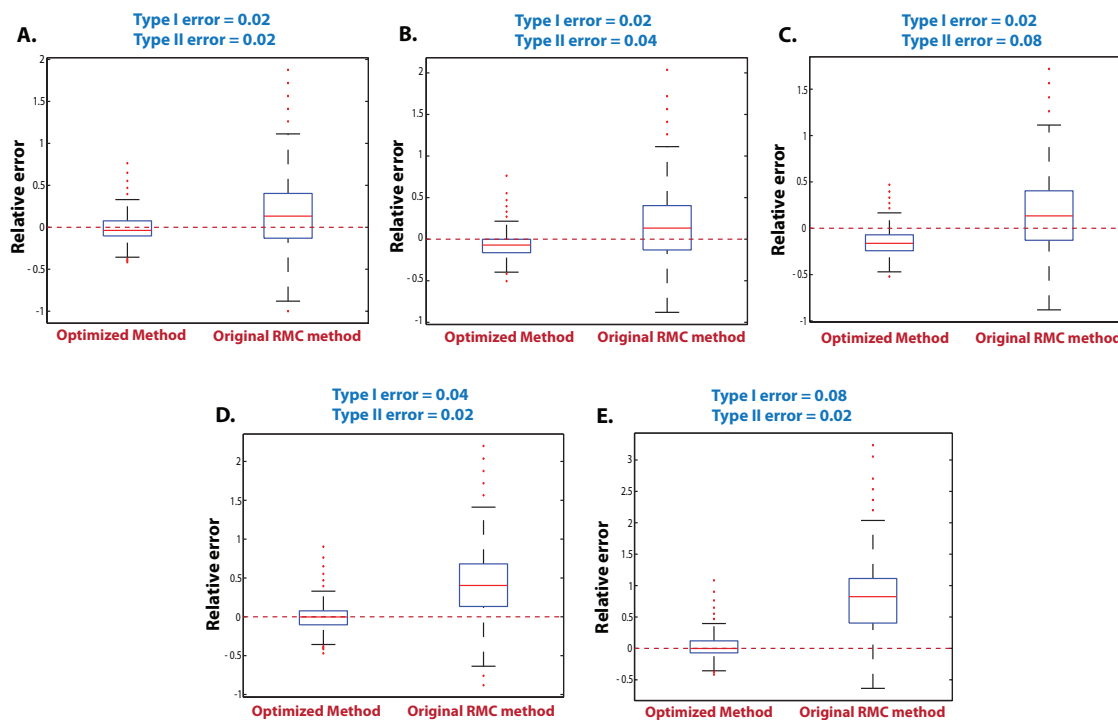


Figure 6.4: Influence of the experimental errors (type - I and II errors) on the performance of both the optimized and original RMC protocol assuming Poisson random sampling. The box plots of the relative difference between $\hat{\lambda}$ and the true value λ were obtained from 10^4 MC independent realizations with a true $\lambda = 1.6$ molecules per well ($p_0 = 0.2$) for the optimized assay and $\lambda = 0.1$ molecules per well for the conventional RMC assay ($p_0 = 0.91$). The frequencies of type-I and II errors are shown in the titles and their selection were done using data shown in Table 6.1. The original RMC protocol is generally more sensitive to these experimental errors than the optimized assay developed in this work.

Table 6.4: Mock RMC experimental results using the optimized and original protocols.

Optimized Protocol $\lambda = 1.6$			Original Protocol $\lambda = 0.1$		
Run number	Fraction of amplified wells $1 - \hat{p}_0$	Average mutant molecules per well $\lambda = -\log(\hat{p}_0)$	Run number	Fraction of amplified wells $1 - \hat{p}_0$	Average mutant molecules per well $1 - \hat{p}_0$
Run 1	36/48	1.39	Run 1	1/48	0.02
Run 2	39/48	1.67	Run 2	6/48	0.13
Run 3	38/48	1.57	Run 3	10/48	0.23
Run 4	35/48	1.31	Run 4	7/48	0.16
Run 5	36/48	1.39	Run 5	6/48	0.13
Run 6	42/48	2.08	Run 6	3/48	0.06
Average (molecule per well)		1.5668	Average (molecule per well)		0.1146
Standard Deviation (molecule per well)		0.2854	Standard Deviation (molecule per well)		0.0565
CV		0.1822	CV		0.5721

Thus, despite the high sensitivity of the RMC assay, the high variability due to the discrete molecular nature of the original assay can lead to a significant uncertainty in measurement estimates. Nevertheless, the subsequent applications of this assay (Appendix A), many of which were concerned with age-dependent mutations in mitochondrial DNA, have predominantly used the naïve form. To illustrate the significance of the optimal protocol, based on the earlier work on the mtDNA point mutation accumulation in murine heart tissue (Chapter 3 or [269]), Monte Carlo simulations of low-level accumulation of mtDNA point mutations during aging and their measurements using the RMC assay with the original and the optimal protocol were performed. Figure 6.5A shows the intrinsic variability in the levels of mtDNA point mutations among heart tissues from a population of mice that may arise from random drift of mtDNA point mutations during the developmental and post-natal stages. When using the original RMC assay, the additional variability (noise) from suboptimal sampling of the original RMC assay, while able to explain the range of experimental data, prevents any definitive conclusion regarding the dynamics of mtDNA mutation accumulation, as shown Figure 6.5B. On the other hand, the reduction in the sampling variability offered by the optimal protocol can provide a much more definitive trend on dynamics (see Figure 6.5C, variability is much more closer to the actual underlying mtDNA point mutation distribution function), which is important in deducing mechanisms of age-related mtDNA mutations.

6.4 Conclusions

Methods for quantifying extremely low levels of mutation frequency ($\sim 10^{-6}$ bp⁻¹), like RMC assay [141], represent an important technological advance, the findings above reveal that the statistics associated with the sample preparation in these assays are of great importance, but often an overlooked aspect. Consequently, sampling protocols need to be carefully designed to maximize signal-to-noise ratio. A sensitive and more robust esti-

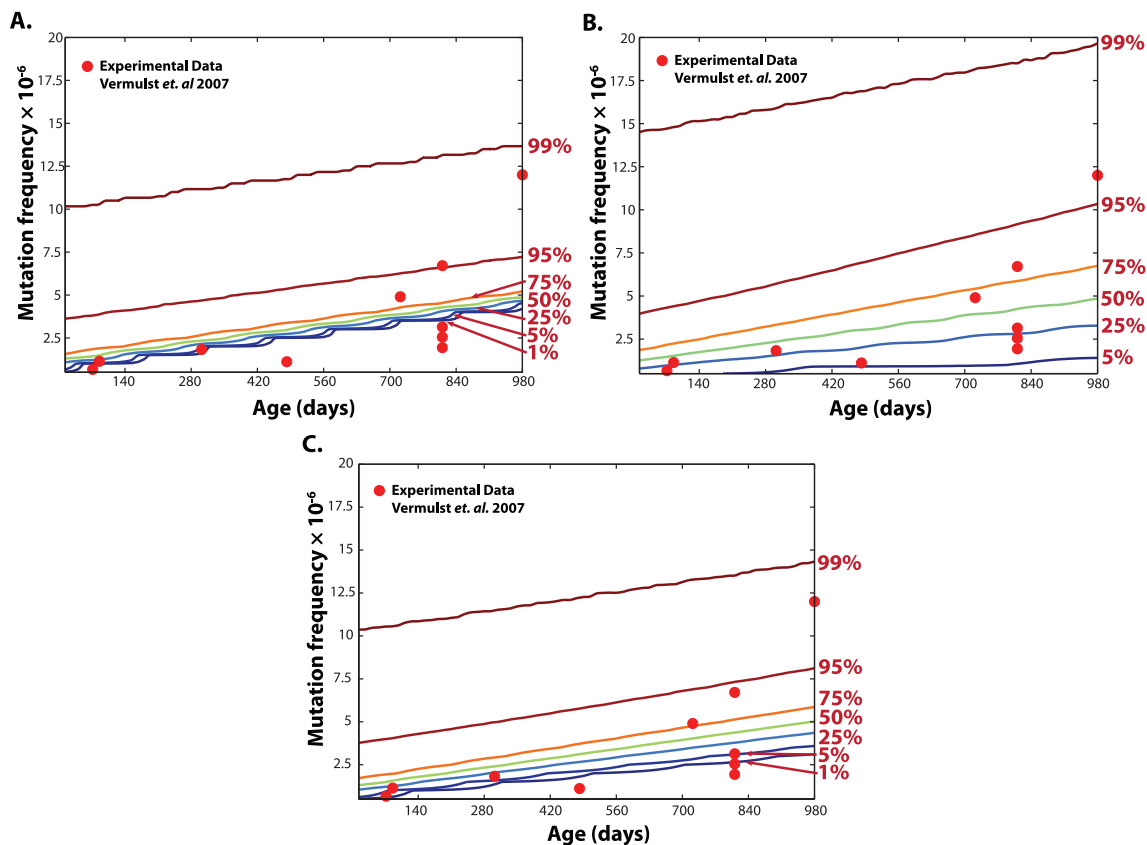


Figure 6.5: **Variability in the point mutation frequency measurement data in mouse heart tissues.** Point mutation burden in wild type mouse population ($n = 1000$ mice) as predicted by a random drift model Chapter 3; [269]. Actual experimental RMC data are shown in red circles [116]. The simulated percentile curves in the plots give the maximum mutation frequency that a given percentage of the mouse population harbors. The percentiles of point mutation frequency distribution in mouse heart tissues as a result of A.) random drift of mtDNA mutation; B.) random drift and sampling variability from the original RMC assay; and C.) random drift and sampling variability from the optimized RMC protocol. The comparison of overall data variability indicates that the optimized protocol developed in this work, provides a substantial reduction in the measurement variability and provides a better estimate of the underlying age-dependent mtDNA mutation accumulation dynamics.

mate of mutation frequency can be obtained by a simple modification of the original assays [116, 141, 226] based on statistical analysis. While the original procedure prescribes a dilution factor of about 0.1 molecules of mutant per well and the counting of *amplified wells*, the maximum signal-to-noise ratio however corresponds to a lesser dilution factor of 1.6 molecules per well and quantifying the fraction of *unamplified wells*. The optimized assay is predicted and confirmed by experimental data, to substantially reduce the measurement variability, and thus provide a more sensible estimate of the age-dependent mtDNA mutation dynamics. While this work pertains to optimization of the RMC assay, the statistics related to sample preparation and sampling protocol maybe more generally applicable to

other DNA quantification techniques requiring high dilution to obtain single molecule end point dilutions.

This chapter highlights the work done in S. K. Poovathingal, J. Gruber, N. Li Fang, B. Halliwell & R. Gunawan. Maximizing Signal-to-Noise Ratio in the Random Mutation Capture Assay. *Nuc. Acids. Res.*, Accepted.

Global parameter estimation methods for stochastic biochemical systems

7.1 Introduction

One of the most critical challenges experienced during the development of different stochastic modeling framework proposed in this dissertation ([Chapter 3](#) and [4](#)), relates to the availability of accurate estimates of critical parameters of mitochondrial turnover parameters like the mtDNA turnover rate, *de novo* mtDNA mutation rate, etc. Obtaining accurate estimates of several of these parameters are technically challenging using the existing measurement technologies. Also, obtaining parameters estimates using different measuring techniques can give vastly diverging values (for example mtDNA turnover rate discussed in [Chapter 3](#)). Furthermore, unlike differential equation models, the identification of stochastic models from experimental data of single cell or cell population data are not yet routine. Thus in this chapter, details of a stochastic parameter estimation framework that has been developed to estimate parameters of stochastic biochemical reaction systems will be discussed.

Despite the availability of high-throughput cell biology, the estimation of unknown (kinetic) model parameters from experimental data is still considered as the bottleneck in biological model identification, especially for dynamical models [[370](#), [371](#)]. The difficulty is generally attributed to the informativeness of the data, or the lack thereof, a property that is proportional to not only the quantity, but also the quality of data. Furthermore, in dynamical models, the time resolution of data is naturally of great importance. In recent years, advances in bio-imaging allow for real time measurements of cellular components

such as mRNAs and proteins in individual cells through the use of fluorescent proteins [172, 372–375]. Such measurements provide more in-depth and informative data about the states of a cell and variability in a cell population, than the traditional lumped measurements from cell culture lysate or tissue homogenate. The purpose of this chapter is to present the development of practical methods that can efficiently use these data in the parameter estimation framework for stochastic biochemical systems.

Chemical master equation (CME) is the most commonly adopted modeling framework to describe stochastic cellular dynamics [172, 178, 374] and thus is used as a benchmark application in this work. The estimation of unknown kinetic parameters from data in CME and other stochastic models has not been adequately addressed in the literature. Many of the published CME models use rate constants that are scaled from deterministic parameter values or selected *ad-hoc* to replicate desired behavior. Since the low-copy-number random events can generate dynamics that are characteristically different from those in thermodynamic or deterministic limit [376, 377], deterministic model parameters identified from data collected under this limit or averaged over cell populations can be misleading. Furthermore, fitting deterministic models (e.g. ODE) to stochastic data has been shown to give poor parameter estimates and model prediction [378]. Among the existing parameter estimation methods for stochastic biological models, some rely on Bayesian inference based on the stochastic differential equation [379, 380], while others are based on maximum likelihood (ML) methods. One ML method obtains parameter estimates by fitting transition density functions of stochastic differential equations in biochemical pathways [378]. A similar approach based on the ML of transitional probabilities requires measurements of the state trajectories at very fast sampling rate, whereby reactions are assumed to occur at most twice in a sampling time interval [381]. The fast sampling requirement makes this approach impractical, since biological data are typically sparse.

In this work, *three* kinetic parameter estimation methods for stochastic models were de-

veloped based on two criteria: maximum likelihood (ML) and density function distance (DFD). Two scenarios of practical application were considered involving both sparsely and densely populated datasets (i.e. low and high replicates). Since the distribution density functions are commonly constructed using histograms, an important aspect related to the binning strategy and the noise associated with finite sampling, has been incorporated in the parameter estimation framework. The efficacy of each method was evaluated and compared based on applications to *four* CME based case studies: RNA dynamics in *Escherichia coli*, gene expression network of galactose uptake model in *Saccharomyces cerevisiae*, and to two bimodal system a.) Schlögl model and b.) a genetic toggle switch in *E. coli*. Despite the use of CME models here, the methods are generally applicable to other stochastic models in which the system behavior or output can be characterized by a PDF of the states.

7.2 Methods

7.2.1 Estimation of state density function

In this work, Stochastic Simulation Algorithm (SSA) [214] was used to generate *in silico* experimental data for the purpose of parameter estimation and to solve for the PDF of the CME model. The histogram should reflect the true state PDF in the limit of the number of realizations tending to infinity. Since only a finite number of data samples are computationally feasible and experimentally practical, the error associated with histogram binning strategy is important, but this is not often discussed in existing literature of the CME. The shape of the resulting density function is known to be sensitive to the number and size of the bins, and the optimal binning distribution need not be of uniform sizes [382]. Characteristic features of a distribution such as bimodality may not be apparent when using bins that are too wide, while histograms can be significantly affected by random fluctuations associated with a small number of data points in bins that are too narrow. Although there is no hard and fast rule on the selection of bin sizes, the minimum number of realizations

in each bin should typically range between 5 and 20 [233]. Unless stated otherwise, the histograms here are constructed such that each bin contains no fewer than ten occurrences. The noise due to the histogram construction using finite size random sample will be taken into account in the parameter estimation below.

In practice, the choice of numerical solvers for model equations determines the performance of any parameter estimation methods. For CME, there has been a tremendous development of numerical algorithms for computing the PDF solution, directly [383–385] or indirectly [175, 214, 386]. The SSA was selected in this work because this algorithm is equivalent to the CME [175, 214], motivating its use to generate *in silico* data. Consequently, the CME model was also solved using SSA, such that the efficacy of the proposed methods can be evaluated independently from the solvers. In this case, deficiencies of SSA will appear equally in both *in silico* data and the model solution.

7.2.2 Parameter estimation methods

The methods developed here are formulated as a minimization of distance measures between model predictions and experimental data. The first method makes use of the common likelihood function and the second involves a distance metric between density functions as predicted by the CME and the data. When experimental error is known or can be determined from data, this noise should be accounted for in the PDF solution. In this work, the error is assumed to be independent and identically distributed (i.i.d.) random samples from a normal distribution with zero mean and variance σ^2 ; ($\mathcal{N}(0, \sigma^2)$), which are then explicitly added to the SSA realizations.

Maximum Likelihood (ML) method

The first estimation criterion is the likelihood function given by:

$$L(\mathbf{k}) = \prod_{j=1}^m \prod_{i=1}^n f(\mathbf{o}_i^j, \mathbf{t}_i; \mathbf{k}), \quad (7.1)$$

where the j -th experimental replicate $\{\mathbf{o}_1^j, \mathbf{o}_2^j, \dots, \mathbf{o}_n^j\}$ are taken at time points $\{t_1, t_2, \dots, t_n\}$ for $j = 1, 2, \dots, m$ (i.e. the experiments are done in m replicates). The likelihood function $f(\mathbf{o}_i^j, \mathbf{t}_i; \mathbf{k})$ is given by the CME model, which in this case is evaluated from the density function histogram of SSA realizations. The parameter estimation is then formulated as maximization of the likelihood function given by:

$$\mathbf{k}^* = \arg \max_{\mathbf{k}} L(\mathbf{k}) \quad (7.2a)$$

$$= \arg \max_{\mathbf{k}} \prod_{j=1}^m \prod_{i=1}^n f(\mathbf{o}_i^j, \mathbf{t}_i; \mathbf{k}) \quad (7.2b)$$

$$= \arg \max_{\mathbf{k}} \prod_{j=1}^m \prod_{i=1}^n P(\mathbf{o}_i^j, t_i | \mathbf{x}_0, t_0) \quad (7.2c)$$

where $P(\mathbf{o}_i^j, t_i | \mathbf{x}_0, t_0)$ is the state PDF reconstructed from SSA simulations, with added Gaussian i.i.d. noise $\epsilon \in \mathcal{N}(0, \sigma^2)$ when appropriate, i.e. the state trajectory is simulated as $\mathbf{o} = \mathbf{x} + \epsilon$ rounded to the nearest integer. For brevity, from hereon $P(\mathbf{o}_i^j, t_i | \mathbf{x}_0, t_0)$ will be denoted by $P(\mathbf{o}, t_i)$. Specific details of the accounting of experimental errors can be found in the description of the case studies in the results section. To avoid numerical underflows, the log-likelihood formulation of the objective function [Equation 7.2](#) is used in this work, giving:

$$\mathbf{k}^* = \arg \min_{\mathbf{k}} -\log L(\mathbf{k}) \quad (7.3a)$$

$$= \arg \min_{\mathbf{k}} \sum_{j=1}^m \sum_{i=1}^n -\log P(\mathbf{o}_i^j, t_i) \quad (7.3b)$$

Density Function Distance (DFD) method

The next two estimation methods are based on the minimization of state density function distance, similar to a divergence measure between two distribution functions, such as the Kullback-Leibler distance [387]. In particular, two estimation criteria are considered using the probability density function (DFD-PDF) and cumulative density function (DFD-CDF). In the PDF distance method, the objective of the parameter estimation is to minimize the difference between the PDF of the experimental data and SSA simulations, as follows:

$$\mathbf{k}^* = \arg \min_{\mathbf{k}} \sum_{i=1}^n \sum_{l=1}^{L-1} \frac{(P_e(\mathbf{o}_l, t_i) - P(\mathbf{o}_l, t_i))^2}{s_{l,i}^2}, \quad (7.4)$$

where $P_e(\mathbf{o}_l, t_i)$ denotes the experimental PDF constructed using a histogram with L bins and \mathbf{o}_l is arbitrarily taken to be the center of each bin. Unless stated otherwise, the binning strategy is referenced to the experimental data and the same binning distribution is used for the SSA simulations. The last bin represents an extra degree of freedom due to normalization of the sum (integral) of the PDF to 1, and thus not included in the optimization procedure. The weighting factor $s_{l,i}^2$ is an estimate of the variance of the l -th bin probability at time t_i arising due to finite random sampling. The process of classifying N elements from either the experimental data or SSA realizations into bins of a histogram can be assumed as a Binomial process and thereby the variance of the bin frequency is computed according to:

$$s_{l,i}^2 = \frac{P_e(\mathbf{o}_l, t_i)(1 - P_e(\mathbf{o}_l, t_i))}{N}. \quad (7.5)$$

As a reliable construction of a PDF typically requires a large number of replicates, the PDF distance may not be appropriate when only few replicates of data are available. On the other hand, the ML method above can be applied to datasets with low replicates, as it does not require the construction of a density function from the experimental data.

The second criterion considers the minimization of the differences between the CDF

constructed using the experimental data and the SSA realizations, given by:

$$\mathbf{k}^* = \arg \min_{\mathbf{k}} \sum_{i=1}^n \sum_{l=1}^{L-1} \frac{(F_e(\mathbf{o}_l, t_i) - F(\mathbf{o}_l, t_i))^2}{S_{l,i}^2}, \quad (7.6)$$

where the CDF $F_e(\mathbf{o}_l, t_i)$ gives the probability to obtain an experimental observation $\mathbf{o} < \mathbf{o}_l$, and $F_e(\mathbf{o}_l, t_i)$ and $F(\mathbf{o}_l, t_i)$ denote the CDF constructed from the cumulative sums of the PDF, $\sum_{k=1}^l P_e(\mathbf{o}_k, t_i)$ and $\sum_{k=1}^l P(\mathbf{o}_k, t_i)$, respectively. Similar to the PDF criteria, the weighting factor $S_{l,i}^2$ is estimated using a binomial assumption to give:

$$S_{l,i}^2 = \frac{F_e(\mathbf{o}_l, t_i)(1 - F_e(\mathbf{o}_l, t_i))}{N}. \quad (7.7)$$

The binning distribution can be kept the same as the PDF, but this need not be necessarily so. Unlike PDF, the shape of CDF is less sensitive to noise from finite sampling, with the exception of the tail ends of the CDF near the minimum and maximum values of the states. The lesser sensitivity to noise also makes the CDF distance method applicable to sparse datasets (low replicates), in which case the binning strategy is done based on the SSA realizations.

7.2.3 Global optimization algorithm

Aside from model solvers, the effectiveness of any parameter estimation methods also depends on the ability to find the global optima to the minimization problems. In the case of stochastic models, the error landscape is anticipated to be highly stochastic due to noise from finite experimental data points, which prevents the use of any optimization algorithms involving gradient search. Here, a variant of evolutionary algorithms, called Differential Evolution (DE), is used as a general purpose global optimization algorithm. This method can effectively handle diversified objective function planes [388], and like other evolutionary algorithms such as genetic algorithm (GA), DE starts with a random popu-

lation member and looks for the global optima by generating new population members using successive recombination and mutations based on the original parent population. However, unlike GA, DE uses floating point instead of bit string encoding, and arithmetic operations instead of logical rules, thereby providing a greater flexibility in the parameter search. Among the settings of DE, the population size and total number of generations are tuned in the case studies below based on the dimensionality of the problem (i.e. number of parameters) and the choice of parameter estimation method, respectively. The remaining parameters are maintained at previously suggested values [388]. The convergence and termination of the optimization can be based on the improvement of the best objective function in the population, standard deviation of the population vector, or maximum difference between the best and worst population member. A combination of several of these criteria can provide an efficient and robust termination criterion [389]. Since the case studies considered in this work involve *in silico* data with known true parameters, a maximum iteration number is used as a termination criteria and the efficacy of each method is judged based on the accuracy of the respective estimates.

The SSA and DE algorithms were implemented using Message Passing Interface (MPI) in C++ and run on a Linux IBM computing cluster (CentOS; GNU C++ compiler (v4.1.1)). A combination of a long period random number generator [366] and multiple independent streams generator [367] were used to guarantee statistically independent streams of random numbers required for both the SSA and DE.

7.3 Results

7.3.1 Case Study 1: RNA dynamics in *E. coli*

The significance of intra-cellular noise arises from the low copy number of genetic materials and gene transcriptional machinery. Thus, the quantification of mRNA would experience a greater influence of such noise than that of proteins, which may have thousands of

copies. A high resolution fluorescence microscopy method has been developed to quantify the molecular count of mRNAs in individual *E. coli* cells [372]. This method is based on the amplification of MS2d-fused fluorescence protein signal by binding to a reporter RNA that has multiple MS2d receptor sites (Figure 7.1A). The transcriptional response was shown to rise and plateau after 70-80 minutes post induction [372]. The molecular counts of the transcripts were obtained by normalizing the fluorescence flux with that generated by a single tagged RNA molecule. A mass-action kinetic model of the average mRNA level was used to fit the experimental data to obtain the kinetic parameter values [372].

A. Genetic construct of RNA expression in *E. Coli*

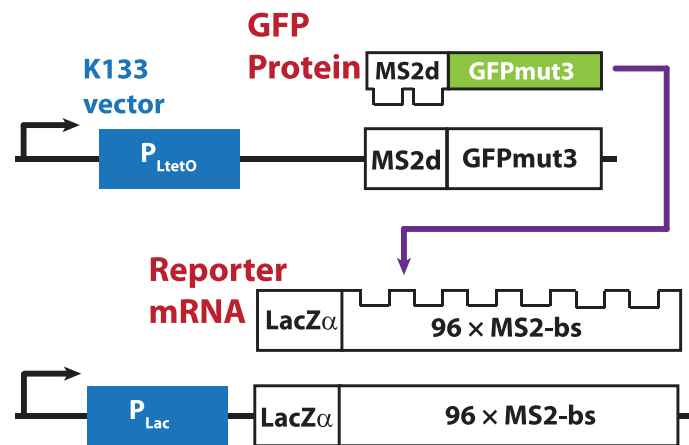
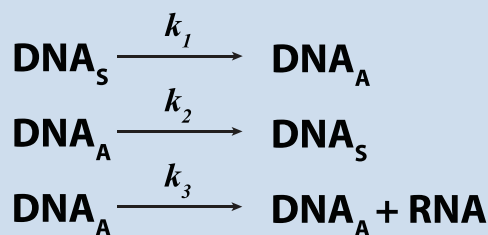


Figure 7.1: mRNA Dynamics Model in *E. coli*. (A) The mRNA detection system comprises two genetic elements; a fluorescence protein fused with bacteriophage protein (MS2d) and a reporter mRNA containing tandem repeats of MS2-binding sites. The GFP binding site repeats facilitate imaging and quantification of cellular mRNA to single molecular level. (B) The transcriptional model constitutes 3 reactions with 3 rate constants. DNA_S represents the silent form, while DNA_A represents the activated form.

B. Reaction Scheme



The first case study uses the CME model corresponding to the reactions and kinetic parameters proposed in the original work, as shown in Figure 7.1B and detailed in Table 7.1 [372]. Considering this model to be the true system, four experimental datasets of mRNA copy numbers with different replicates ($m = 10, 20, 100,$ and $10,000$) were simulated using the SSA. The simulated data were contaminated with measurement errors arising due to the normalization of the fluorescence flux, were taken to be discrete rounded values

of normal random samples ($\mathcal{N}(0, 0.25)$), consistent with the actual wet-lab experiments [372]. The mRNA transcripts per cell generation were recorded every 0.5 minutes until 75 minutes, mimicking the original experimental protocol.

Table 7.1: SSA formulation of the RNA dynamics model.

SSA implementation		
Reactions	Propensities	Stoichiometric change
1	$k_1 \times DNA_S$	$DNA_S = 0; DNA_A = 1$
2	$k_2 \times DNA_A$	$DNA_S = 1; DNA_A = 0$
3	$k_3 \times DNA_A$	$mRNA = mRNA + 1$

The parameter search was constrained to a space bounded by $\mathbf{k} \in [0, 5]$. The density functions predicted by the CME were constructed using 10,000 SSA realizations with added i.i.d and ($\mathcal{N}(0, 0.25)$) noise. In the case of low replicate datasets ($m = 10, 20$, and 100), only the DFD-CDF method was applied, in which the CDF of the experimental data was constructed according to: [233]

$$F_e(o_l, t_i) = \frac{l - 0.5}{m}. \quad (7.8)$$

where l now denotes the index of the state in replicate vector after arranging the data in ascending order (i.e., $o_1 \leq o_2 \leq \dots \leq o_m$). This construction implicitly uses the differences between sorted data values as the bin sizes. As stated earlier, since the DFD-PDF method requires the histograms of experimental data, which in the case of low replicate datasets, are highly inaccurate, this method was only performed for cell population data ($m = 10,000$). The DE optimization was implemented with a population size of 30 ($10 \times$ the number of parameters) for 4,000 generations and the optimization routine took about 1.5 hours for completion.

Table 7.2 presents the parameter values estimated using the ML and DFD methods for all datasets. In general, the parameter estimates were closer to the true values with increasing number of replicates, as expected from the increase of information with higher

replicates. The DFD(-CDF) method generally performed better than the ML. Amongst the parameters, k_1 is the most accurately determined parameter by all methods. At higher replicates, the DFD-CDF method converged to the true solution faster than the PDF and ML methods, in this order, which could be attributed to the difference in the shape of the objective function surface. As seen in [Figure 7.2A](#) and [7.2C](#), the DFD-CDF criterion produced a higher surface curvature (second derivatives) than those of ML and DFD-PDF ([Figure 7.2B](#), [7.2D](#), [7.2E](#)). Using a larger population size and higher number of iterations (100 population members and 20,000 generations), the ML method was able to match the accuracy of the CDF estimates (see [Table 7.2](#), $m = 10$).

Table 7.2: Parameter estimation of *E.coli* RNA dynamics model. Parameter estimates in the mRNA dynamics model in *E. coli*. The true parameter values are $\mathbf{k} = [0.0277; 0.1667; 0.4]$. The search bound for the optimization algorithm was $[0, 5]$.

‡ DE optimization performed with 100 population members and 20,000 generations.

Replicates	ML			DFD-CDF			DFD-PDF		
	k_1	k_2	k_3	k_1	k_2	k_3	k_1	k_2	k_3
10	0.0235	1.304	3.2201	0.02	0.1029	0.3643	—	—	—
	(0.02333) [‡]	(0.3231) [‡]	(0.7232) [‡]	—	—	—	—	—	—
20	0.0227	0.1095	0.2858	0.0371	0.2124	0.5263	—	—	—
100	0.0362	0.2930	0.5533	0.0273	0.1702	0.4121	—	—	—
10000	0.0279	0.2354	0.4872	0.0276	0.1659	0.4102	0.0273	0.1532	0.3837

7.3.2 Case Study 2: Galactose uptake model in *S. cerevisiae*

The inherent stochastic nature of gene expression can lead to diversified responses in a (clonal) cell population, even when subjected to uniform external conditions. This diversity has been demonstrated in a cell population using fluorescence techniques such as flow cytometry (FACS). The second case study used in this work looks at the problem of estimating CME parameters from a cell population data. The model describes an artificial genetic construct with the green fluorescence protein (GFP) gene downstream of a galactose activated promoter UAS_G and a TetR repressor binding element $2xtetO_2$ ([Figure 7.3A](#)). In the presence of galactose, the GFP expression can be modulated rheostatically by varying the level of inducer ATc [177]. The original publication utilized a clonal population

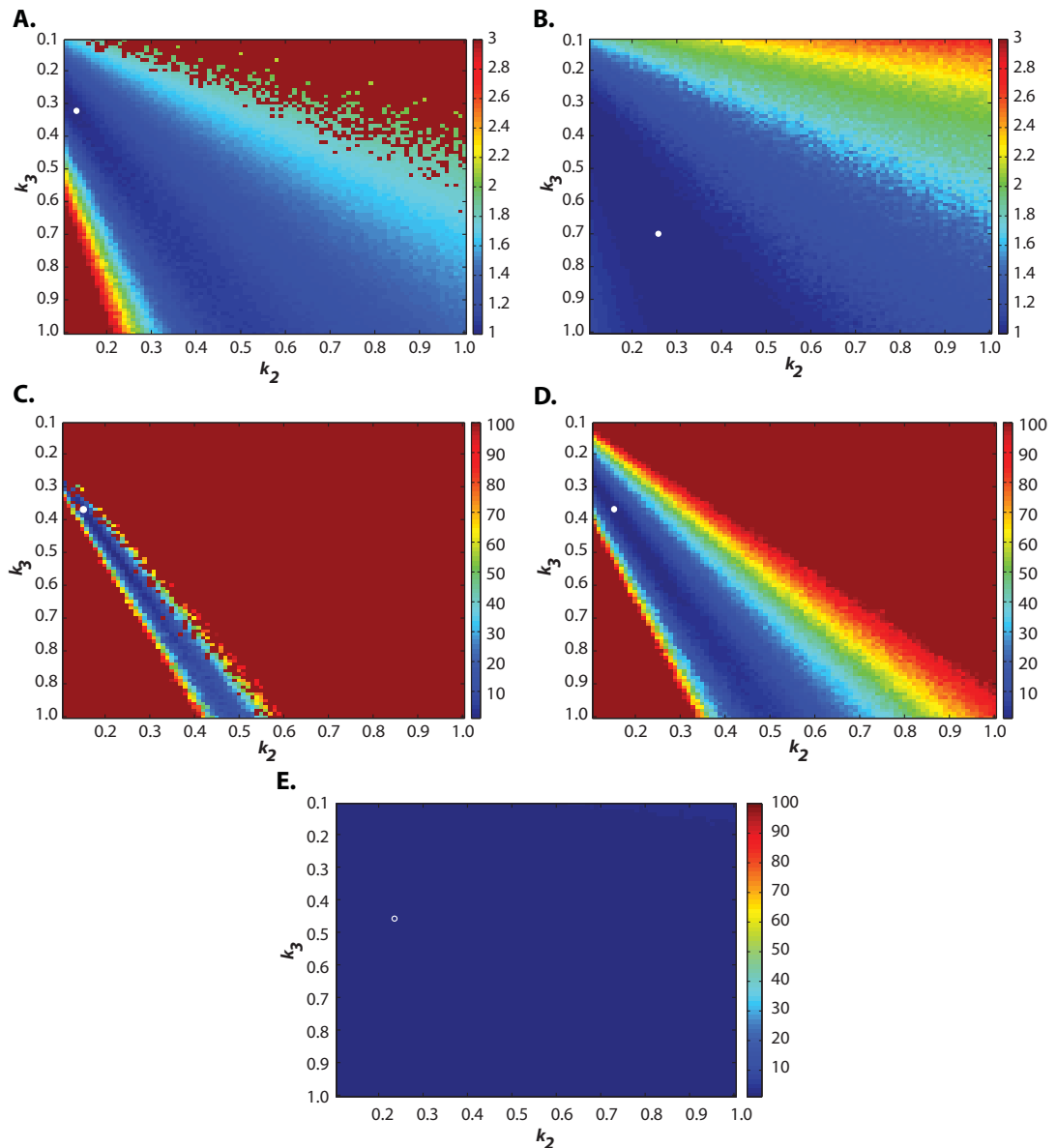


Figure 7.2: **Normalized objective function contours of the ML and DFD methods in the *E. coli* RNA dynamics model.** The parameter values k_2 and k_3 were varied between 0.1 and 1 while keeping the value of k_1 at its original value. The normalization was done with respect to the optimal solution from each parameter estimation plane, where the white circles represent the extrema on the normalized objective function plane. (A-B) Normalized objective function contours of the DFD-CDF and ML methods using sparse datasets ($m = 10$), respectively. (C-E) Normalized objective functions of the DFD-CDF, -PDF and ML methods using population datasets ($m = 10,000$).

of *S. cerevisiae* (yeast) to investigate the inherent cellular noise in the GFP gene expression, which is measured as the heterogeneity of fluorescence among the cells.

The CME model adapted from this work captures the random transitions among all possible promoter states as shown in Figure 7.3B. The states PC_1 , PC_2 and PC_3 represent free/silent, intermediate complex, and pre-initiation complex promoter configurations,

respectively, while the states RC_1 and RC_2 describe different forms of repressed promoter configurations. The transcriptional (RNA synthesis) and translational (protein synthesis) processes are modeled as single-step irreversible reactions (Figure 7.3B).

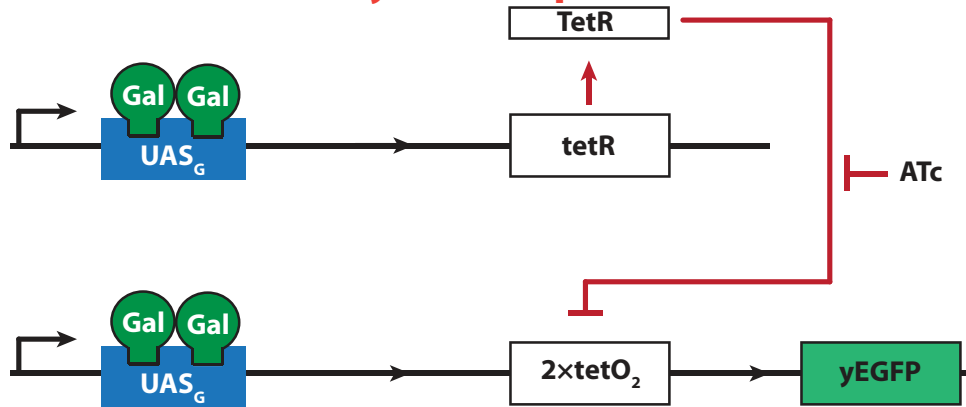
In the simplified model, the different promoter configurations are assumed to be in equilibrium, which reduces the model to a set of 8 irreversible reactions, 4 states, and 8 kinetic parameters, as shown in Figure 7.3B (dashed boxes) [177]. As in the first case study, this model was considered to be the true system and the molecular data of yEGFP and TetR were generated using SSA, giving 10^4 realizations at every 5 dimensionless time units up to 50 (or about 18 times the half life of yEGFP [390]). This condition corresponds to 440 minutes of post induction by 2% galactose and 40 ng ml^{-1} ATc. To study the scalability of the proposed methods, the parameter estimation of the full network with 18 reactions, 9 states, and 15 kinetic parameters was also done using a second *in silico* dataset with 10^4 SSA realizations from the complete model. The details on the CME formulation for both the reduced and the complete model of the yEGFP gene expression pathway have been described in Table 7.3 and Table 7.4.

Both ML and DFD methods were first applied to the reduced model, in which the DE optimization was done with 80 population members for 4000 generations, which took about 50 hours for convergence. The bounds on the parameter search space are given in Table 7.5. As mentioned above, the binning strategy in the DFD methods was based

Table 7.3: SSA formulation for the reduced model of the reduced - yEGFP galactose utilization pathway.

Reactions	SSA implementation	
	Propensities	Stoichiometric change
1	$\kappa_R \langle PC_3 \rangle$	$\text{mRNA}_G = \text{mRNA}_G + 1$
2	$\gamma_R \times \text{mRNA}_G$	$\text{mRNA}_G = \text{mRNA}_G - 1$
3	$\kappa_P \times \text{mRNA}_G$	$\text{yEGFP} = \text{yEGFP} + 1$
4	$\gamma_P \times \text{yEGFP}$	$\text{yEGFP} = \text{yEGFP} - 1$
5	κ_R^t	$\text{mRNA}_R = \text{mRNA}_R + 1$
6	$\gamma_R^t \times \text{mRNA}_R$	$\text{mRNA}_R = \text{mRNA}_R - 1$
7	$\kappa_P^t \times \text{mRNA}_R$	$\text{TetR} = \text{TetR} + 1$
8	$\gamma_R^t \times \text{TetR}$	$\text{TetR} = \text{TetR} - 1$

A. Genetic construct of yEGFP expression in *S. cerevisiae*



B. Repressor pathway

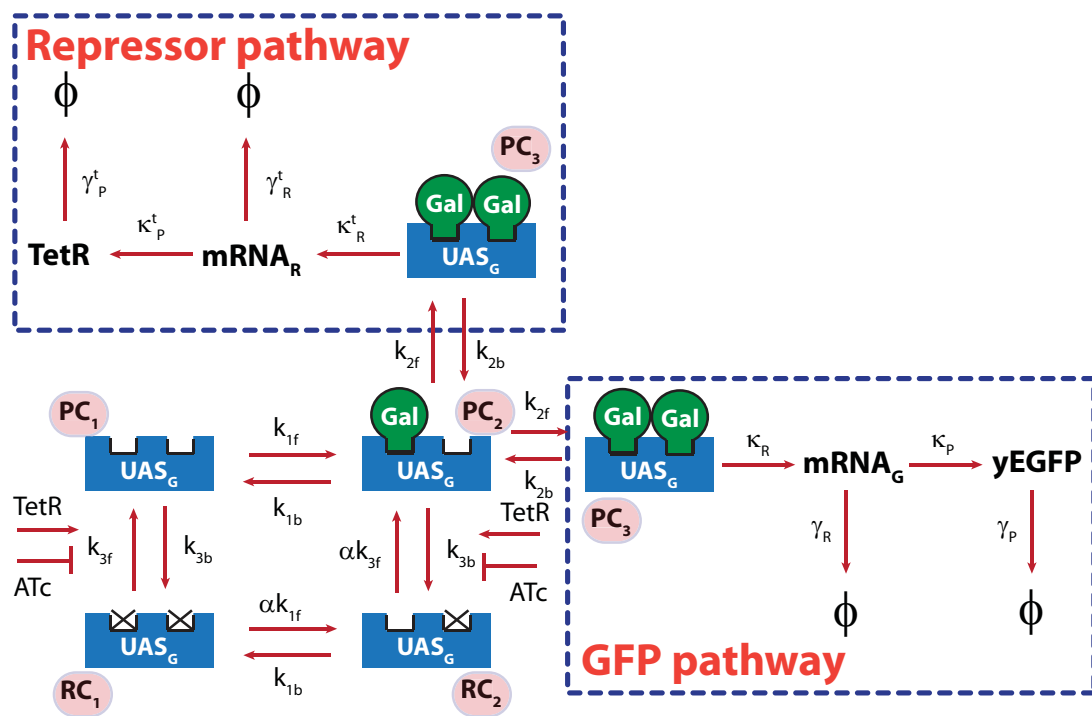


Figure 7.3: **Gene Expression Model for the Preferential Galactose Uptake in Yeast Cells.** (A) Genetic construct of the transcriptional control of the yeast-enhanced green fluorescent protein expression in the galactose utilization pathway of yeast. (B) The complete gene expression pathway includes (fast) reversible transformations among different promoter configurations and subsequent irreversible RNA and protein synthesis pathways. The reduced model assumes pseudo-equilibrium among the promoter configurations, and thus only describes dynamics of processes in the dashed boxes.

on the simulated experimental data, while the likelihood function in the ML method was constructed based on the histogram of SSA simulations. Table 7.5 presents the parameter estimates from the ML and the two DFD methods along with the true parameter values.

Table 7.4: SSA formulation for the complete model of the full yEGFP galactose utilization pathway.

Transcriptional reversible processes			Irreversible process		
Rxns	Propensities	Stoichiometric change	Rxns	Propensities	Stoichiometric change
1	$k_{1f} \times PC_1$	$PC_1=0; PC_2=1$	1	$\kappa \langle PC_3 \rangle$	$mRNA_G = mRNA_G + 1$
2	$k_{1b} \times PC_2$	$PC_1=1; PC_2=0$	2	$\gamma_R \times mRNA_G$	$mRNA_G = mRNA_G - 1$
3	$k_{2f} \times PC_2$	$PC_2=0; PC_3=1$	3	$\kappa_P \times mRNA_G$	$yEGFP = yEGFP + 1$
4	$k_{2b} \times PC_3$	$PC_2=1; PC_3=0$	4	$\gamma_P \times yEGFP$	$yEGFP = yEGFP - 1$
5	$k_{3f} \times PC_1$	$PC_1=0; RC_1=1$	5	κ_R^t	$mRNA_R = mRNA_R + 1$
6	$k_{3b} \times RC_1$	$PC_1=1; RC_1=0$	6	$\gamma_R^t \times mRNA_R$	$mRNA_R = mRNA_R - 1$
7	$\alpha \cdot k_{1f} \times RC_1$	$RC_1=0; RC_2=1$	7	$\kappa_P^t \times mRNA_R$	$TetR = TetR + 1$
8	$k_{1b} \times RC_2$	$RC_1=1; RC_2=0$	8	$\gamma_P^t \times TetR$	$TetR = TetR - 1$
9	$\alpha \cdot k_{3f} \times PC_2$	$PC_2=0; RC_2=1$			
10	$k_{3b} \times RC_2$	$PC_2=1; RC_2=0$			

As in the first example, the DFD-CDF method gave the most accurate estimates, followed by the DFD-PDF and ML methods, respectively. As illustrated in Figure 7.2C, 7.2D & 7.2E, the differences in the performance of these methods again arises from the steepness of the objective function plane. However, the lesser performing methods can potentially match the accuracy of the CDF method if population size and number of iterations in the DE optimization are increased.

The scalability of the methods discussed in this work was evaluated by performing the estimation of the complete yEGFP Galactose uptake model (ref. Table 7.6). In this case, the DE optimization was performed using 150 population members for 4000 generations and took approximately 60 hours for convergence. In this case also, the CDF method again generally outperformed the PDF and ML (Table 7.6). The parameter estimates from DFD-CDF gave yEGFP-PDF that is in agreement with wet-lab data (Figure 7.4)[177]. But

Table 7.5: Parameter estimation of *S. cerevisiae* reduced yEGFP model.

Parameters	ML	DFD-CDF	DFD-PDF	Bounds	True values
κ_R	1.1443	1	1.0478	[0, 5]	1
κ_P	1.0382	1.005	1.2174	[0, 5]	1
γ_R	4.5036	5.0306	5.7355	[0, 10]	5
γ_P	0.0128	0.0126	0.012	[0, 5]	0.0125
κ_R^t	0.428	0.432	0.431	[0, 5]	0.417
κ_P^t	2.1254	1.0542	1.24	[0, 5]	1
γ_R^t	6.2433	2.9966	3.4982	[0, 10]	3
γ_P^t	0.0102	0.0114	0.0115	[0, 5]	0.0125

some of the parameters, especially those involving fast reversible processes, cannot be accurately identified from data. The lack of complete parameter identifiability is perhaps not surprising, when one considers that measurements of only few states are available and that the time scale of these measurements better reflects the slow kinetics of the irreversible processes.

Table 7.6: Parameter estimation results of *S. cerevisiae* full yEGFP model.

Parameters	ML	DFD-CDF	DFD-PDF	Bounds	True value
Transcriptional processes					
k_{1f}	0.4061	0.4082	0.4292	[0, 5]	0.42
k_{1b}	0.211	0.1171	0.8296	[0, 5]	0.2485
k_{2f}	74.1848	25.9882	99.7701	[0, 100]	50
k_{2b}	4.1423	18.8779	2.0815	[0, 20]	10
k_{3f}	3.2×10^{-3}	3.87×10^{-3}	0.0166	[0, 5]	3.023×10^{-3}
k_{3b}	17.2405	19.9408	19.7665	[0, 20]	10
α	0.1	0.0183	0.0211	[0, 5]	0.025
Irreversible processes					
κ_R	0.8939	0.9296	0.8078	[0, 5]	1
κ_P	2.0345	1.1103	1.0995	[0, 5]	1
γ_R	7.3543	5.2431	5.4116	[0, 10]	5
γ_P	0.0116	0.0124	0.012	[0, 5]	0.0125
κ_R^t	0.4376	0.4157	0.4152	[0, 5]	0.417
κ_P^t	1.7641	0.9755	1.3732	[0, 5]	1
γ_R^t	4.3235	2.9034	3.9315	[0, 10]	3
γ_P^t	0.0107	0.0116	0.0103	[0, 5]	0.0125

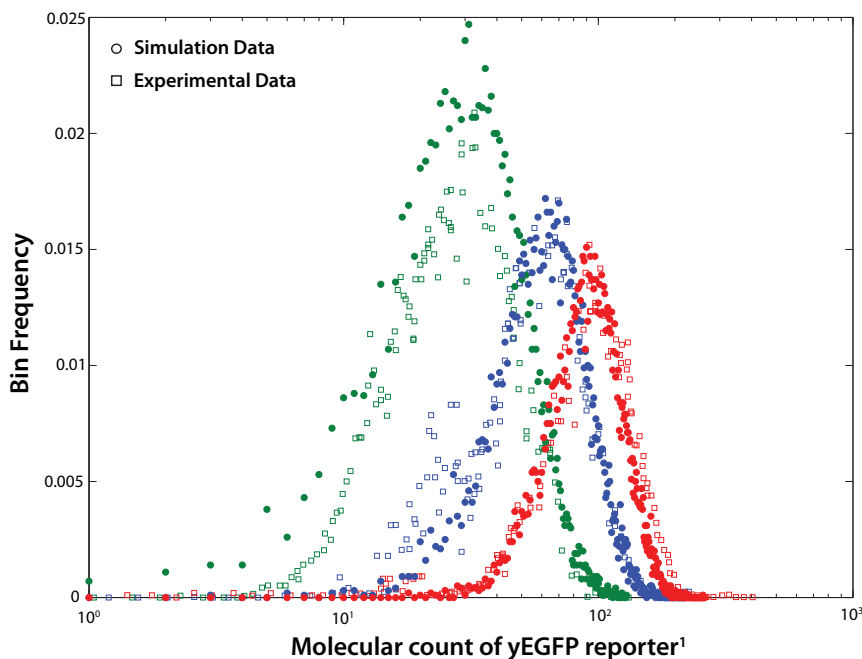


Figure 7.4: Validation of the parameter estimates obtained from the optimization of the complete yEGFP model. Comparison of actual experimental data and CME model prediction using SSA and parameters estimated in case study 2 (galactose uptake model in *S. cerevisiae*) (green - 150 minutes, blue - 290 minutes, and red - 440 minutes post-induction of 40 ng/ml ATc) [177].

¹Blake WJ, M KA, Cantor CR, Collins JJ: Noise in eukaryotic gene expression. *Nature* 2003, 422: 633- 637

Two other estimation criteria based on the maximum density function distance, were performed using PDF and CDF:

$$\mathbf{k}^* = \arg \min_{\mathbf{k}} \sum_{i=1}^m \max_{L-1} \frac{|P_e(\mathbf{o}_l, t_i) - P(\mathbf{o}_l, t_i)|}{\sqrt{s_{l,i}^2}} \quad (7.9)$$

and

$$\mathbf{k}^* = \arg \min_{\mathbf{k}} \sum_{i=1}^m \max_{L-1} \frac{|F_e(\mathbf{o}_l, t_i) - F(\mathbf{o}_l, t_i)|}{\sqrt{S_{l,i}^2}} \quad (7.10)$$

The outcome of the application of these criteria to the reduced and complete yEGFP gene expression pathway is summarized in Table 7.7 and 7.8. A comparison of optimization results in these tables indicates that the estimation criteria based on the maximum density distance can provide similar performance as those in the previous case.

Table 7.7: Parameter estimation of *S. cerevisiae* reduced yEGFP model using maximum probability distance measure. Estimated kinetic rate constants of the reduced yEGFP genetic transcriptional process using the DFD method with maximum distance measure.

Parameters	DFD-CDF	DFD-PDF	Bounds	True values
κ_R	1.0196	1.0311	[0, 5]	1
κ_P	0.9997	1.0735	[0, 5]	1
γ_R	4.92	5.1962	[0, 10]	5
γ_P	0.0127	0.0127	[0, 5]	0.0125
κ_R^t	0.4373	0.4285	[0, 5]	0.417
κ_P^t	1.1713	1.1122	[0, 5]	1
γ_R^t	3.276	3.1543	[0, 10]	3
γ_P^t	0.0115	0.01103	[0, 5]	0.0125

7.3.3 Case Study 3: Stochastic model of a synthetic toggle switch

Multi-stability is often seen in biological networks, such as in λ -phage decision circuit [391], MAPK cascade [173], and cell cycle regulation [392]. In particular, bistability is a common motif encountered in cellular signaling pathways [393]. Motivated by this, a genetic toggle switch had previously been engineered in *E. coli* to show the ability to synthesize such motif. The synthetic switch consisted of two repressor-promoter pairs, with (i) $P_{Ls1con} - lacI$ repressing P_{trc-2} promoter and (ii) vice versa $P_{trc-2} - cIts$ (thermal

Table 7.8: **Parameter estimation results of *S. cerevisiae* full yEGFP model using maximum probability distance measure.** Estimated kinetic rate constants of the complete yEGFP genetic transcriptional process using the DFD method with maximum distance measure.

Parameters	DFD-CDF	DFD-PDF	Bounds	True value
Transcriptional processes				
k_{1f}	0.4428	0.4192	[0, 5]	0.42
k_{1b}	0.1468	1.0296	[0, 5]	0.2485
k_{2f}	98.2117	45.7701	[0, 100]	50
k_{2b}	19.9595	1.0834	[0, 20]	10
k_{3f}	9.4113×10^{-3}	0.0232	[0, 5]	3.023×10^{-3}
k_{3b}	15.042	12.3425	[0, 20]	10
α	0.0707	0.0834	[0, 5]	0.025
Irreversible processes				
κ_R	0.9296	1.1078	[0, 5]	1
κ_P	1.1451	2.2095	[0, 5]	1
γ_R	5.6745	6.4432	[0, 10]	5
γ_P	0.0106	0.0222	[0, 5]	0.0125
κ_R^t	0.4062	0.5323	[0, 5]	0.417
κ_P^t	1.1087	2.3425	[0, 5]	1
γ_R^t	3.4129	4.1232	[0, 10]	3
γ_P^t	0.013	0.0324	[0, 5]	0.0125

sensitive) repressing P_{Ls1con} promoter [375], such that they are mutually inhibitory (see Figure 7.5A). The switching behavior was visualized by means of green fluorescence protein (GFP), inserted downstream of $cIts$. The ON switch was accomplished by an inducer, isopropyl β -D-thiogalactosepyronoside (IPTG), that represses the activity of $lacI$ (Figure 7.5A). By modulating the concentrations of the IPTG, the genetic toggle system could exhibit bistability with hysteresis [375].

A simple deterministic model was proposed to examine the behavior of the toggle switch and to analyze different conditions of bistability [375]. The corresponding CME formulation is described in the Figure 7.5B and 7.5C [394]. Here, the propensity functions are taken directly from the deterministic model and they give effective rates of reaction following a canonical Hill equation. Taking this model to be the true system, *in silico* data of GFP fluorescence at IPTG concentration of 6×10^{-5} M were simulated using 10^4 independent SSA realizations, emulating flow cytometry data.

As the ML performed consistently poorer than the DFD methods in the previous case studies, the stochastic rate constants here ($\alpha_1, \alpha_2, \beta, \gamma, \eta, K$) were estimated using the DFD-

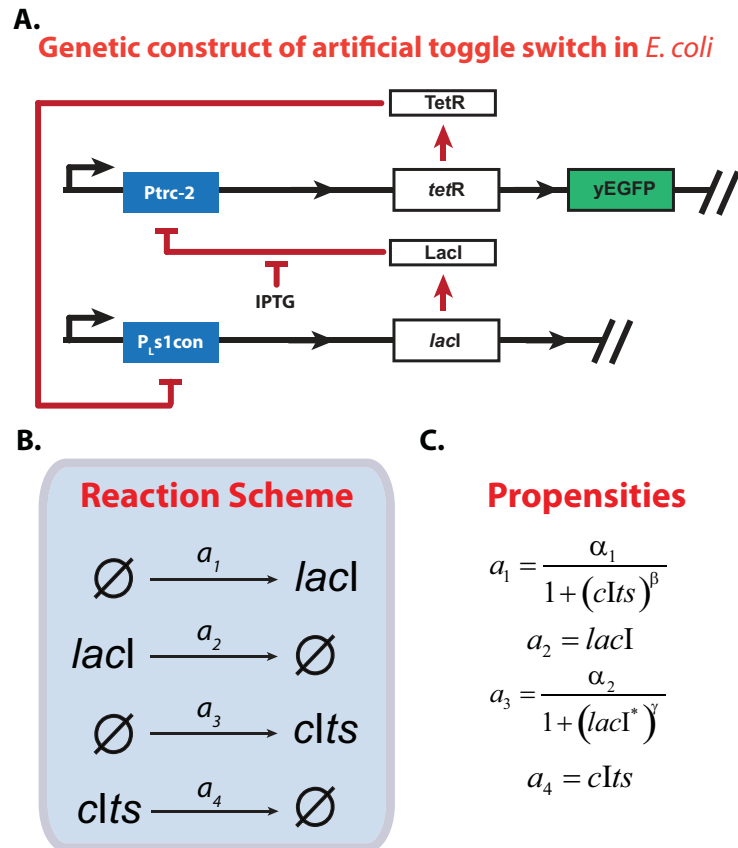


Figure 7.5: **Stochastic dynamics of synthetic gene toggle switch engineered in *E. coli*.** (A) Synthetic circuit of the genetic toggle switch of *E. coli* [375]. (B) The genetic model of the toggle switch comprising of 4 reactions and (C) the corresponding propensity functions.

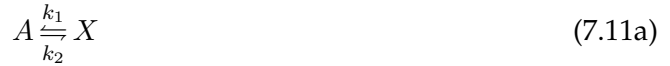
CDF and -PDF methods, with DE parameters: 150 population members and 4000 generations. Both CDF and PDF criteria took about 48 hours for completion. The parameter bounds and estimates are given in Table 7.9. Comparing to the true values, this case study, like the previous case studies, again showed that the DFD-CDF method performed better than DFD-PDF with more accurate and robust estimates of the kinetic rate constants.

Table 7.9: **Parameter estimation of synthetic toggle switch in *E. coli*.**

Parameters	DFD-CDF	DFD-PDF	Bounds	True values
α_1	137.716	99.456	[0, 200]	156.25
α_2	15.644	15.391	[0, 20]	15.6
β	2.309	2.543	[0, 10]	2.5
γ	1.071	1.015	[0, 10]	1.0
η	2.065	8.434	[0, 10]	2.0015
K	7.331×10^{-5}	5.831×10^{-4}	[0, 1]	6.0×10^{-5}

7.3.4 Case Study 4: The Schlögl Model

The Schlögl reaction has commonly been used as a typical example of a bistable system [395–397]



The concentrations of A and B are typically assumed constant (buffered). The propensity functions for the above reaction can be formulated as:

$$a_1 = k_1 [A] \Omega \quad (7.12a)$$

$$a_2 = k_2 X \quad (7.12b)$$

$$a_3 = k_3 X \frac{X-1}{\Omega} \frac{X-2}{\Omega} \quad (7.12c)$$

$$a_4 = k_4 [B] X \frac{X-1}{\Omega} \quad (7.12d)$$

where Ω ($\Omega = 10$ units is chosen here) is called the extensivity, a parameter that is proportional to the system volume. The actual values of the kinetic rate constants used for the SSA realizations are presented in [Table 7.10](#).

The molecular data of component X is formed as a density function constructed using 10^4 independent realizations of SSA, simulated for a time period of 20 units and recorded for every 4 time units. Parameter estimation was performed using the DFD method with both CDF and PDF criteria. The DE optimization was implemented with a population size of 40 and for 4000 generations and the estimation took about 72 hours for completion. [Table 7.10](#) shows the results of the parameters estimation, which again indicates a better performance of the DFD-CDF method compared to the DFD-PDF method.

Table 7.10: Parameter estimation of Schlögl model.

Parameters	DFD-CDF	DFD-PDF	Bounds	True values
$k_1 [A]$	0.4563	0.3125	$[0, 1]$	0.5
k_2	3.04418	3.0123	$[0, 5]$	3.0
k_3	0.9148	1.1362	$[0, 5]$	1.0
k_4	1.1083	1.4563	$[0, 5]$	1.0

7.4 Discussion

In this chapter, three practical methods are proposed for the estimation of the parameters from (noisy) single cell datasets with low and high replicates. As the methods rely on a histogram construction of density functions from a finite sample of experimental data and Monte Carlo simulations, the objective function evaluation has a trade-off between low accuracy when using bins that are too wide, and high sensitivity to noise when bins are too small. In order to balance this trade-off, the binning was done such that the narrowest bin has at least ten occurrences. The noise associated with this binning strategy is also taken into account in the objective function in the DFD methods, which is modeled according to a Binomial distribution.

The proposed methods are developed while considering a few practical issues when dealing with real biological datasets, such as data sparsity (low replicates), data noise and relatively coarse sampling intervals. The methods developed here do not require fast time-sampling like in [381], which might pose a restrictive constraint in practice. When population data are available, the DFD methods can fully exploit the additional information and rigorously handle the noise associated with the finite sample construction of a density function through the weighting factors. Although the examples considered in this work are represented by the CME, the methodologies developed in this work are generally applicable to parameter estimation of other stochastic models (e.g. Langevin), as long as the distribution density function can be constructed. Furthermore, the different methods developed in this work can be used to robustly estimate the rate constants of large scale

gene expression networks as well as systems with multistability and general nonlinear propensity equations.

The case studies above showed that methods based on matching density function shapes between model and data generally performed better than maximizing likelihood function. Furthermore, the DFD-CDF distance is more sensitive to parameters than both the DFD-PDF and ML, and thus is the most effective method developed in this work. The higher sensitivity of the CDF with respect to parameter variations is expected as a result of the cumulative sum of the PDF sensitivity. This is evident from comparing the normalized objective function surfaces as shown in [Figure 7.2](#), in which the CDF objective functions have the steepest curvature. The increased curvature leads to a faster convergence to the minima in the DE optimization of the CDF than the PDF, though both methods eventually converge to optimal parameter estimates with similar accuracy. In addition, the CDF is generally less sensitive to noise from finite sampling as can be seen from the noise weighting factor $S_{l,i}$ when normalized with the respective probability, i.e. the coefficient of variation (CoV) $S_{l,i}/F_e(\mathbf{o}_l, t_i) = \sqrt{1 - F_e(\mathbf{o}_l, t_i)}/n \cdot \sqrt{F_e(\mathbf{o}_l, t_i)}$. The monotonically decreasing CoV as a function of $F_e(\mathbf{o}_l, t_i)$ indicates that the CDF construction becomes less affected by finite sampling noise with increasing $F_e(\mathbf{o}_l, t_i)$.

Similar to the parameter estimation in deterministic models, parameter identifiability is a key issue in the estimation of the CME parameters. Such problem is commonly encountered in the parameter estimation of deterministic ODE models [398]. Following the same arguments from the deterministic estimation, the identifiability problem is caused by the limited information contained in the data about the parameters governing the fast transformations among the different promoter configurations. Such problem can be alleviated by getting additional measurements with a faster sampling rate and if possible, measuring the variables that are directly affected by the parameters, e.g. the fractions of promoters in each configuration of the second case study. An analogue of deterministic

parameter identifiability analysis can be performed using the parametric sensitivity of the density function and experiments can be designed to maximize the degree of information in the data [394, 399, 400].

Another aspect which is worth discussing the context of this work relates to the precision of the estimated parameters. Using identical experimental conditions, it is possible to generate several independent datasets. These different datasets can be potentially used to generate different parameter estimates using different, independent parameter estimation runs. Based on these independent parameter estimates, precision of the parameters in terms of confidence interval/region on the parameter estimates can be obtained. However using the stochastic parameter estimation framework developed in this work for estimating the parameter precision presents one of the significant challenges related to the computational burden. It would be computationally very expensive to run several, separate and independent runs of parameter estimation routine to obtain the parameter precision. Future work in extending this algorithm to determine the parameter precision has to look into avenues where these values can be obtained without having to resort to expensive Monte Carlo simulations.

Having said that, most of the computational cost of the parameter estimation related to CME is due to the large number of SSA realizations needed to construct the solution of the CME. Furthermore, every generation of DE requires multiple computations of the objective function according to the population size setting and each of population members in turn requires the SSA solution as mentioned previously. One way to alleviate the computational burden would be to lower the SSA realizations in constructing the density function. This would however increase the binning noise, and could possibly reduce the speed of convergence to the optimal solution and the accuracy of parameter estimates (see [Figure 7.6A-C](#)). Nevertheless, there is a diminishing return with increasing number of SSA realizations, since noise variance generally scales with the inverse of the number of sam-

ples (i.e. the standard deviation is only halved for every 4 times increase in the number of data). Alternatively, efficient approximation methods for simulating the CME can be used in place of the exact SSA [384, 386, 401–404], again at the cost of reduced estimation accuracy. In addition, the optimization parameters, namely population size and generations, can be further tuned for the proposed methods. Unfortunately, the relationship between these two parameters is most likely nonlinear and problem specific, which may require trial and error methods to find the best setting for a particular problem.

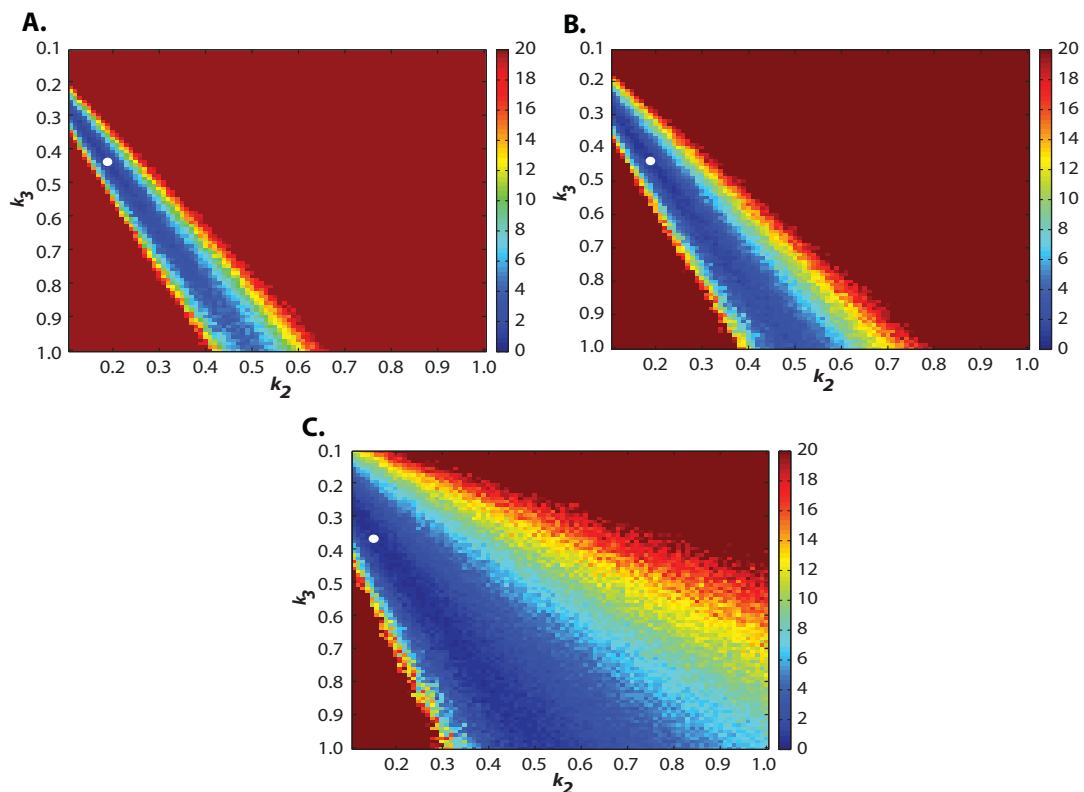


Figure 7.6: **Effect of the finite sampling noise on the parameter estimation of *E. coli* RNA dynamics model.** Normalized objective function contours of the DFD-PDF method for SSA realizations of 10,000 (A), 5000 (B), and 1000 (C). The parameter values k_2 and k_3 were varied between 0.1 and 1 while keeping the value of k_1 at its original value. The normalization was done with respect to the optimal solution from each case, where the white circles represent the extrema on the normalized objective function plane.

7.5 Conclusions

The inherent stochasticity associated with low copy number processes in the cellular genetic milieu can introduce significant noise in gene expression profiles. The modeling of such noisy system requires a careful consideration of random processes and the parameters governing the probability of random events [178, 179]. Three parameter estimation methods for stochastic models have been proposed based on the maximum likelihood criterion and density function distances of PDF and CDF. Since state density functions of stochastic systems are often constructed from a finite number of experimental data points or Monte Carlo realizations, a careful consideration has been taken to characterize the influence of noise arising from the histogram binning. Specifically, the effects of histogram noise are directly incorporated into the parameter estimation objective function as weighting functions. Applications to different case studies have shown that the proposed methods are both effective and practical. Amongst the proposed methods, the CDF-DFD method has been found to be the most efficient in estimating the kinetic rate constant than the others (i.e., the ML and DFD-PDF methods) due to the higher sensitivity of CDF to the parameters.

This chapter highlights the work done in S. K. Poovathingal & R. Gunawan. Global parameter estimation methods for stochastic biochemical systems. *BMC Bioinfo.*, 11:414, 2010.

Conclusions & future directions

8.1 Conclusions

The mFRTA proposes that ROS produced during cellular respiratory process (OXPHOS) in mitochondria causes degeneration of mitochondrial genome integrity resulting in cellular dysfunction and aging [1, 30, 227]. Several premises of mFRTA are intensely debated; however the direct role of mtDNA mutations in aging has been recently demonstrated in several single cell studies and in studies involving transgenic mitochondrial mutator mice. Thus, an understanding of the mitochondrial genetic integrity and its breakdown with time represents an important and active area of gerontological research. Many of the debates related to the fundamental premises of mitochondrial theory of aging relate to the uncertainty in the age-dependent mtDNA mutation accumulation data. The uncertainty associated with measurement data is a resultant of the inherent complexity of biological processes, especially related to the inherent stochasticity in the mtDNA maintenance process and variability arising from the measurement protocols. With the availability of efficient computational tools, it will be possible to develop and apply novel numerical algorithms to complement the state-of-the-art experimental techniques to have a greater comprehension of these complexities. Thus, the focus of the present dissertation is to develop a stochastic modeling framework, to capture the roles of different sources of stochasticity, like the intrinsic stochasticity associated with different mitochondrial genetic processes and randomness involved with the mutation quantification assays, in understanding the underlying dynamics of mtDNA mutation accumulation. The following

paragraphs summarize specific contributions of the present dissertation in addressing the knowledge gap in understanding the mechanisms of origin and accumulation dynamics of mtDNA mutations during aging.

To capture the stochasticity associated with the inherent mtDNA turnover dynamics, a stochastic model of mouse heart tissue was developed based on CME. The point mutation model captured the random drift of mtDNA mutations in heart tissue of mouse population during the embryonic development and post-natal life for wild-type mice and POLG mutator mice. The model was simulated using a Monte Carlo approach based on the Stochastic Simulation Algorithm (Chapter 3) [175, 229, 405]. The model constituted two main processes of mtDNA maintenance: mtDNA turnover and the associated *de novo* mtDNA mutations. The simulation outcome highlighted the importance of the coupling between experimental variability and intrinsic stochasticity of mtDNA mutation dynamics within the measurement data. The intrinsic stochasticity of mtDNA point mutation burden of the *in silico* mouse population was found to have a long tailed non-Gaussian density function, with a small fraction of the population harboring significantly higher mutation burden. The contribution of measurement variability to the overall data uncertainty was observed to be substantial, causing significant uncertainty in determining the underlying mtDNA mutation accumulation dynamics. Simulation results further indicated that the interpretation of the data can be significantly flawed if the statistical properties of mtDNA mutation quantification protocol are not considered critically. Unlike the experimental predictions [116], simulations of both the mouse models strongly indicated that the assumption of ROS mediated acceleration of mtDNA point mutation with age is not necessary to explain the measurement data. Model simulations indicated that *de novo* mutations arising during the cell lineage of embryonic development play a critical role in post-birth mutation accumulation in post mitotic tissues like heart. This effect is much more prominent in the case of error prone POLG mutator mouse models, where much of the high mutation burden in

the heart tissue of the transgenic mice is already acquired at birth.

Single cell biochemical analyses of respiratory deficient cells have indicated abundance of single type of mtDNA mutations, strongly hinting of the existence of a replicative mechanism favoring one type of mtDNA template, a process referred to as clonal expansion. However, the understanding of the functionality and the mechanism of this process is still incomplete. Several hypotheses have been proposed to explain the underlying clonal expansion mechanism. Considering the discreteness of mtDNA mutations dynamics, the stochastic model of mtDNA turnover process (Chapter 3) was modified to include specific features of cellular energy logistics in rat cardiomyocytes, to test different clonal expansion hypotheses and to elucidate the underlying mechanisms, which best explains the experimental data (Chapter 4). Consistent with our earlier model predictions (Chapter 3), simulation results in this work strongly indicated that the mosaicity seen in cellular mtDNA mutation burden in a tissue arise due to the stochasticity in mitochondrial turnover process, predominantly during developmental cell divisions. However, for explaining the clonal expansion dynamics of functional mtDNA mutations observed in post-mitotic tissues, both replicative advantage of mutant mtDNA and the up-regulation of nuclear retrograde response, which is modulated by the cellular energy deficit, are required. Simulation results also indicated that the extent of replicative advantage of mtDNA mutations plays a very important role in the progression of clonal expansion dynamics.

The stochastic mtDNA mutation models developed in this work were further extended into a multi-scale hybrid modeling (stochastic and deterministic model) framework to gain understanding of the etiology and progression of sarcopenia, an important pathology associated with aging (Chapter 5). Sarcopenia refers to an age related condition in which decline in skeletal muscle mass and strength occurs with age and causes frailty in aged organisms. The multi-scale mathematical model describes the following process: a.) developmental occurrences and expansion of mtDNA mutations in myo-nuclear regions of

skeletal muscle, b.) age-dependent spread and expansion of mtDNA along the muscle fiber, and c.) the associated progressive loss of muscle strength. The turnover of mtDNA in progenitor cells during embryogenesis was simulated using stochastic model. A reaction-diffusion model approximating the active transport of mtDNA nucleoids along with the mtDNA turnover dynamics in fused skeletal muscle fiber were formulated as partial differential equations. Observations of sarcopenic skeletal muscle fiber, such as the extent of clonal expansion in muscle fibers, and the number of fibers becoming respiratory deficient were validated with experimental data. It was observed that mtDNA diffusivity, which is a parameter that may be modulated by mitochondrial fusion-fission, plays a crucial role in determining the mtDNA mutation spread and thus in the progression of sarcopenia. Here, faster diffusion, though intuitively not desirable as it enhances the spread of mtDNA mutations, in fact led to fewer focal accumulations of respiratory deficient regions in skeletal muscle tissue. This arises predominantly due to the reduction of the nuclear retrograde response at diluted mtDNA mutation burden. The hybrid model was further used to perform parametric sensitivity analysis to determine the effect of different mtDNA turnover parameters on the severity and progression of sarcopenia, thereby providing important insights into the possible interventions to retard severity of sarcopenia. Specifically, using the hybrid model, different commonly hypothesized mechanisms for the beneficial effects of caloric restriction in retarding the progression of sarcopenia were analyzed. Simulation results indicated that somatic cells ability to sustain enhanced stress, i.e. to withstand higher cellular mtDNA mutation burden, might be a potential mechanism to explain the intervention of CR in reducing the progression of sarcopenia. Further, simulation results also indicated that the beneficial effects of CR accrue gradually and are neither rapidly inducible nor reversible upon the dietary regime transfer.

As the simulation results of mtDNA point mutation model pointed to large uncertainty in the mutation measurement data, especially when using mutation quantification proto-

col necessitating extremely high levels of molecular dilutions ($\sim 10^{-6}$ bp $^{-1}$) like the RMC assay [116, 141, 226]. While the RMC assay, represent an important technological advance, our findings in this thesis indicated that the statistics associated with the sample preparation in these assays is of great importance, but often an overlooked aspect. Consequently, sampling protocols need to be carefully designed to maximize signal-to-noise ratio. A statistical design of experiment to optimize the RMC protocol was developed using linearized variance propagation analysis and Monte Carlo methods (Chapter 6). The DOE revealed that the optimized RMC protocol is based on optimal dilution factor of about 1.6 mutant DNA template per well and counting *unamplified wells*, compared to 0.1 template per well of the conventional RMC assay and quantifying the fraction of *amplified wells*. Experimental validation was carried out using a mock RMC assay based on a plasmid system. Both the simulation results and experimental validation indicated that the optimized protocol can provide a substantial improvement in measurement accuracy and reduction in variability, with 10-times improvement in the information offered per PCR well, i.e. the optimal protocol can achieve the same coefficient of variation using one-tenth the number of wells as in the original assay. The reduction of the measurement variability obtained by the RMC optimization further provided more sensible estimate of the age-dependent mutation dynamics by reducing the noise related to mtDNA mutation burden data. While this work pertains to the optimization of RMC assay, the statistics related to sample preparation and sampling protocol maybe more generally applicable to other DNA quantification techniques requiring high dilution to obtain single molecule end point dilutions.

One of the most critical challenges experienced during the development of the present modeling framework relates to the availability of "kinetic" rate constants of stochastic biochemical systems. Three parameter estimation methods for stochastic models have been proposed in this dissertation based on the maximum likelihood (ML) criterion and density function distances of DFD-PDF and DFD-CDF (Chapter 7). Since state density functions of

stochastic systems are often constructed from a finite number of experimental data points or Monte Carlo realizations, a careful consideration has been taken to characterize the influence of noise arising from the histogram binning. Specifically, in this work, the effect of histogram noise are directly incorporated into the parameter estimation objective function as weightage functions. Applications to different case studies of gene transcription models had shown that the proposed methods are both effective and practical. Amongst the proposed methods, the DFD-CDF method has been found to be the most efficient and robust in estimating the kinetic rate constants than the others (i.e., the ML and DFD-PDF methods) due to the higher sensitivity of CDF to noise.

8.2 Recommendations

The stochastic mtDNA mutation models developed in this work are based on the mtDNA turnover process in murine tissues (primarily heart tissue). This model can be further developed to study the origin and accumulation dynamics of mtDNA mutations in different murine tissues and tissues of other commonly used animal models in gerontological research (like *Caenorhabditis elegans*, *Drosophila melanogaster*, Rhesus monkey, Human and Bats). One of the motivations in developing such *in silico* models of different organisms is to understand the importance of mtDNA turnover process during different stages of an organism life. For example, in this dissertation it was observed that mouse development plays a critical role in determining the tissue level mtDNA mutation burden. Developmental programs and the associated cell lineage differ significantly between different organisms. Similarly, the organism's life span also varies considerably between different species, and it is significantly challenging/expensive to estimate the tissue level mutation burden data using experimental techniques. By developing accurate models that captures the mtDNA turnover process, it would be relatively straightforward to estimate the tissue level mtDNA mutation burden at different stages of organism's lifespan. Another inter-

esting aspect that can be studied with this kind of modeling framework is to estimate the tissue level mtDNA mutation burden in different organisms at the end of life span and to understand the existence of the correlation between mtDNA mutation burden and its effect on the cellular/organismal aging, if any.

One interesting aspect that could be studied as an extension of the present thesis is the mtDNA mutation burden in the mitotic tissues like (blood cells, epithelial cells, spleen etc). Quantification of the mtDNA mutations in the context of gerontological research has largely been with somatic cells having only relaxed replication of mtDNA. In mitotic cells, aside from the normal mtDNA turnover, the mtDNA population undergoes enormous quanta of replication due to mitotic cell division. Also, unlike the post mitotic cells, the mitotic cells are turned over on a regular basis (removed and replaced from the tissue with new cells). Since very limited experimental data are available on the mtDNA mutation load in the mitotic tissues, it would be interesting to estimate the mtDNA mutation burden in these tissues, by modifying the simulation framework developed in this work to include specific features related to the cell division. Additionally, one of the interesting insights which is largely neglected in the original works [130–133, 287], relates to the mtDNA mutation frequency in fast dividing tissues of POLG mutator mice. Experimental observations indicated that one of the most prominent pathologies associated with the fast dividing tissue of POLG mutator mouse models manifest in the form of alopecia (skin epithelial follicles), spleen enlargement and anemia [130, 131, 287]. These data indicate that the mtDNA mutation burden associated with these tissues might be much higher and much more significant than the mtDNA mutation burden found in post mitotic tissues. Treatment of cell division and selection pressure for mitochondrial turnover in the mitotic tissues of POLG mutator/normal mouse models might be a promising area of investigation for the future work.

Since the discovery of mtDNA degenerative diseases, several remarkable features of

different syndromes were identified. For instance, these conditions are often inherited, though not in the same way as the disorder arising from mutation in nuclear genes [25, 108, 123]. Similarly, the resulting symptoms are more unpredictable than those caused by nuclear genetic mutations. In the cellular proliferation during animal development, the cells approximately double the mtDNA population before undergoing division and thus provide roughly equal amount of mtDNA contents to the dividing daughter cells. However, the cell does not regulate the distribution of the mtDNA content to the dividing daughter cells. Consequently, one daughter cell may inherit large fraction of mutant mtDNA and the other may inherit a large fraction of normal mtDNA. Thus the level of heteroplasmy in different organs that arise from these early developmental progenitor cells may be markedly different. The laws of probability that governs the segregation of the different mtDNA templates between the dividing cells are not clearly known. In the simulation framework developed in this dissertation, equi-partition of the different mtDNA templates are assumed between the dividing daughter cells based on a well mixed assumption. However, in reality this assumption might not be suitable, and it would be interesting to extend this work to study the underlying random process that governs the distribution of different mtDNA templates during the embryonic cell division and its consequence on the tissue-level heteroplasmy.

Another aspect that deserves further investigation relates to the mitochondrial fusion-fission process. Due to constant fusion-fission process, mitochondria are considered to be a highly dynamic organelle. However, given that the mtDNA are bound to each other in nucleoids and since nucleoid are tethered to cellular structures, it is non-trivial to intuitively understand the dynamics of mtDNA mutations arising from the mtDNA turnover process in the presence of mitochondrial fusion-fission process. A well mixed assumption of mtDNA population was used in the stochastic models developed in this thesis. An extension of the present model to include the specific feature of mitochondrial fusion-fission

process might provide a greater understanding of the dispersion of different mitochondrial templates (both normal and mutant mtDNA) across the cellular mitochondrial population [406]. Asymmetric segregation of the mtDNA between the mitochondria undergoing fission could result in stochastically giving rise to one of the daughter mitochondria that has significantly higher mutation burden. Such asymmetric segregation can give rise to mutational hot spots within a cell which could potentially result in clonal expansion process. Our group is presently investigating the role of mitochondrial fusion-fission process in the mtDNA mutation accumulation dynamics.

As discussed in the earlier chapters, accurate estimation of rate constants associated with mitochondrial turnover process are significantly difficult with the existing technological advances in measurement technology. Extension of the stochastic parameter estimation framework (Chapter 7) based on CME, could be an attractive alternate to estimate parameters which are difficult to estimate. For example, in studying the tissue level mtDNA mutation burden, generally the aggregate level or the mtDNA mutation burden in the complete tissue homogenate is quantified. Using such dynamical data and using distribution function of mtDNA mutation frequency compounded with the uncertainty arising from the mtDNA mutation quantification protocol, the maximum likelihood parameter estimation framework proposed in Chapter 7 can be exploited to estimate the rate constants of mtDNA turnover process.

In the context of mtDNA mutation clonal expansion dynamics, there are two aspects that are intricately related to each other and are involved in the mtDNA mutagenesis process: a.) initial formation and b.) subsequent mutation accumulation. While the present dissertation focuses on the accumulation dynamics related to single type of mutation (Chapter 3 and Chapter 4), understanding the mechanisms of mtDNA mutation formation and its subsequent role in mtDNA mutation accumulation and clonal expansion dynamics is important, especially since obtaining the experimental data of the complete mtDNA

deletion spectrum is technically challenging. In the context of mtDNA deletions, the formation of deletions have been hypothesized to be influenced by sequence motifs. Study of mtDNA sequence motifs in the context of distribution of direct repeats, stemloops and tRNA structures, which may be directly involved in mtDNA deletion formation can provide data on the frequency of occurrence of different types of mtDNA deletions, some of which are currently under investigation in our group. With this information, the *de novo* mtDNA deletions arising during the mtDNA turnover process (Chapter 4) can be sampled from the predicted *in silico* mtDNA deletion spectrum. The *in silico* mouse model simulated this way will provide information about the intra-cellular heteroplasmy and tissue distribution of different mtDNA deletions in the cells. As additional model validation measure, the *in silico* mtDNA deletion frequencies in post-mitotic cells obtained from the stochastic simulation of mtDNA turnover process, using the initial distribution function of deletion breakpoint can be compared with the experimental data on distribution frequencies of different deletions found in post mitotic tissues.

Based on the beneficial effects of aerobic exercise on the general muscle mass development and function, it has been suggested as a therapeutic intervention for sarcopenia and mitochondrial myopathies. There are two types of exercise training based on the initiating stimulus: i.) Endurance and, ii.) Resistance training [407]. Endurance training involves low intensity physical endurance for longer time. Endurance training has generally been associated with enhanced mitochondrial biogenesis, ATP production and oxygen consumption. Beneficial effects of endurance training on sarcopenia and mitochondrial disorders is not clearly understood and the extension of the multi-scale hybrid model of sarcopenic skeletal muscle fiber developed in this thesis (Chapter 7) can be adapted to determine the influence of endurance training on the mitochondrial mutation heteroplasmy and associated progression of sarcopenia. Resistance training involves high intensity physical bouts of stimulus for shorter duration. Resistance training has been

suggested to cause muscle hypertrophy by recruiting muscle satellite cells [408]. These satellite cells are generally devoid of mtDNA mutations, and their fusion with the existing myofibers results in shift in mtDNA mutation heteroplasmy levels. This shift of mtDNA mutation heteroplasmy is termed as gene shifting [408, 409]. Gene shift is considered as a promising intervention strategy to treat myopathies [409], and may also be possible intervention for sarcopenia. The multi-scale hybrid model developed in this thesis can be modified to include specific features of gene shifting process by modeling the satellite cell recruitment. By developing these models and subjecting them to parametric perturbation analysis can provide insights about possible clinical or pharmacological interventions to treat sarcopenia/mitochondrial myopathies.

Spatial aspects of biochemical reactions are important in modeling the intra-cellular processes [410]. Diffusion process is important, because the interacting components do not diffuse sufficiently fast to make the system well-stirred between the individual reactions or interactions. Stochasticity arises because the number of reacting components within the diffusion range is often small and this gives rise to probabilistic and nonlinear nature of molecular interactions of the reacting components in the system. The hybrid model of skeletal muscle fiber (Chapter 7), considers a deterministic framework for simulating the mtDNA turnover process in skeletal muscle fibers, during the organism's postnatal life. However, in the case mtDNA diffusion in the skeletal muscle fiber, mtDNA diffusivity is very low ($D_{eff} = 1.1 \times 10^{-15} \text{ m}^2/\text{s}$) and thus the stochastic origin and accumulation dynamics of mtDNA mutations during the post natal stage is likely to play important role in the development and progression of mtDNA clonal expansion process. The spatial heterogeneity is generally handled using a spatially-discretized reaction-diffusion models [228]. The reaction-diffusion master equation (RDME) is generally used to obtain exact solution of biochemical reaction systems where diffusion cannot be neglected and conventionally the method averages out the kinetics at microscopic lengths and time scales

[228, 411, 412]. However, the computational complexity associated with the simulation related to RDME is enormous for even very simple systems [376, 413]. Thus, a direct extension of this method for simulating the reaction-diffusion physics in skeletal muscle fiber might not be computationally feasible and reliance on hybrid method might be warranted, where approximation like Poisson process and Langevin formulation can be used instead of the exact method.

References

- [1] D. Harman. Aging: a theory based on free radical and radiation chemistry. *J Gerontol*, 11(3):298–300, 1956.
- [2] T. B. Kirkwood and S. N. Austad. Why do we age? *Nature*, 408(6809):233–8, 2000.
- [3] E. Bell. Evolutionary and nonevolutionary theories of senescence. *American Naturalist*, 124:600–603, 1984.
- [4] T. B. Kirkwood. Evolution of ageing. *Nature*, 270(5635):301–4, 1977.
- [5] T. B. L. Kirkwood and R. Holiday. The evolution of ageing and longevity. *Proc R Soc Lond B*, 205:513–546, 1979.
- [6] T. B. L. Kirkwood. Understanding the odd science of aging. *Cell*, 120(4):437–447, 2005.
- [7] Z. A. Medvedev. An attempt at a rational classification of theories of ageing. *Biol Rev Camb Philos Soc*, 65(3):375–98, 1990.
- [8] A. Kowald and T. B. Kirkwood. A network theory of ageing: the interactions of defective mitochondria, aberrant proteins, free radicals and scavengers in the ageing process. *Mutat Res*, 316(5-6):209–36, 1996.
- [9] T. B. Kirkwood, R. J. Boys, C. S. Gillespie, and *et al.* Towards an e-biology of ageing: integrating theory and data. *Nat Rev Mol Cell Biol*, 4(3):243–9, 2003.
- [10] C. Franceschi, S. Valensin, M. Bonafe, and *et al.* The network and the remodeling theories of aging: historical background and new perspectives. *Exp Gerontol*, 35(6-7):879–96, 2000.
- [11] T. B. Kirkwood. A systematic look at an old problem. *Nature*, 451(7179):644–7, 2008.
- [12] A. Plisko and B. A. Gilchrest. Growth factor responsiveness of cultured human fibroblasts declines with age. *J Gerontol*, 38(5):513–8, 1983.
- [13] J. Campisi, S. H. Kim, C. S. Lim, and M. Rubio. Cellular senescence, cancer and aging: the telomere connection. *Exp Gerontol*, 36(10):1619–37, 2001.
- [14] J. Campisi. Aging, tumor suppression and cancer: high wire-act! *Mech Ageing Dev*, 126(1):51–8, 2005.
- [15] D. M. Baird and D. Kipling. The extent and significance of telomere loss with age. *Ann N Y Acad Sci*, 1019:265–8, 2004.
- [16] T. von Zglinicki. Oxidative stress shortens telomeres. *Trends Biochem Sci*, 27(7):339–44, 2002.
- [17] S. Goldstein. Replicative senescence: the human fibroblast comes of age. *Science*, 249(4973):1129–33, 1990.
- [18] A. Terman and U. T. Brunk. Aging as a catabolic malfunction. *Int J Biochem Cell Biol*, 36(12):2365–75, 2004.

- [19] G. Carrard, A. L. Bulteau, I. Petropoulos, and B. Friguet. Impairment of proteasome structure and function in aging. *Int J Biochem Cell Biol*, 34(11):1461–74, 2002.
- [20] C. Soti and P. Csermely. Aging and molecular chaperones. *Exp Gerontol*, 38(10):1037–40, 2003.
- [21] B. M. Stanulis-Praeger. Cellular senescence revisited: a review. *Mech Ageing Dev*, 38(1):1–48, 1987.
- [22] M. P. Cohen and V. Yu-Wu. Age-related changes in non-enzymatic glycosylation of human basement membranes. *Exp Gerontol*, 18(6):461–9, 1983.
- [23] R. Bucala, P. Model, and A. Cerami. Modification of dna by reducing sugars: a possible mechanism for nucleic acid aging and age-related dysfunction in gene expression. *Proc Natl Acad Sci U S A*, 81(1):105–9, 1984.
- [24] G. O. Ivy, F. Schottler, J. Wenzel, and *et al.* Inhibitors of lysosomal enzymes: accumulation of lipofuscin-like dense bodies in the brain. *Science*, 226(4677):985–7, 1984.
- [25] D. C. Wallace. Mitochondrial DNA mutations and neuromuscular disease. *Trends Genet*, 5(1):9–13, 1989.
- [26] A. Herbst, J. W. Pak, D. McKenzie, and *et al.* Accumulation of mitochondrial DNA deletion mutations in aged muscle fibers: evidence for a causal role in muscle fiber loss. *J Gerontol A Biol Sci Med Sci*, 62(3):235–45, 2007.
- [27] A. Bender, K. J. Krishnan, C. M. Morris, and *et al.* High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nat Genet*, 38(5):515–517, 2006.
- [28] J. Wanagat, M. R. Wolff, and J. M. Aiken. Age-associated changes in function, structure and mitochondrial genetic and enzymatic abnormalities in the Fischer 344 × Brown Norway F(1) hybrid rat heart. *J Mol Cell Cardiol*, 34(1):17–28, 2002.
- [29] R. W. Taylor, M. J. Barron, G. M. Borthwick, and *et al.* Mitochondrial DNA mutations in human colonic crypt stem cells. *J Clin Invest*, 112:1351–1360, 2003.
- [30] D. Harman. The biologic clock: the mitochondria? *J Am Geriatr Soc*, 20(4):145–7, 1972.
- [31] R. M. Schwartz and M. O. Dayhoff. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science*, 199(4327):395–403, 1978.
- [32] B. Alberts. *Molecular biology of the cell*. Garland Science, New York, 2002. 4th.
- [33] N. G. Larsson. Somatic mitochondrial DNA mutations in mammalian aging. *Annu Rev Biochem*, 79:683–706, 2010.
- [34] H. Chen and D. C. Chan. Emerging functions of mammalian mitochondrial fusion and fission. *Hum Mol Genet*, 14 Spec No. 2:R283–9, 2005.
- [35] E. J. Masoro and S. N. Austad. *Handbook of the biology of aging*. Elsevier Academic Press, Amsterdam ; Boston, 2006. 6th.
- [36] G. Barja. Mitochondrial oxygen radical generation and leak: sites of production in states 4 and 3, organ specificity, and relation to aging and longevity. *J Bioenerg Biomembr*, 31(4):347–66, 1999.
- [37] G. Lenaz. The mitochondrial production of reactive oxygen species: mechanisms and implications in human pathology. *IUBMB Life*, 52(3-5):159–64, 2001.
- [38] I. Sipos, L. Tretter, and V. Adam-Vizi. Quantitative relationship between inhibition of respiratory complexes and formation of reactive oxygen species in isolated nerve terminals. *J Neurochem*, 84(1):112–8, 2003.

- [39] I. V. Grigolava, M. I. Ksenzenko, A. A. Konstantinob, and *et al.* [tiron as a spin-trap for superoxide radicals produced by the respiratory chain of submitochondrial particles]. *Biokhimiia*, 45(1):75–82, 1980.
- [40] G. Loschen, A. Azzi, C. Richter, and L. Flohe. Superoxide radicals as precursors of mitochondrial hydrogen peroxide. *FEBS Lett*, 42(1):68–72, 1974.
- [41] J. F. Turrens and A. Boveris. Generation of superoxide anion by the nadh dehydrogenase of bovine heart mitochondria. *Biochem J*, 191(2):421–7, 1980.
- [42] E. Cadenas, A. Boveris, C. I. Ragan, and A. O. Stoppani. Production of superoxide radicals and hydrogen peroxide by nadh-ubiquinone reductase and ubiquinol-cytochrome c reductase from beef-heart mitochondria. *Arch Biochem Biophys*, 180(2):248–57, 1977.
- [43] A. Herrero and G. Barja. Localization of the site of oxygen radical generation inside the complex i of heart and nonsynaptic brain mammalian mitochondria. *J Bioenerg Biomembr*, 32(6):609–15, 2000.
- [44] P. C. Hinkle, R. A. Butow, E. Racker, and B. Chance. Partial resolution of the enzymes catalyzing oxidative phosphorylation. xv. reverse electron transfer in the flavin-cytochrome beta region of the respiratory chain of beef heart submitochondrial particles. *J Biol Chem*, 242(22):5169–73, 1967.
- [45] K. Takeshige and S. Minakami. Nadh- and nadph-dependent formation of superoxide anions by bovine heart submitochondrial particles and nadh-ubiquinone reductase preparation. *Biochem J*, 180(1):129–35, 1979.
- [46] M. L. Genova, B. Ventura, G. Giuliano, and *et al.* The site of production of superoxide radical in mitochondrial complex i is not a bound ubisemiquinone but presumably iron-sulfur cluster n2. *FEBS Lett*, 505(3):364–8, 2001.
- [47] R. G. Hansford, B. A. Hogue, and V. Mildaziene. Dependence of h₂o₂ formation by rat heart mitochondria on substrate availability and donor age. *J Bioenerg Biomembr*, 29(1):89–95, 1997.
- [48] L. K. Kwong and R. S. Sohal. Substrate and site specificity of hydrogen peroxide generation in mouse mitochondria. *Arch Biochem Biophys*, 350(1):118–26, 1998.
- [49] Y. Kushnareva, A. N. Murphy, and A. Andreyev. Complex I-mediated reactive oxygen species generation: modulation by cytochrome c and NAD(P)⁺ oxidation-reduction state. *Biochem J*, 368(Pt 2):545–53, 2002.
- [50] A. A. Starkov and G. Fiskum. Regulation of brain mitochondrial h₂o₂ production by membrane potential and nad(p)h redox state. *J Neurochem*, 86(5):1101–7, 2003.
- [51] A. Boveris, N. Oshino, and B. Chance. The cellular production of hydrogen peroxide. *Biochem J*, 128(3):pp. 617–630, 1972.
- [52] A. Boveris, E. Cadenas, and A. O. Stoppani. Role of ubiquinone in the mitochondrial generation of hydrogen peroxide. *Biochem J*, 156(2):435–44, 1976.
- [53] B. L. Trumpower. The protonmotive q cycle. energy transduction by coupling of proton translocation to electron transfer by the cytochrome bc₁ complex. *J Biol Chem*, 265(20):11409–12, 1990.
- [54] D. Han, R. Canali, D. Rettori, and N. Kaplowitz. Effect of glutathione depletion on sites and topology of superoxide and hydrogen peroxide production in mitochondria. *Mol Pharmacol*, 64(5):1136–44, 2003.
- [55] K. B. Beckman and B. N. Ames. The free radical theory of aging matures. *Physiol Rev*, 78(2):547–81, 1998.

- [56] R. S. Sohal and R. Weindruch. Oxidative stress, caloric restriction, and aging. *Science*, 273(5271):59–63, 1996.
- [57] W. C. Burkans and M. Weinberger. DNA replication stress, genome instability and aging. *Nucleic Acids Res*, 35(22):7545–56, 2007.
- [58] K. J. Krishnan, A. K. Reeve, D. C. Samuels, and *et al.* What causes mitochondrial DNA deletions in human cells? *Nat Genet*, 40(3):275–9, 2008.
- [59] C. Richter. Oxidative damage to mitochondrial DNA and its relationship to ageing. *Int J Biochem Cell Biol*, 27(7):647–53, 1995.
- [60] Y. H. Wei. Mitochondrial DNA alterations as ageing-associated molecular events. *Mutat Res*, 275(3-6):145–55, 1992.
- [61] R. A. Gus'kova, II Ivanov, V. K. Kol'tover, and *et al.* [permeability of bilayer phospholipid membranes to superoxide oxygen radicals]. *Biokhimiia*, 49(5):758–66, 1984.
- [62] T. R. Pushpa-Rekha, A. L. Burdsall, L. M. Oleksa, and *et al.* Rat phospholipid-hydroperoxide glutathione peroxidase. cDNA cloning and identification of multiple transcription and translation start sites. *J Biol Chem*, 270(45):26993–9, 1995.
- [63] F. Antunes, A. Salvador, and R. E. Pinto. PhgpX and phospholipase a2/gpX: comparative importance on the reduction of hydroperoxides in rat liver mitochondria. *Free Radic Biol Med*, 19(5):669–77, 1995.
- [64] M. Lundberg, C. Johansson, J. Chandra, and *et al.* Cloning and expression of a novel human glutaredoxin (grx2) with mitochondrial and nuclear isoforms. *J Biol Chem*, 276(28):26269–75, 2001.
- [65] E. S. Arner and A. Holmgren. Physiological functions of thioredoxin and thioredoxin reductase. *Eur J Biochem*, 267(20):6102–9, 2000.
- [66] H. Z. Chae, S. W. Kang, and S. G. Rhee. Isoforms of mammalian peroxiredoxin that reduce peroxides in presence of thioredoxin. *Methods Enzymol*, 300:219–26, 1999.
- [67] S. Watabe, T. Hiroi, Y. Yamamoto, and *et al.* Sp-22 is a thioredoxin-dependent peroxide reductase in mitochondria. *Eur J Biochem*, 249(1):52–60, 1997.
- [68] T. S. Chang, C. S. Cho, S. Park, and *et al.* Peroxiredoxin iii, a mitochondrion-specific peroxidase, regulates apoptotic signaling by mitochondria. *J Biol Chem*, 279(40):41975–84, 2004.
- [69] A. Kowald. The mitochondrial theory of aging. *Biol Signals Recept*, 10(3-4):162–75, 2001.
- [70] M. D. Brand. Uncoupling to survive? the role of mitochondrial inefficiency in ageing. *Exp Gerontol*, 35(6-7):811–20, 2000.
- [71] D. G. Nicholls and D. G. Nicholls. Hamster brown-adipose-tissue mitochondria. the control of respiration and the proton electrochemical potential gradient by possible physiological effectors of the proton conductance of the inner membrane. *Eur J Biochem*, 49(3):573–83, 1974.
- [72] C. D. Nobes, G. C. Brown, P. N. Olive, and M. D. Brand. Non-ohmic proton conductance of the mitochondrial inner membrane in hepatocytes. *J Biol Chem*, 265(22):12903–9, 1990.
- [73] D. F. Rolfe, J. M. Newman, J. A. Buckingham, and *et al.* Contribution of mitochondrial proton leak to respiration rate in working skeletal muscle and liver and to smr. *Am J Physiol*, 276(3 Pt 1):C692–9, 1999.
- [74] D. F. Rolfe and M. D. Brand. Contribution of mitochondrial proton leak to skeletal muscle respiration and to standard metabolic rate. *Am J Physiol*, 271(4 Pt 1):C1380–9, 1996.
- [75] L. J. Sweetlove, A. Lytovchenko, M. Morgan, and *et al.* Mitochondrial uncoupling protein is required for efficient photosynthesis. *Proc Natl Acad Sci U S A*, 103(51):19587–92, 2006.

- [76] S. S. Korshunov, V. P. Skulachev, and A. A. Starkov. High protonic potential actuates a mechanism of production of reactive oxygen species in mitochondria. *FEBS Lett*, 416(1):15–8, 1997.
- [77] A. J. Lambert and M. D. Brand. Superoxide production by nadh:ubiquinone oxidoreductase (complex i) depends on the ph gradient across the mitochondrial inner membrane. *Biochem J*, 382(Pt 2):511–7, 2004.
- [78] M. D. Brand, J. L. Pakay, A. Ocloo, and *et al.* The basal proton conductance of mitochondria depends on adenine nucleotide translocase content. *Biochem J*, 392(Pt 2):353–62, 2005.
- [79] G. M. Heaton, R. J. Wagenvoord, Jr. Kemp, A., and D. G. Nicholls. Brown-adipose-tissue mitochondria: photoaffinity labelling of the regulatory site of energy dissipation. *Eur J Biochem*, 82(2):515–21, 1978.
- [80] B. Cannon and J. Nedergaard. Brown adipose tissue: function and physiological significance. *Physiol Rev*, 84(1):277–359, 2004.
- [81] H. M. Feldmann, V. Golozoubova, B. Cannon, and J. Nedergaard. Ucp1 ablation induces obesity and abolishes diet-induced thermogenesis in mice exempt from thermal stress by living at thermoneutrality. *Cell Metab*, 9(2):203–9, 2009.
- [82] A. Dlaskova, K. J. Clarke, and R. K. Porter. The role of ucp 1 in production of reactive oxygen species by mitochondria isolated from brown adipose tissue. *Biochim Biophys Acta*, 1797(8):1470–6, 2010.
- [83] R. Oelkrug, M. Kutschke, C. W. Meyer, and *et al.* Uncoupling protein 1 decreases superoxide production in brown adipose tissue mitochondria. *J Biol Chem*, 285(29):21961–8, 2010.
- [84] V. Azzu, C. Affourtit, E. P. Breen, and *et al.* Dynamic regulation of uncoupling protein 2 content in ins-1e insulinoma cells. *Biochim Biophys Acta*, 1777(10):1378–83, 2008.
- [85] C. Fleury, M. Neverova, S. Collins, and *et al.* Uncoupling protein-2: a novel gene linked to obesity and hyperinsulinemia. *Nat Genet*, 15(3):269–72, 1997.
- [86] E. Aguirre and S. Cadenas. Gdp and carboxyatractylate inhibit 4-hydroxynonenal-activated proton conductance to differing degrees in mitochondria from skeletal muscle and heart. *Biochim Biophys Acta*, 1797(10):1716–26, 2010.
- [87] O. Boss, S. Samec, A. Paoloni-Giacobino, and *et al.* Uncoupling protein-3: a new member of the mitochondrial carrier family with tissue-specific expression. *FEBS Lett*, 408(1):39–42, 1997.
- [88] P. Jezek, A. D. Costa, and A. E. Vercesi. Reconstituted plant uncoupling mitochondrial protein allows for proton translocation via fatty acid cycling mechanism. *J Biol Chem*, 272(39):24272–8, 1997.
- [89] M. Laloi, M. Klein, J. W. Riesmeier, and *et al.* A plant cold-induced uncoupling protein. *Nature*, 389(6647):135–6, 1997.
- [90] S. Raimbault, S. Dridi, F. Denjean, and *et al.* An uncoupling protein homologue putatively involved in facultative muscle thermogenesis in birds. *Biochem J*, 353(Pt 3):441–4, 2001.
- [91] C. R. Vianna, T. Hagen, C. Y. Zhang, and *et al.* Cloning and functional characterization of an uncoupling protein homolog in hummingbirds. *Physiol Genomics*, 5(3):137–45, 2001.
- [92] J. Diao, E. M. Allister, V. Koshkin, and *et al.* Ucp2 is highly expressed in pancreatic alpha-cells and influences secretion and survival. *Proc Natl Acad Sci U S A*, 105(33):12057–62, 2008.
- [93] Y. Emre, C. Hurtaud, M. Karaca, and *et al.* Role of uncoupling protein ucp2 in cell-mediated immunity: how macrophage-mediated insulinitis is accelerated in a model of autoimmune diabetes. *Proc Natl Acad Sci U S A*, 104(48):19085–90, 2007.

- [94] L. J. Toime and M. D. Brand. Uncoupling protein-3 lowers reactive oxygen species production in isolated mitochondria. *Free Radic Biol Med*, 49(4):606–11, 2010.
- [95] N. Garrido, L. Griparic, E. Jokitalo, and *et al.* Composition and dynamics of human mitochondrial nucleoids. *Mol Biol Cell*, 14(4):1583–96, 2003.
- [96] C. T. Moraes, L. Kenyon, and H. Hao. Mechanisms of human mitochondrial DNA maintenance: the determining role of primary sequence and length over function. *Mol Biol Cell*, 10(10):3345–56, 1999.
- [97] C. T. Moraes. What regulates mitochondrial DNA copy number in animal cells? *Trends Genet*, 17(4):199–205, 2001.
- [98] B. A. Kaufman, S. M. Newman, R. L. Hallberg, and *et al.* *in organello* formaldehyde crosslinking of proteins to mtDNA: identification of bifunctional proteins. *Proc Natl Acad Sci U S A*, 97(14):7772–7, 2000.
- [99] A. F. Davis and D. A. Clayton. *In situ* localization of mitochondrial DNA replication in intact mammalian cells. *J Cell Biol*, 135(4):883–93, 1996.
- [100] D. Ojala, J. Montoya, and G. Attardi. tRNA punctuation model of RNA processing in human mitochondria. *Nature*, 290:470–474, 1981.
- [101] T. Strachan and A. P. Read. *Human molecular genetics*. Wiley-Liss, New York ; Chichester [England], 1999. 2nd.
- [102] B. Halliwell and O. I. Aruoma. DNA damage by oxygen-derived species. its mechanism and measurement in mammalian systems. *FEBS Lett*, 281(1-2):9–19, 1991.
- [103] M. Y. Yang, M. Bowmaker, A. Reyes, and *et al.* Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell*, 111(4):495–505, 2002.
- [104] M. Bowmaker, M. Y. Yang, T. Yasukawa, and *et al.* Mammalian mitochondrial DNA replicates bidirectionally from an initiation zone. *J Biol Chem*, 278:pp. 50961–50969, 2003.
- [105] G. S. Shadel and D. A. Clayton. Mitochondrial DNA maintenance in vertebrates. *Annu Rev Biochem*, 66:409–35, 1997.
- [106] G. Barja. Free radicals and aging. *Trends Neurosci*, 27(10):595–600, 2004.
- [107] R. J. Wiesner, G. Zsurka, and W. S. Kunz. Mitochondrial DNA damage and the aging process: facts and imaginations. *Free Radic Res*, 40(12):1284–94, 2006.
- [108] D. C. Wallace. Diseases of the mitochondrial DNA. *Ann Rev Biochem*, 61:1175–1212, 1992.
- [109] M. Corral-Debrinski, T. Horton, M. T. Lott, and *et al.* Mitochondrial DNA deletions in human brain: regional variability and increase with advanced age. *Nat Genet*, 2(4):324–9, 1992.
- [110] G. A. Cortopassi and N. Arnheim. Detection of a specific mitochondrial DNA deletion in tissues of older humans. *Nucleic Acids Res*, 18(23):6927–33, 1990.
- [111] J. Muller-Hocker. Cytochrome-c-oxidase deficient cardiomyocytes in the human heart—an age-related phenomenon. a histochemical ultracytochemical study. *Am J Pathol*, 134(5):1167–73, 1989.
- [112] B. Halliwell. Oxygen and nitrogen are pro-carcinogens. damage to DNA by reactive oxygen, chlorine and nitrogen species: measurement, mechanism and the effects of nutrition. *Mutat Res*, 443(1-2):37–52, 1999.
- [113] B. Halliwell and J. M. C. Gutteridge. *Free radicals in biology and medicine*. Clarendon Press ; Oxford University Press, Oxford New York, 1999. 3rd.

- [114] Y. Michikawa, F. Mazzucchelli, N. Bresolin, and *et al.* Aging-dependent large accumulation of point mutations in human mtDNA control region for accumulation. *Science*, 286(5440):774–779, 1999.
- [115] M. Khaidakov, R. H. Heflich, M. G. Manjanatha, and *et al.* Accumulation of point mutations in mitochondrial DNA of aging mice. *Mutat Res*, 526(1-2):1–7, 2003.
- [116] M. Vermulst, J. H. Bielas, G. C. Kujoth, and *et al.* Mitochondrial point mutations do not limit the natural lifespan of mice. *Nat Genet*, 39(4):540–3, 2007.
- [117] A. Trifunovic. Mitochondrial DNA and ageing. *Biochim Biophys Acta*, 1757:611–617, 2006.
- [118] K. Khrapko, E. Nekhaeva, Y. Kraytsberg, and W. Kunz. Clonal expansions of mitochondrial genomes: implications for *in vivo* mutational spectra. *Mutat Res*, 522(1-2):13–9, 2003.
- [119] D. C. Wallace. Mitochondrial diseases in man and mouse. *Science*, 283:1482–1488, 1999.
- [120] D. C. Wallace. Mitochondrial DNA sequence variation in human evolution and disease. *Proc Natl Acad Sci USA*, 91:8739–8746, 1994.
- [121] I. J. Holt, A. E. Harding, and J. A. Morgan-Hughes. Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature*, 331(6158):717–9, 1988.
- [122] R. W. Taylor and D. M. Turnbull. Mitochondrial DNA mutations in human disease. *Nat Rev Genet*, 6(5):389–402, 2005.
- [123] S. DiMauro and D. C. Wallace. *Mitochondrial DNA in human pathology*. Raven Press, New York, 1993.
- [124] B. S. Mandavilli, J. H. Santos, and B. V. Houten. Mitochondrial DNA repair and aging. *Mutat Res*, 509:127–151, 2002.
- [125] S. DiMauro, M. Hirano, and E. A. Schon. *Mitochondrial Medicine*. Informa Healthcare, 2006. 1.
- [126] J. Miquel, A. C. Economos, J. Fleming, and J. E. Jr. Johnson. Mitochondrial role in cell aging. *Exp Gerontol*, 15(6):575–91, 1980.
- [127] J. Gruber, S. Schaffer, and B. Halliwell. The mitochondrial free radical theory of ageing—where do we stand? *Front Biosci*, 13:6554–79, 2008.
- [128] T. Lu, Y. Pan, S. Y. Kao, and *et al.* Gene regulation and DNA damage in the ageing human brain. *Nature*, 429(6994):883–91, 2004.
- [129] A. Panov, S. Dikalov, N. Shalbuyeva, and *et al.* Rotenone model of Parkinson disease: multiple brain mitochondria dysfunctions after short term systemic rotenone intoxication. *J Biol Chem*, 280(51):42026–35, 2005.
- [130] G. C. Kujoth, A. Hiona, T. D. Pugh, and *et al.* Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science*, 309(5733):481–4, 2005.
- [131] A. Trifunovic, A. Wredenberg, M. Falkenberg, and *et al.* Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature*, 429(6990):417–23, 2004.
- [132] A. Trifunovic, A. Hansson, A. Wredenberg, and *et al.* Somatic mtDNA mutations cause aging phenotypes without affecting reactive oxygen species production. *Proc Natl Acad Sci U S A*, 102(50):17993–8, 2005.
- [133] G. C. Kujoth, P. C. Bradshaw, S. Haroon, and T. A. Prolla. The role of mitochondrial DNA mutations in mammalian aging. *PLoS Genet*, 3(2):e24, 2007.
- [134] M. I. Ekstrand, M. Terzioglu, D. Galter, and *et al.* Progressive parkinsonism in mice with respiratory-chain-deficient dopamine neurons. *Proc Natl Acad Sci U S A*, 104(4):1325–30, 2007.

- [135] S. E. Schriener, N. J. Linford, G. M. Martin, and *et al.* Extension of murine life span by overexpression of catalase targeted to mitochondria. *Science*, 308(5730):1909–1911, 2005.
- [136] J. L. Mott, D. Zhang, P. L. Farrar, and *et al.* Low frequencies of mitochondrial DNA mutations cause cardiac disease in the mouse. *Ann N Y Acad Sci*, 893:353–7, 1999.
- [137] D. Zhang, J. L. Mott, S. W. Chang, and *et al.* Construction of transgenic mice with tissue-specific acceleration of mitochondrial DNA mutagenesis. *Genomics*, 69(2):151–61, 2000.
- [138] J. N. Spelbrink, F. Y. Li, V. Tiranti, and *et al.* Human mitochondrial DNA deletions associated with mutations in the gene encoding TWINKLE, a phage T7 gene 4-like protein localized in mitochondria. *Nat Genet*, 28(3):223–31, 2001.
- [139] H. Tynynmaa, K. P. Mjosund, S. Wanrooij, and *et al.* Mutant mitochondrial helicase TWINKLE causes multiple mtDNA deletions and a late-onset mitochondrial disease in mice. *Proc Natl Acad Sci U S A*, 102(49):17687–92, 2005.
- [140] K. Inoue, K. Nakada, A. Ogura, and *et al.* Generation of mice with mitochondrial dysfunction by introducing mouse mtDNA carrying a deletion into zygotes. *Nat Genet*, 26(2):176–81, 2000.
- [141] J. H. Bielas and L. A. Loeb. Quantification of random genomic mutations. *Nat Methods*, 2(4):285–290, 2005.
- [142] L. Piko, A. J. Hougham, and K. J. Bulpitt. Studies of sequence heterogeneity of mitochondrial DNA from rat and mouse tissues: evidence for an increased frequency of deletions/additions with aging. *Mech Ageing Dev*, 43(3):279–93, 1988.
- [143] Y. Goto, I. Nonaka, and S. Horai. A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature*, 348(6302):651–3, 1990.
- [144] J. M. Shoffner, M. T. Lott, A. M. Lezza, and *et al.* Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial dna tRNA(Lys) mutation. *Cell*, 61(6):931–7, 1990.
- [145] R. Rossignol, B. Faustin, C. Rocher, and *et al.* Mitochondrial threshold effects. *Biochem J*, 370(Pt 3):751–62, 2003.
- [146] J. Hayashi, S. Ohta, A. Kikuchi, and *et al.* Introduction of disease-related mitochondrial DNA deletions into HeLa cells lacking mitochondrial DNA results in mitochondrial dysfunction. *Proc Natl Acad Sci U S A*, 88(23):10614–8, 1991.
- [147] A. Chomyn, A. Martinuzzi, M. Yoneda, and *et al.* 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 U S A*, 89(10):4221–5, 1992.
- [148] D. A. Cottrell, E. L. Blakely, M. A. Johnson, and *et al.* Mitochondrial enzyme-deficient hippocampal neurons and choroidal cells in AD. *Neurology*, 57(2):260–4, 2001.
- [149] D. A. Cottrell, E. L. Blakely, M. A. Johnson, and *et al.* Cytochrome c oxidase deficient cells accumulate in the hippocampus and choroid plexus with age. *Neurobiol Aging*, 22(2):265–72, 2001.
- [150] E. Bua, J. Johnson, A. Herbst, and *et al.* Mitochondrial DNA-deletion mutations accumulate intracellularly to detrimental levels in aged human skeletal muscle fibers. *Am J Hum Genet*, 79(3):469–80, 2006.
- [151] J. Wanagat, Z. Cao, P. Pathare, and J. M. Aiken. 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):322–32, 2001.

- [152] J. W. Pak, A. Herbst, E. Bua, and *et al.* Mitochondrial DNA mutations as a fundamental mechanism in physiological declines associated with aging. *Aging Cell*, 2(1):1–7, 2003.
- [153] G. A. Cortopassi and N. Arnheim. Using the polymerase chain reaction to estimate mutation frequencies and rates in human cells. *Mutat Res*, 277(3):239–49, 1992.
- [154] Y. Kraytsberg, E. Kudryavtseva, A. C. McKee, and *et al.* Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat Genet*, 38(5):518–20, 2006.
- [155] H. Kitano. Computational systems biology. *Nature*, 420(6912):206–10, 2002.
- [156] H. Kitano. Systems biology: a brief overview. *Science*, 295(5560):1662–4, 2002.
- [157] W. R. Ashby, A. L. Hodgkin, and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117(4):500–44, 1952.
- [158] R. Rosen. *Anticipatory systems : philosophical, mathematical, and methodological foundations*. Pergamon Press, Oxford, England ; New York, 1985. 1st.
- [159] N. Wiener. *Cybernetics; or, Control and communication in the animal and the machine*. J. Wiley, New York,, 1948.
- [160] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117(4):500–44, 1952.
- [161] N. Le Novere. The long journey to a systems biology of neuronal function. *BMC Syst Biol*, 1:28, 2007.
- [162] D. Noble. Cardiac action and pacemaker potentials based on the Hodgkin-Huxley equations. *Nature*, 188:495–7, 1960.
- [163] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, 3:318–356, 1960.
- [164] E. O. Voit. *Computational analysis of biochemical systems : a practical guide for biochemists and molecular biologists*. Cambridge University Press, Cambridge ; New York, 2000.
- [165] M. A. Savageau, E. O. Voit, and D. H. Irvine. Biochemical systems theory and metabolic control theory: 1. fundamental similarities and differences. *Math Biosci*, 86:127–145, 1987.
- [166] O. Wolkenhauer. Systems biology: the reincarnation of systems theory applied in biology? *Brief Bioinform*, 2(3):258–70, 2001.
- [167] R. N. Stuart and E. W. Branscomb. Quantitative theory of *in vivo lac* regulation: significance of repressor packaging. I. equilibrium considerations. *J Theor Biol*, 31(2):313–329, 1971.
- [168] J. M. Pedraza and A. van Oudenaarden. Noise propagation in gene networks. *Science*, 307:1965–1969, 2005.
- [169] M. Kaern, T. C. Elston, W. J. Blake, and J. J. Collins. Stochasticity in gene expression: From theories to phenotypes. *Nat Rev Genet*, 6:451–464, 2005.
- [170] J. Paulsson. Models of stochastic gene expression. *Physics of Life*, 2:157–175, 2005.
- [171] J. Paulsson and M. Ehrenberg. Noise in a minimal regulatory network: plasmid copy number control. *J Rev Biophys*, 34(1):1–59, 2001.
- [172] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403:335–338, 2000.

- [173] E. M. Ozbudak, M. Thattai, H. N. Lim, and *et al.* Multistability in the lactose utilization network of *escherichia coli*. *Nature*, 427(6976):737–40, 2004.
- [174] D. A. McQuarrie. Stochastic approach to chemical kinetics. *J Appl Probab*, 4 (3):413–478, 1967.
- [175] D. T. Gillespie. *Markov Processes: An Introduction for Physical Scientists*. Academic Press, San Diego, 1991.
- [176] P. S. Swain, M. B. Elowitz, and E. D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA*, 99(20):12795–12800, 2002.
- [177] W. J. Blake, M. Kaern, C. R. Cantor, and J. J. Collins. Noise in eukaryotic gene expression. *Nature*, 422(6932):633–7, 2003.
- [178] H. H. McAdams and A. Arkin. It’s a noisy business! genetic regulation at the nanomolar scale. *Trends Genet*, 15(2):65–9, 1999.
- [179] H. H. McAdams and A. P. Arkin. Stochastic mechanisms in gene expression. *Proc Natl Acad Sci USA*, 94:814–819, 1999.
- [180] E. Cevenini, L. Invidia, F. Lescai, and *et al.* Human models of aging and longevity. *Expert Opin Biol Ther*, 8(9):1393–405, 2008.
- [181] G. A. Cortopassi and A. Wong. Mitochondria in organismal aging and degeneration. *Biochim Biophys Acta*, 1410(2):183–93, 1999.
- [182] G. A. Cortopassi, D. Shibata, N. W. Soong, and N. Arnheim. A pattern of accumulation of a somatic deletion of mitochondrial DNA in aging human tissues. *Proc Natl Acad Sci U S A*, 89(16):7370–4, 1992.
- [183] M. F. Alexeyev, S. P. LeDoux, and G. L. Wilson. Mitochondrial DNA and aging. *Clin Sci*, 107:355–364, 2004.
- [184] M. Mishto, E. Bellavista, A. Santoro, and *et al.* Immunoproteasome and LMP2 polymorphism in aged and Alzheimer’s disease brains. *Neurobiol Aging*, 27(1):54–66, 2006.
- [185] L. A. Herndon, P. J. Schmeissner, J. M. Dudaronek, and *et al.* Stochastic and genetic factors influence tissue-specific decline in ageing *c. elegans*. *Nature*, 419(6909):808–814, 2002.
- [186] J. W. Vaupel, J. R. Carey, K. Christensen, and *et al.* Biodemographic trajectories of longevity. *Science*, 280(5365):855–860, 1998.
- [187] A. Kriete, B. A. Sokhansanj, D. L. Coppock, and G. B. West. Systems approaches to the networks of aging. *Ageing Res Rev*, 5(4):434–48, 2006.
- [188] D. Gems and L. Partridge. Insulin/IGF signalling and ageing: seeing the bigger picture. *Curr Opin Genet Dev*, 11(3):287–92, 2001.
- [189] G. J. Lithgow. Aging mechanisms from nematodes to mammals. *Nutrition*, 14(6):522–4, 1998.
- [190] M. Rincon, R. Muzumdar, G. Atzmon, and N. Barzilai. The paradox of the insulin/IGF-1 signaling pathway in longevity. *Mech Ageing Dev*, 125(6):397–403, 2004.
- [191] O. Arino, M. Kimmel, and G. F. Webb. Mathematical modeling of the loss of telomere sequences. *J Theor Biol*, 177(1):45–57, 1995.
- [192] I. Rubelj and Z. Vondracek. Stochastic mechanism of cellular aging—abrupt telomere shortening as a model for stochastic nature of cellular aging. *J Theor Biol*, 197(4):425–38, 1999.
- [193] C. J. Proctor, C. Soti, R. J. Boys, and *et al.* Modelling the actions of chaperones and their role in ageing. *Mech Ageing Dev*, 126(1):119–31, 2005.

- [194] T. von Zglinicki, R. Pilger, and N. Sitte. Accumulation of single-strand breaks is the major cause of telomere shortening in human fibroblasts. *Free Radic Biol Med*, 28(1):64–74, 2000.
- [195] P. J. Lindop and J. Rotblat. Shortening of life and causes of death in mice exposed to a single whole-body dose of radiation. *Nature*, 189:645–8, 1961.
- [196] L. Szilard. On the nature of the aging process. *Proc Natl Acad Sci U S A*, 45(1):30–45, 1959.
- [197] R. Holliday and T. B. Kirkwood. Predictions of the somatic mutation and mortalization theories of cellular ageing are contrary to experimental observations. *J Theor Biol*, 93(3):627–42, 1981.
- [198] T. B. Kirkwood and C. J. Proctor. Somatic mutations and ageing *in silico*. *Mech Ageing Dev*, 124(1):85–92, 2003.
- [199] A. Kowald and T. B. Kirkwood. Mitochondrial mutations, cellular instability and ageing: modelling the population dynamics of mitochondria. *Mutat Res*, 295(3):93–103, 1993.
- [200] A. Kowald and T. B. L. Kirkwood. Towards a network theory of aging: A model combining the free radical theory and the protein error theory. *J Theor Biol*, 168:75–94, 1994.
- [201] A. Kowald and T. B. Kirkwood. Accumulation of defective mitochondria through delayed degradation of damaged organelles and its possible role in the ageing of post-mitotic and dividing cells. *J Theor Biol*, 202(2):145–60, 2000.
- [202] A. D. de Grey. A proposed refinement of the mitochondrial free radical theory of aging. *Bioessays*, 19(2):161–6, 1997.
- [203] H. Chen, S. A. Detmer, A. J. Ewald, and *et al.* Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. *J Cell Biol*, 160(2):189–200, 2003.
- [204] F. J. Iborra, H. Kimura, and P. R. Cook. The functional organization of mitochondrial genomes in human cells. *BMC Biol*, 2:9, 2004.
- [205] K. Nakada, K. Inoue, T. Ono, and *et al.* Inter-mitochondrial complementation: Mitochondria-specific system preventing mice from expression of disease phenotypes by mutant mtDNA. *Nat Med*, 7(8):934–940, 2001.
- [206] L. Mao, C. Zabel, M. A. Wacker, and *et al.* Estimation of the mtDNA mutation rate in aging mice by proteome analysis and mathematical modeling. *Exp Gerontol*, 41(1):11–24, 2006.
- [207] J. Muller-Hocker. Cytochrome c oxidase deficient fibres in the limb muscle and diaphragm of man without muscular disease: an age-related alteration. *J Neurol Sci*, 100(1-2):14–21, 1990.
- [208] G. J. Capps, D. C. Samuels, and P. F. Chinnery. A model of the nuclear control of mitochondrial DNA replication. *J Theor Biol*, 221(4):565–83, 2003.
- [209] P. F. Chinnery and D. C. Samuels. Relaxed replication of mtDNA: A model with implications for the expression of disease. *Am J Hum Genet*, 64(4):1158–65, 1999.
- [210] J. L. Elson, D. C. Samuels, D. M. Turnbull, and P. F. Chinnery. Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age. *Am J Hum Genet*, 68:802–806, 2001.
- [211] E. A. Shoubridge, G. Karpati, and K. E. Hastings. Deletion mutants are functionally dominant over wild-type mitochondrial genomes in skeletal muscle fiber segments in mitochondrial disease. *Cell*, 62(1):43–9, 1990.
- [212] A. Kowald, M. Jendrach, S. Pohl, and *et al.* On the relevance of mitochondrial fusions for the accumulation of mitochondrial deletion mutants: a modelling study. *Aging Cell*, 4(5):273–83, 2005.

- [213] I. J. Holt, H. E. Lorimer, and H. T. Jacobs. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell*, 100(5):515–24, 2000.
- [214] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J Phys Chem*, 81:2340–2361, 1977.
- [215] R. Shenkar, W. Navidi, S. Tavaré, and *et al.* The mutation rate of the human mtDNA deletion mtDNA4977. *Am J Hum Genet*, 59(4):772–80, 1996.
- [216] Z. Cao, J. Wanagat, S. H. McKiernan, and J. M. Aiken. Mitochondrial DNA deletion mutations are concomitant with ragged red regions of individual, aged muscle fibers: analysis by laser-capture microdissection. *Nucleic Acids Res*, 29(21):4502–8, 2001.
- [217] E. Bua, S. H. McKiernan, and J. M. Aiken. Calorie restriction limits the generation but not the progression of mitochondrial abnormalities in aging skeletal muscle. *FASEB J*, 18(3):582–4, 2004.
- [218] N. J. Gross, G. S. Getz, and M. Rabionowitz. Apparent turnover of mitochondrial Deoxyribonucleic acid and mitochondrial phospholipids in the tissues of the rat. *J Biol Chem*, 244(6):1552–1562, 1969.
- [219] H. A. Collier, K. Khrapko, N. D. Bodyak, and *et al.* High frequency of homoplasmic mitochondrial DNA mutations in human tumors can be explained without selection. *Nat Genet*, 28(2):147–50, 2001.
- [220] K. B. Beckman and B. N. Ames. Endogenous oxidative damage of mtDNA. *Mutat Res*, 424(1-2):51–8, 1999.
- [221] ESCODD. Inter-laboratory validation of procedures for measuring 8-oxo-7,8-dihydroguanine/8-oxo-7,8-dihydro-2'-deoxyguanosine in DNA. *Free Radic Res*, 36(3):239–245, 2002.
- [222] M. L. Hamilton, H. Van Remmen, J. A. Drake, and *et al.* Does oxidative damage to DNA increase with age? *Proc Natl Acad Sci U S A*, 98(18):10469–74, 2001.
- [223] M. Calleja, P. Pena, C. Ugalde, and *et al.* Mitochondrial DNA remains intact during *Drosophila* aging, but the levels of mitochondrial transcripts are significantly reduced. *J Biol Chem*, 268(25):18891–7, 1993.
- [224] Y. Kraytsberg, A. Nicholas, P. Caro, and K. Khrapko. Single molecule PCR in mtDNA mutational analysis: Genuine mutations vs. damage bypass-derived artifacts. *Methods*, 46(4):269–73, 2008.
- [225] Y. Kraytsberg and K. Khrapko. Single-molecule PCR: an artifact-free PCR approach for the analysis of somatic mutations. *Expert Rev Mol Diagn*, 5(5):809–15, 2005.
- [226] M. Vermulst, J. H. Bielas, and L. A. Loeb. Quantification of random mutations in the mitochondrial genome. *Methods*, 46(4):263–8, 2008.
- [227] D. Harman. Free radical theory of aging: dietary implication. *Am J Clin Nutr*, 25:839–843, 1972.
- [228] C. W. Gardiner. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences (Springer Series in Synergetics)*. Springer, Berlin, 2004.
- [229] D. T. Gillespie. A rigorous derivation of the Chemical Master Equation. *Physica A*, 188:404–425, 1992.
- [230] R. J. Wiesner, J. C. Ruegg, and I. Morano. Counting target molecules by exponential polymerase chain reaction: copy number of mitochondrial DNA in rat tissue. *Biochem Biophys Res Co*, 183(2):553–559, 1992.

- [231] Z. Liu and R. A. Butow. Mitochondrial retrograde signaling. *Annu Rev Genet*, 40:159–185, 2006.
- [232] D. A. Clayton. Replication of animal mitochondrial DNA. *Cell*, 28(4):693–705, 1982.
- [233] D. C. Montgomery and G. C. Runger. *Applied Statistics and Probability for Engineers*. Wiley, New York, 2006.
- [234] T. A. Kunkel. DNA replication fidelity. *J Biol Chem*, 267(26):18251–18254, 1992.
- [235] N. Steuerwald, J. A. Barritt, R. Adler, and *et al.* Quantification of mtDNA in single oocytes, polar bodies and subcellular components by real-time rapid cycle fluorescence monitored PCR. *Zygote*, 8(3):209–15, 2000.
- [236] K. Elliott and M. O'Connor. *Embryogenesis in mammals (Ciba Foundation symposium ; 40)*. Elsevier, 1976.
- [237] L. Piko and K. D. Taylor. Amounts of mitochondrial DNA and abundance of some mitochondrial gene transcripts in early mouse embryos. *Dev Biol*, 123(2):364–74, 1987.
- [238] N. G. Larsson and D. A. Clayton. Molecular genetic aspects of human mitochondrial disorders. *Annu Rev Genet*, 29:151–178, 1995.
- [239] D. M. Turnbull and R. N. Lightowlers. An essential guide to mtDNA maintenance. *Nat Genet*, 18(3):199–200, 1998.
- [240] M. L. Collins, S. Eng, R. Hoh, and M. K. Hellerstein. Measurement of mitochondrial DNA synthesis *in vivo* using a stable isotope-mass spectrometric technique. *J Appl Physiol*, 94:2203–2211, 2003.
- [241] R. B. Cervantes, J. R. Stringer, C. Shao, and *et al.* Embryonic stem cells and somatic cells differ in mutation frequency and type. *Proc Natl Acad Sci USA*, 99(6):3586–3590, 2002.
- [242] N. G. Larsson, J. Wang, H. Wilhelmsson, and *et al.* Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat Genet*, 18:231–236, 1998.
- [243] C. Karatza, W. D. Stein, and S. Shall. Kinetics of *in vitro* ageing of mouse embryo fibroblasts. *J Cell Sci*, 65:163–75, 1984.
- [244] N. J. Sissman. Developmental landmarks in cardiac morphogenesis: Comparative chronology. *Am J Cardiol*, 25:141–148, 1970.
- [245] J. N. Weiss. The Hill equation revisited: uses and misuses. *FASEB J*, 11:835–841, 1997.
- [246] A. R. J. Cadet, J. Cadet, L. Moller, and *et al.* Are we sure we know how to measure 8-oxo-7,8-dihydroguanine in DNA from human cells? *Arch Biochem Biophys*, 423(1):57–65, 2004.
- [247] ESCODD. Comparative analysis of baseline 8-oxo-7,8-dihydroguanine in mammalian cell DNA, by different methods in different laboratories: an approach to consensus. *Free Rad Res*, 23:2129–2133, 2002.
- [248] M. Limson and C. M. Jackson. Changes in the weights of various organs and systems of young rats maintained on a low-protein diet. *J Nutr*, 5(2):163–174, 1931.
- [249] S. Chanda and H. M. Mehendale. Hepatic cell division and tissue repair: a key to survival after liver injury. *Mol Med Today*, 2(2):82–9, 1996.
- [250] F. Diaz and C. T. Moraes. Mitochondrial biogenesis and turnover. *Cell Calcium*, 44(1):24–35, 2008.
- [251] I. Kim, S. Rodriguez-Enriquez, and J. J. Lemasters. Selective degradation of mitochondria by mitophagy. *Arch Biochem Biophys*, 462(2):245–53, 2007.

- [252] H. Nakatogawa, K. Suzuki, Y. Kamada, and Y. Ohsumi. Dynamics and diversity in autophagy mechanisms: lessons from yeast. *Nat Rev Mol Cell Biol*, 10(7):458–67, 2009.
- [253] E. L. Axe, S. A. Walker, M. Manifava, and *et al.* Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *J Cell Biol*, 182(4):685–701, 2008.
- [254] K. Nowikovsky, S. Reipert, R. J. Devenish, and R. J. Schweyen. Mdm38 protein depletion causes loss of mitochondrial k^+/h^+ exchange activity, osmotic swelling and mitophagy. *Cell Death Differ*, 14(9):1647–56, 2007.
- [255] G. Twig, A. Elorza, A. J. Molina, and *et al.* Fission and selective fusion govern mitochondrial segregation and elimination by autophagy. *EMBO J*, 27(2):433–46, 2008.
- [256] B. Westermann. Mitochondrial fusion and fission in cell life and death. *Nat Rev Mol Cell Biol*, 11(12):872–84, 2010.
- [257] I. Kissova, M. Deffieu, S. Manon, and N. Camougrand. Uth1p is involved in the autophagic degradation of mitochondria. *J Biol Chem*, 279(37):39068–74, 2004.
- [258] R. Tal, G. Winter, N. Ecker, and *et al.* Aup1p, a yeast mitochondrial protein phosphatase homolog, is required for efficient stationary phase mitophagy and cell survival. *J Biol Chem*, 282(8):5617–24, 2007.
- [259] T. Kanki, K. Wang, Y. Cao, and *et al.* Atg32 is a mitochondrial protein that confers selectivity during mitophagy. *Dev Cell*, 17(1):98–109, 2009.
- [260] K. Okamoto, N. Kondo-Okamoto, and Y. Ohsumi. Mitochondria-anchored receptor atg32 mediates degradation of mitochondria via selective autophagy. *Dev Cell*, 17(1):87–97, 2009.
- [261] C. W. Birky. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proc Natl Acad Sci U S A*, 92(25):11331–11338, 1995.
- [262] D. Bogenhagen and D. A. Clayton. Mouse L cell mitochondrial DNA molecules are selected randomly for replication throughout the cell cycle. *Cell*, 11:pp. 719–727, 1977.
- [263] G. Biswas, O. A. Adebajo, B. D. Freedman, and *et al.* Retrograde Ca^{2+} signaling in C2C12 skeletal myocytes in response to mitochondrial genetic and metabolic stress: a novel mode of inter-organelle crosstalk. *EMBO J*, 18:522–533, 1999.
- [264] A. Terman and U. T. Brunk. Autophagy in cardiac myocyte homeostasis, aging, and pathology. *Cardiovasc Res*, 68(3):355–365, 2005.
- [265] G. K. Michalopoulos and M. C. DeFrances. Liver regeneration. *Science*, 276(5309):60–6, 1997.
- [266] D. Wang, D. A. Kreuzer, and J. M. Essigmann. Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions. *Mutat Res*, 400(1-2):99–115, 1998.
- [267] K. S. Lim, K. Jeyaseelan, M. Whiteman, and *et al.* Oxidative damage in mitochondrial DNA is not extensive. *Ann N Y Acad Sci*, 1042:210–20, 2005.
- [268] B. Halliwell. Oxidative stress and cancer: have we moved forward? *Biochem J*, 401(1):1–11, 2007.
- [269] S. K. Poovathingal, J. Gruber, B. Halliwell, and R. Gunawan. Stochastic drift in mitochondrial DNA point mutations: a novel perspective *ex silico*. *PLoS Comput Biol*, 5(11):e1000572, 2009.
- [270] A. Nicholas, Y. Kraytsberg, X. Guo, and K. Khrapko. On the timing and the extent of clonal expansion of mtDNA deletions: evidence from single-molecule PCR. *Exp Neurol*, 218(2):316–9, 2009.
- [271] R. A. Menzies and P. H. Gold. The turnover of mitochondria in a variety of tissues of young adult and aged rats. *J Biol Chem*, 246(8):2425–9, 1971.

- [272] N. G. Lipsky and P. L. Pedersen. Mitochondrial turnover in animal cells. half-lives of mitochondria and mitochondrial subfractions of rat liver based on [¹⁴C]bicarbonate incorporation. *J Biol Chem*, 256(16):8652–7, 1981.
- [273] S. Miwa, C. Lawless, and T. von Zglinicki. Mitochondrial turnover in liver is fast *in vivo* and is accelerated by dietary restriction: application of a simple dynamic model. *Aging Cell*, 7(6):920–3, 2008.
- [274] H. Korr, C. Kurz, T. O. Seidler, and *et al.* Mitochondrial DNA synthesis studied autoradiographically in various cell types *in vivo*. *Braz J Med Biol Res*, 31(2):289–98, 1998.
- [275] M. J. Fletcher and D. R. Sanadi. Turnover of rat-liver mitochondria. *Biochim Biophys Acta*, 51:356–360, 1961.
- [276] R. P. Huemer, K. D. Lee, A. E. Reeves, and C. Bickert. Mitochondrial studies in senescent mice-II. specific activity, buoyant density, and turnover of mitochondrial DNA. *Experimental Gerontology*, 6(5):327–334, 1971.
- [277] V. K. Mootha, J. Bunkenborg, J. V. Olsen, and *et al.* Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell*, 115(5):629–640, 2003.
- [278] D. Mijaljica, M. Prescott, and R. J. Devenish. Different fates of mitochondria: alternative ways for degradation? *Autophagy*, 3(1):4–9, 2007.
- [279] A. Jaleel, K. R. Short, Y. W. Asmann, and *et al.* *In vivo* measurement of synthesis rate of individual skeletal muscle mitochondrial proteins. *Am J Physiol Endocrinol Metab*, 295(5):E1255–1268, 2008.
- [280] R. J. Burgess, J. H. Walker, and R. J. Mayer. Choice of precursors for the measurement of protein turnover by the double-isotope method. application to the study of mitochondrial proteins. *Biochem J*, 176(3):919–26, 1978.
- [281] D. F. Dai, L. F. Santana, M. Vermulst, and *et al.* Overexpression of catalase targeted to mitochondria attenuates murine cardiac aging. *Circulation*, 119(21):2789–97, 2009.
- [282] M. M. Ip, P. Y. Chee, and R. W. Swick. Turnover of hepatic mitochondrial ornithine aminotransferase and cytochrome oxidase using [¹⁴C]carbonate as tracer. *Biochimica et Biophysica Acta*, 354(1):29–38, 1974.
- [283] M. Nicoletti, C. Guerri, and S. Grisolia. Turnover of carbamyl phosphate synthase, of other mitochondrial enzymes and of rat tissues. effect of diet and of thyroidectomy. *European Journal of Biochemistry*, 75(2):583–592, 1977.
- [284] H. S. Marver, A. Collins, D. P. Tschudy, and M. Rechcigl Jr. Delta-aminolevulinic acid synthetase. II. induction in rat liver. *Journal of Biological Chemistry*, 241(19):4323–4329, 1966.
- [285] R. Wallace, E. Knecht, and S. Grisolia. Turnover of rat liver ornithine transcarbamylase. *FEBS lett*, 208(2):427–30, 1986.
- [286] M. D. Grozdova and I. K. Starostina. Protein synthesis in the myocardium in experimental allergic lesion of the heart in rabbits. *Vopr Med Khim*, 19(4):403–7, 1973.
- [287] A. D. N. J. deGrey. Mitochondrial mutations in mammalian aging: An over-hasty about-turn. *Rejuv Res*, 7(3):171–174, 2004.
- [288] N. D. Bodyak, E. Nekhaeva, J. Y. Wei, and K. Khrapko. Quantification and sequencing of somatic deleted mtDNA in single cells: evidence for partially duplicated mtDNA in aged human tissues. *Hum Mol Genet*, 10(1):17–24, 2001.
- [289] K. Khrapko, N. Bodyak, W. G. Thilly, and *et al.* 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):2434–41, 1999.

- [290] E. Nekhaeva, N. D. Bodyak, Y. Kravtsov, and *et al.* Clonally expanded mtDNA point mutations are abundant in individual cells of human tissues. *Proc Natl Acad Sci USA*, 99:5521–5526, 2002.
- [291] L. Cao, H. Shitara, T. Horii, and *et al.* The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. *Nat Genet*, 39(3):386–90, 2007.
- [292] S. J. Dubec, R. Aurora, and H. P. Zassenhaus. Mitochondrial DNA mutations may contribute to aging via cell death caused by peptides that induce cytochrome c release. *Rejuvenation Res*, 11(3):611–9, 2008.
- [293] F. J. Miller, F. L. Rosenfeldt, C. Zhang, and *et al.* Precise determination of mitochondrial DNA copy number in human skeletal and cardiac muscle by a PCR-based assay: lack of change of copy number with age. *Nucleic Acids Res*, 31(11):e61, 2003.
- [294] M. Sciacco, E. Bonilla, E. A. Schon, and *et al.* Distribution of wild-type and common deletion forms of mtDNA in normal and respiration-deficient muscle fibers from patients with mitochondrial myopathy. *Hum Mol Genet*, 3(1):13–9, 1994.
- [295] F. Diaz, M. P. Bayona-Bafaluy, M. Rana, and *et al.* Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control. *Nucleic Acids Res*, 30(21):4626–33, 2002.
- [296] H. Fukui and C. T. Moraes. Mechanisms of formation and accumulation of mitochondrial DNA deletions in aging neurons. *Hum Mol Genet*, 18(6):1028–36, 2009.
- [297] D. Bogenhagen, C. Lowell, and Clayton D. A. Mechanism of mitochondrial DNA replication in mouse L-cells. replication of unicircular dimer molecules. *J Mol Biol*, 148(1):77–93, 1981.
- [298] R. S. Williams, S. Salmons, E. A. Newsholme, and *et al.* Regulation of nuclear and mitochondrial gene expression by contractile activity in skeletal muscle. *J Biol Chem*, 261(1):376–80, 1986.
- [299] H. A. Coller, K. Khrapko, P. Herrero-Jimenez, and *et al.* Clustering of mutant mitochondrial DNA copies suggests stem cells are common in human bronchial epithelium. *Mutat Res*, 578(1-2):256–71, 2005.
- [300] D. Narendra, A. Tanaka, D. F. Suen, and R. J. Youle. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *J Cell Biol*, 183(5):795–803, 2008.
- [301] N. Arnheim and G. Cortopassi. Deleterious mitochondrial DNA mutations accumulate in aging human tissues. *Mutat Res*, 275(3-6):157–67, 1992.
- [302] G. J. Wang, L. M. Nutter, and S. A. Thayer. Insensitivity of cultured rat cortical neurons to mitochondrial DNA synthesis inhibitors: evidence for a slow turnover of mitochondrial DNA. *Biochem Pharmacol*, 54(1):181–7, 1997.
- [303] R. Del Bo, M. Crimi, M. Sciacco, and *et al.* High mutational burden in the mtDNA control region from aged muscles: a single-fiber study. *Neurobiol Aging*, 24(6):829–38, 2003.
- [304] G. Fayet, M. Jansson, D. Sternberg, and *et al.* Ageing muscle: clonal expansions of mitochondrial DNA point mutations and deletions cause focal impairment of mitochondrial function. *Neuromuscul Disord*, 12(5):484–93, 2002.
- [305] E. E. Jazin, L. Cavellier, I. Eriksson, and *et al.* Human brain contains high levels of heteroplasmy in the noncoding regions of mitochondrial DNA. *Proc Natl Acad Sci U S A*, 93(22):12382–7, 1996.
- [306] R. S. Cha, W. G. Thilly, and H. Zarbl. N-nitroso-N-methylurea-induced rat mammary tumors arise from cells with preexisting oncogenic Hras1 gene mutations. *Proc Natl Acad Sci U S A*, 91(9):3749–53, 1994.

- [307] D. J. Winton, M. A. Blount, and B. A. Ponder. A clonal marker induced by mutation in mouse intestinal epithelium. *Nature*, 333(6172):463–6, 1988.
- [308] I. Janssen, D. S. Shepard, P. T. Katzmarzyk, and R. Roubenoff. The healthcare costs of sarcopenia in the United States. *J Am Geriatr Soc*, 52(1):80–5, 2004.
- [309] P. Balagopal, O. E. Rooyackers, D. B. Adey, and *et al.* Effects of aging on *in vivo* synthesis of skeletal muscle myosin heavy-chain and sarcoplasmic protein in humans. *Am J Physiol*, 273(4 Pt 1):E790–800, 1997.
- [310] S. A. Hawkins, R. A. Wiswell, T. J. Marcell, and T. J. Marcell. Sarcopenia: causes, consequences, and preventions. *J Gerontol A Biol Sci Med Sci*, 58(10):M911–6, 2003.
- [311] J. Lexell, C. C. Taylor, and M. Sjöström. What is the cause of the ageing atrophy? total number, size and proportion of different fiber types studied in whole *vastus lateralis* muscle from 15- to 83-year-old men. *J Neurol Sci*, 84(2-3):275–94, 1988.
- [312] T. Sato, H. Akatsuka, K. Kito, and *et al.* Age changes in size and number of muscle fibers in human minor pectoral muscle. *Mech Ageing Dev*, 28(1):99–109, 1984.
- [313] J. Lexell. Human aging, muscle mass, and fiber type composition. *J Gerontol A Biol Sci Med Sci*, 50 Spec No:11–6, 1995.
- [314] E. Gutmann and V. Hanzlikova. Motor unit in old age. *Nature*, 209(5026):921–2, 1966.
- [315] S. V. Brooks and J. A. Faulkner. Contraction-induced injury: recovery of skeletal muscles in young and old mice. *Am J Physiol*, 258(3 Pt 1):C436–42, 1990.
- [316] L. Mosoni, M. C. Valluy, B. Serrurier, and *et al.* Altered response of protein synthesis to nutritional state and endurance training in old rats. *Am J Physiol*, 268(2 Pt 1):E328–35, 1995.
- [317] J. E. Morley. Anorexia of aging: physiologic and pathologic. *Am J Clin Nutr*, 66(4):760–73, 1997.
- [318] J. S. Tenover, A. M. Matsumoto, D. K. Clifton, and W. J. Bremner. Age-related alterations in the circadian rhythms of pulsatile luteinizing hormone and testosterone secretion in healthy men. *J Gerontol*, 43(6):M163–9, 1988.
- [319] J. S. Tenover. Effects of testosterone supplementation in the aging male. *J Clin Endocrinol Metab*, 75(4):1092–8, 1992.
- [320] B. M. Carlson. Factors influencing the repair and adaptation of muscles in aged individuals: satellite cells and innervation. *J Gerontol A Biol Sci Med Sci*, 50 Spec No:96–100, 1995.
- [321] R. Weindruch. The retardation of aging by caloric restriction: studies in rodents and primates. *Toxicol Pathol*, 24(6):742–5, 1996.
- [322] C. M. Lee, L. E. Aspnes, S. S. Chung, and *et al.* Influences of caloric restriction on age-associated skeletal muscle fiber characteristics and mitochondrial changes in rats and mice. *Ann N Y Acad Sci*, 854:182–91, 1998.
- [323] L. E. Aspnes, C. M. Lee, R. Weindruch, and *et al.* Caloric restriction reduces fiber loss and mitochondrial abnormalities in aged rat muscle. *FASEB J*, 11(7):573–81, 1997.
- [324] E. A. Bua, S. H. McKiernan, J. Wanagat, and *et al.* Mitochondrial abnormalities are more frequent in muscles undergoing sarcopenia. *J Appl Physiol*, 92(6):2617–24, 2002.
- [325] S. H. McKiernan, E. Bua, J. McGorray, and J. Aiken. Early-onset calorie restriction conserves fiber number in aging rat skeletal muscle. *FASEB J*, 18(3):580–1, 2004.
- [326] S. R. Schwarze, C. M. Lee, S. S. Chung, and *et al.* High levels of mitochondrial DNA deletions in skeletal muscle of old rhesus monkeys. *Mech Ageing Dev*, 83(2):91–101, 1995.

- [327] Y. Rolland, S. Czerwinski, G. Abellan Van Kan, and *et al.* Sarcopenia: its assessment, etiology, pathogenesis, consequences and future perspectives. *J Nutr Health Aging*, 12(7):433–50, 2008.
- [328] R. E. Allen, R. A. Merkel, and R. B. Young. Cellular aspects of muscle growth: myogenic cell proliferation. *J Anim Sci*, 49(1):115–27, 1979.
- [329] D. S. Gokhin, S. R. Ward, S. N. Bremner, and R. L. Lieber. Quantitative analysis of neonatal skeletal muscle functional improvement in the mouse. *J Exp Biol*, 211(Pt 6):837–43, 2008.
- [330] C. M. Guerin and S. G. Kramer. Cytoskeletal remodeling during myotube assembly and guidance: coordinating the actin and microtubule networks. *Commun Integr Biol*, 2(5):452–7, 2009.
- [331] C. J. Snow, M. Goody, M. W. Kelly, and *et al.* Time-lapse analysis and mathematical characterization elucidate novel mechanisms underlying muscle morphogenesis. *PLoS Genet*, 4(10):e1000219, 2008.
- [332] E. Schultz. A quantitative study of the satellite cell population in postnatal mouse lumbrical muscle. *Anat Rec*, 180(4):589–95, 1974.
- [333] D. McKenzie, E. Bua, S. McKiernan, and *et al.* Mitochondrial DNA deletion mutations: a causal role in sarcopenia. *Eur J Biochem*, 269(8):2010–5, 2002.
- [334] M. P. Yaffe. Dynamic mitochondria. *Nat Cell Biol*, 1(6):E149–50, 1999.
- [335] J. Vazquez, A. S. Belmont, and J. W. Sedat. Multiple regimes of constrained chromosome motion are regulated in the interphase *Drosophila* nucleus. *Curr Biol*, 11(16):1227–39, 2001.
- [336] C. M. Lee, R. Weindruch, and J. M. Aiken. Age-associated alterations of the mitochondrial genome. *Free Radic Biol Med*, 22(7):1259–69, 1997.
- [337] M. E. Lopez, N. L. Van Zeeland, D. B. Dahl, and *et al.* Cellular phenotypes of age-associated skeletal muscle mitochondrial abnormalities in rhesus monkeys. *Mutat Res*, 452(1):123–38, 2000.
- [338] R. Cartoni, B. Leger, M. B. Hock, and *et al.* Mitofusins-1/2 and ERRalpha expression are increased in human skeletal muscle after physical exercise. *J Physiol*, 567(Pt 1):349–58, 2005.
- [339] V. P. Skulachev. Mitochondrial filaments and clusters as intracellular power-transmitting cables. *Trends Biochem Sci*, 26(1):23–9, 2001.
- [340] F. Legros, F. Malka, P. Frachon, and *et al.* Organization and dynamics of human mitochondrial DNA. *J Cell Sci*, 117(Pt 13):2653–62, 2004.
- [341] T. Wenz, F. Diaz, D. Hernandez, and C. T. Moraes. Endurance exercise is protective for mice with mitochondrial myopathy. *J Appl Physiol*, 106(5):1712–9, 2009.
- [342] J. F. Passos, T. von Glinicki, and T. B. L. Kirkwood. Mitochondria and ageing: winning and losing in the numbers game. *BioEssays*, 29:908–917, 2007.
- [343] H. Lantz, M. Peltonen, L. Agren, and J. S. Torgerson. Intermittent versus on-demand use of a very low calorie diet: a randomized 2-year clinical trial. *J Intern Med*, 253(4):463–71, 2003.
- [344] D. C. Willcox, B. J. Willcox, H. Todoriki, and *et al.* Caloric restriction and human longevity: what can we learn from the Okinawans? *Biogerontology*, 7(3):173–7, 2006.
- [345] B. J. Merry. Molecular mechanisms linking calorie restriction and longevity. *Int J Biochem Cell Biol*, 34(11):1340–54, 2002.
- [346] R. S. Sohal, H. H. Ku, S. Agarwal, and *et al.* Oxidative damage, mitochondrial oxidant generation and antioxidant defenses during aging and in response to food restriction in the mouse. *Mech Ageing Dev*, 74(1-2):121–33, 1994.

- [347] P. Kapahi, M. E. Boulton, and T. B. Kirkwood. Positive correlation between mammalian life span and cellular resistance to stress. *Free Radic Biol Med*, 26(5-6):495–500, 1999.
- [348] G. Rao, E. Xia, M. J. Nadakavukaren, and A. Richardson. Effect of dietary restriction on the age-dependent changes in the expression of antioxidant enzymes in rat liver. *J Nutr*, 120(6):602–9, 1990.
- [349] R. Sreekumar, J. Unnikrishnan, A. Fu, and *et al.* Effects of caloric restriction on mitochondrial function and gene transcripts in rat muscle. *Am J Physiol Endocrinol Metab*, 283(1):E38–43, 2002.
- [350] A. J. Lambert and B. J. Merry. Lack of effect of caloric restriction on bioenergetics and reactive oxygen species production in intact rat hepatocytes. *J Gerontol A Biol Sci Med Sci*, 60(2):175–80, 2005.
- [351] A. E. Civitaresse, S. Carling, L. K. Heilbronn, and *et al.* Calorie restriction increases muscle mitochondrial biogenesis in healthy humans. *PLoS Med*, 4(3):e76, 2007.
- [352] G. Lopez-Lluch, N. Hunt, B. Jones, and *et al.* Calorie restriction induces mitochondrial biogenesis and bioenergetic efficiency. *Proc Natl Acad Sci U S A*, 103(6):1768–73, 2006.
- [353] E. Nisoli, C. Tonello, A. Cardile, and *et al.* Calorie restriction promotes mitochondrial biogenesis by inducing the expression of eNOS. *Science*, 310(5746):314–7, 2005.
- [354] C. R. Hancock, D. H. Han, K. Higashida, and *et al.* Does calorie restriction induce mitochondrial biogenesis? a re-evaluation. *FASEB J*, 2010.
- [355] R. A. Butow and N. G. Avadhani. Mitochondrial signaling: the retrograde response. *Mol Cell*, 14(1):1–15, 2004.
- [356] J. J. Lehman, P. M. Barger, A. Kovacs, and *et al.* Peroxisome proliferator-activated receptor gamma coactivator-1 promotes cardiac mitochondrial biogenesis. *J Clin Invest*, 106(7):847–56, 2000.
- [357] D. Xu and T. Finkel. A role for mitochondria as potential regulators of cellular life span. *Biochem Biophys Res Commun*, 294(2):245–8, 2002.
- [358] B. S. Kristal and B. P. Yu. *Aging and its modulation by dietary restriction: In Modulation of Aging Processes by Dietary Restriction*. CRC Press, Boca Raton, FL, 1994.
- [359] B. P. Yu and H. Y. Chung. Stress resistance by caloric restriction for longevity. *Ann N Y Acad Sci*, 928:39–47, 2001.
- [360] M. P. Mattson and R. Wan. Beneficial effects of intermittent fasting and caloric restriction on the cardiovascular and cerebrovascular systems. *J Nutr Biochem*, 16(3):129–37, 2005.
- [361] E. S. Hars, H. Qi, A. G. Ryazanov, and *et al.* Autophagy regulates ageing in *c. elegans*. *Autophagy*, 3(2):93–5, 2007.
- [362] M. J. Forster, P. Morris, and R. S. Sohal. Genotype and age influence the effect of caloric intake on mortality in mice. *FASEB J*, 17(6):690–2, 2003.
- [363] J. H. Bielas and J. A. Heddle. Proliferation is necessary for both repair and mutation in transgenic mouse cells. *Proc Natl Acad Sci U S A*, 97(21):11391–6, 2000.
- [364] L. C. Greaves, N. E. Beadle, G. A. Taylor, and *et al.* Quantification of mitochondrial DNA mutation load. *Aging Cell*, 8(5):566–72, 2009.
- [365] P. Wonnapijit, P. F. Chinnery, and D. C. Samuels. Previous estimates of mitochondrial DNA mutation level variance did not account for sampling error: comparing the mtDNA genetic bottleneck in mice and humans. *Am J Hum Genet*, 86(4):540–50, 2010.

- [366] M. Matsumoto and T. Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, 8(1):3–30, 1998.
- [367] P. LeCuyer, R. Simard, E. J. Chen, and W. D. Kelton. An Object-Oriented random-number package with many long streams and substreams. *Oper Res*, 50(6):1073, 2002.
- [368] E. Stadlober. The ratio of uniforms approach for generating discrete random variates. *J Comput Appl Math*, 31(1):181–189, 1990.
- [369] E. Stadlober. Sampling from Poisson, Binomial and Hypergeometric distributions: Ratio of uniforms as a simple and fast alternative. *Berichte der mathematisch-statistischen Sektion in der Forschungsgesellschaft Joanneum*, 303, 1989.
- [370] E. Yang, E. van Nimwegen, M. Zavolan, and *et al.* Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. *Genome Res*, 13(8):1863–72, 2003.
- [371] I. C. Chou and E. O. Voit. Recent developments in parameter estimation and structure identification of biochemical and genomic systems. *Math Biosci*, 219(2):57–83, 2009.
- [372] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox. Real-time kinetics of gene activity in individual bacteria. *Cell*, 123(6):1025–36, 2005.
- [373] J. Yu, J. Xiao, X. Ren, and *et al.* Probing gene expression in live cells, one protein molecule at a time. *Science*, 311(5767):1600–3, 2006.
- [374] A. Colman-Lerner, A. Gordon, E. Serra, and *et al.* Regulated cell-to-cell variation in a cell-fate decision system. *Nature*, 437:699–706, 2005.
- [375] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in *escherichia coli*. *Nature*, 403:339–342, 2000.
- [376] D. Fange and J. Elf. Noise-induced *min* phenotypes in *E. coli*. *PLoS Comput Biol*, 2(6):e80, 2006.
- [377] M. Samoilov, S. Plyasunov, and A. P. Arkin. Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc Natl Acad Sci U S A*, 102(7):2310–5, 2005.
- [378] T. Tian, S. Xu, J. Gao, and K. Burrage. Simulated maximum likelihood method for estimating kinetic rates in gene expression. *Bioinformatics*, 23(1):84–91, 2007.
- [379] A. Golightly and D. J. Wilkinson. Bayesian sequential inference for stochastic kinetic biochemical network models. *J Comput Biol*, 13(3):838–51, 2006.
- [380] A. Golightly and D. J. Wilkinson. Bayesian inference for a discretely observed stochastic kinetic model. *Stat Comput*, (18):125–135, 2008.
- [381] S. Reinker, R. M. Altman, and J. Timmer. Parameter estimation in stochastic biochemical reactions. *Syst Biol (Stevenage)*, 153(4):168–78, 2006.
- [382] D. W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization (Wiley Series in Probability and Statistics)*. Wiley, 1992.
- [383] S. Macnamara, A. M. Bersani, K. Burrage, and R. B. Sidje. Stochastic chemical kinetics and the total quasi-steady-state assumption: application to the stochastic simulation algorithm and chemical master equation. *J Chem Phys*, 129(9):095105, 2008.
- [384] S. Macnamara, K. Burrage, and R. B. Sidje. Multiscale modeling of chemical kinetics via the master equation. *SIAM J.; Multiscale Modeling & Simulation*, 6(4):1146–1168, 2008.
- [385] B. Munsky and M. Khammash. The finite state projection algorithm for the solution of the chemical master equation. *J Chem Phys*, 124(4):044104, 2006.

- [386] M. A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem. A*, 104:1876–89, 2000.
- [387] S. Kullback and S. Leibler. On information and sufficiency. *Ann Math Stat*, 22(1):79–86, 1951.
- [388] R. Storn and K. Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *J Global Optim*, 1997.
- [389] K. Zielinski, D. Peters, and R. Laur. Stopping criteria for single-objective optimization. Singapore, 2005. Proceedings of the Third International Conference on Computational Intelligence, Robotics and Autonomous Systems.
- [390] M. T. Chen and R. Weiss. Artificial cell-cell communication in yeast *saccharomyces cerevisiae* using signaling elements from *arabidopsis thaliana*. *Nat Biotechnol*, 23(12):1551–5, 2005.
- [391] A. Arkin, J. Ross, and H. H. McAdams. Stochastic kinetic analysis of developmental pathway bifurcation in Phage *lambda*-infected *escherichia coli* cells. *Genetics*, 149:1633–1648, 1998.
- [392] J. R. Pomerening, E. D. Sontag, and J. E. Jr. Ferrell. Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nat Cell Biol*, 5(4):346–51, 2003.
- [393] U. S. Bhalla and R. Iyengar. Emergent properties of networks of biological signaling pathways. *Science*, 283(5400):381–7, 1999.
- [394] R. Gunawan, Y. Cao, L. Petzold, and F. J. Doyle. Sensitivity analysis of discrete stochastic systems. *Biophys J*, 88(4):2530–40, 2005.
- [395] P. Gaspard. Fluctuation theorem for nonequilibrium reactions. *J Chem Phys*, 120(19):8898–905, 2004.
- [396] F. Schlögl. On thermodynamics near a steady state. *Z. Physik*, 248:446–458, 1971.
- [397] F. Schlögl. Chemical reaction models for non-equilibrium phase transitions. *Z. Physik*, 253:147–161, 1972.
- [398] I. E. Nikerel, W. A. van Winden, P. J. Verheijen, and J. J. Heijnen. Model reduction and *a priori* kinetic parameter identifiability analysis using metabolome time series for metabolic reaction networks with linlog kinetics. *Metab Eng*, 11(1):20–30, 2009.
- [399] K. G. Gadkar, R. Gunawan, and 3rd Doyle, F. J. Iterative approach to model identification of biological networks. *BMC Bioinformatics*, 6:155, 2005.
- [400] S. Plyasunov and A.P. Arkin. Efficient stochastic sensitivity analysis of discrete event systems. *J Comp Phys*, 221(2):724–738, 2006.
- [401] Y. Cao, D. T. Gillespie, and L. R. Petzold. Efficient step size selection for the τ -leaping simulation method. *J Chem Phys*, 124:044109, 2006.
- [402] A. Chatterjee and D. G. Vlachos. Binomial distribution based τ -leap accelerated stochastic simulation. *J Chem Phys*, 122:024112, 2005.
- [403] E. L. Haseltine and J. B. Rawlings. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J Chem Phys*, 117:6959, 2002.
- [404] T. Tian and K. Burrage. Binomial leap methods for simulating stochastic chemical kinetics. *J Chem Phys*, 121:10356, 2004.
- [405] D. T. Gillespie. Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem*, 58:35, 2007.
- [406] P. K. Mouli, G. Twig, and O. S. Shirihai. Frequency and selectivity of mitochondrial fusion are key to its quality maintenance function. *Biophys J*, 96(9):3509–18, 2009.

- [407] T. Taivassalo and R. G. Haller. Implications of exercise training in mtDNA defects—use it or lose it? *Biochim Biophys Acta*, 1659(2-3):221–31, 2004.
- [408] T. Taivassalo, E. A. Shoubridge, J. Chen, and *et al.* Aerobic conditioning in patients with mitochondrial myopathies: physiological, biochemical, and genetic effects. *Ann Neurol*, 50(2):133–41, 2001.
- [409] T. Taivassalo, K. Fu, T. Johns, and *et al.* Gene shifting: a novel therapy for mitochondrial myopathy. *Hum Mol Genet*, 8(6):1047–52, 1999.
- [410] B. Di Ventura, C. Lemerle, K. Michalodimitrakis, and L. Serrano. From *in vivo* to *in silico* biology and back. *Nature*, 443(7111):527–33, 2006.
- [411] S. Isaacson. Relationship between the reaction-diffusion master equation and particle tracking models. *J Phys A-Math Theor*, 41(065003), 2008.
- [412] R. Erban and J. Chapman. Stochastic modelling of reaction-diffusion process: algorithms for bimolecular reactions. *Phys Biol*, 6(046001), 2009.
- [413] J. Elf and M. Ehrenberg. Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. *Syst Biol (Stevenage)*, 1(2):230–6, 2004.

Appendix A

Citations of Random Mutation Capture Assay

A.1 Citations related to the direct application of RMC assay, since its inception. (source: Pubmed)

Year of Publication	Author	Article	Journal	Volume	Issue	Page No.	Impact Factor (JCR 2009)
1	2010	Perzzi, B <i>et al.</i> The use of PIG-A as a sentinel gene for the study of the somatic mutation rate and of mutagenic agents <i>in vivo</i>	<i>Mut. Res.</i>	705	1	pp 3-10	7.097
2	2010	Gorman, S <i>et al.</i> Mitochondrial mutagenesis induced by tumor-specific radiation bystander effects.	<i>J. Mol. Med.</i>	88	7	701-8	5.004
3	2010	Chen, H <i>et al.</i> Mitochondrial fusion is required for mtDNA stability in skeletal muscle and tolerance of mtDNA mutations.	<i>Cell</i>	141	2	280-9	31.152
4	2010	Dai, D.F. <i>et al.</i> Age-dependent cardiomyopathy in mitochondrial mutator mice is attenuated by overexpression of catalase targeted to mitochondria.	<i>Aging Cell</i>	9	4	536-44	7.554
5	2010	Salk, J.J. <i>et al.</i> Mutational Heterogeneity in Human Cancers: Origin and Consequences	<i>Ann. Rev. Path. Mech.</i>	5	51-75	13.5	
6	2009	Poovathingal, S.K. <i>et al.</i> Stochastic Drift in Mitochondrial DNA Point Mutations: A Novel Perspective <i>Ex Silico</i>	<i>PLOS Comp. Biol.</i>	5	11	e1000572	5.759
7	2009	Milbury, C.A. <i>et al.</i> PCR-Based Methods for the Enrichment of Minority Alleles and Mutations	<i>Clin. Chem.</i>	55	4	632-40	6.263
8	2009	Khrapko, K. Mitochondrial DNA mutations and aging: devils in the details?	<i>Trend. Genet.</i>	25	2	91-98	8.689
9	2009	Martin, G.M. <i>et al.</i> Aging and Cancer: Two Sides of the Same Coin?	<i>J. Gerontol. - A</i>	64	6	615-17	3.083
10	2009	Dai, D.F. <i>et al.</i> Overexpression of catalase targeted to mitochondria attenuates murine cardiac aging.	<i>Circulation</i>	119	21	2789-97	14.816
11	2009	Greaves, L.C. <i>et al.</i> Quantification of mitochondrial DNA mutation load.	<i>Aging Cell</i>	8	5	566-72	7.554
12	2009	Edgar, D & Trifunovic, A. The mtDNA mutator mouse: Dissecting mitochondrial involvement in aging.	<i>Aging</i>	1	12	1028-32	
13	2009	Gruber, J. <i>et al.</i> The mitochondrial free radical theory of ageing - Where do we stand?	<i>Front. Biosci.</i>	13	6554-79	3.603	

14	2008	Kujoth, G.C. <i>et al.</i>	Evolving insight into the role of mitochondrial DNA mutations in aging	<i>Exp. Gerontol.</i>	43	1	20-23	3,342
15	2008	Vermulst, M <i>et al.</i>	Quantification of random mutations in the mitochondrial genome.	<i>Methods</i>	46	4	263-8	3,763
16	2008	Vermulst, M <i>et al.</i>	DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice.	<i>Nat. Genet.</i>	40	4	392-4	34,284
17	2008	Loeb, L.A. <i>et al.</i>	Cancers exhibit a mutator phenotype: clinical implications.	<i>Cancer Res.</i>	68	10	3551-7	7,543
18	2008	Kraytberg, Y <i>et al.</i>	Single molecule PCR in mtDNA mutational analysis: Genuine mutations vs. damage bypass-derived artifacts.	<i>Methods</i>	46	4	269-73	3,763
19	2008	Salvioli, S. <i>et al.</i>	The impact of mitochondrial DNA on human lifespan: a view from studies on centenarians.	<i>Biotech. J.</i>	3	6	740-9	3,146
20	2007	Garcia, A.M. <i>et al.</i>	A model system for analyzing somatic mutations in <i>Drosophila melanogaster</i>	<i>Nat. Meth.</i>	4	5	401-403	16,874
21	2007	Bielas, J.H. <i>et al.</i>	LOH-proficient embryonic stem cells: a model of cancer progenitor cells?	<i>Trends Genet.</i>	23	4	154-7	8,689
22	2007	Khrapko, K.	Mitochondrial DNA mutations and aging: a case closed?	<i>Nat. Genet.</i>	39	4	445-6	34,284
23	2007	Diehl, F. <i>et al.</i>	Digital quantification of mutant DNA in cancer patients	<i>Cur. Opin. Onc.</i>	19	1	36-42	4,088
24	2007	Laun, P. <i>et al.</i>	Yeast mother cell-specific ageing, genetic (in)stability, and the somatic mutation theory of ageing	<i>Nuc. Acids Res.</i>	35	22	7514-26	7,479
25	2007	Vermulst, M <i>et al.</i>	Mitochondrial point mutations do not limit the natural lifespan of mice.	<i>Nat. Genet.</i>	39	4	540-3	34,284
26	2007	Zheng, L. <i>et al.</i>	Fen1 mutations result in autoimmunity, chronic inflammation and cancers.	<i>Nat. Med.</i>	13	7	812-9	27,136
27	2007	Kraytberg, Y <i>et al.</i>	Are somatic mitochondrial DNA mutations relevant to our health? A challenge for mutation analysis techniques	<i>Exp. Opin. Med. Diag.</i>	1	1	109-16	4,218

28	2006	Smilenov, L.B.	Tumor development: Haploinsufficiency and local network assembly	<i>Canc. Lett.</i>	240	1	17-28	3.741
29	2006	Venkatesan, R.N., <i>et al.</i>	Generation of mutator mutants during carcinogenesis	<i>DNA Rep.</i>	5	3	95-7	4.199
30	2006	Li, M. <i>et al.</i>	BEAMing up for detection and quantification of rare sequence variants	<i>Nat. Meth.</i>	3	2	95-7	16.874
31	2006	Klein, C.A.	Random mutations, selected mutations: A PIN opens the door to new genetic landscapes	<i>Proc. Natl. Acad. Sci.</i>	103	48	18033-34	9.432
32	2006	Bielas, J.H. <i>et al.</i>	Human cancers express a mutator phenotype.	<i>Proc. Natl. Acad. Sci.</i>	103	48	18238-42	9.432
33	2005	Beckman, R.A. & Loeb, L.A.	Negative clonal selection in tumor evolution	<i>Genetics</i>	171	4	2123-31	3.889
34	2005	Bielas, J.H. & Loeb, L.A.	Quantification of random genomic mutations	<i>Nat. Meth.</i>	2		285-90	16.874

Appendix B

COMSOL report — sarcopenic model

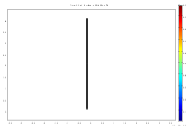
B.1 COMSOL report on the reaction-diffusion multiphysics model of sarcopenic skeletal muscle

2011-02-22 system:/home/CP Codes/MiceCodes/Compre...CRE_Sarcopenia_Module_New_working.html

#1



COMSOL Model Report



1. Table of Contents

- Title - COMSOL Model Report
- Table of Contents
- Model Properties
- Constants
- Global Expressions
- Geometry
- Geom1
- Solver Settings
- Postprocessing
- Variables

2. Model Properties

Property	Value
Model name	
Author	
Company	
Department	
Reference	
URL	
Saved date	Nov 17, 2010 1:03:29 AM
Creation date	Feb 13, 2010 11:31:46 AM
COMSOL version	COMSOL 3.5.0.603

File name:

/data/home/suresh/CP Codes/MiceCodes/Comprehensive/Working/Sarcopenia_Study/COMSOLTrials/NonCRE_Sarcopenia_Module_New_working.mph

Application modes and modules used in this model:

- Geom1 (Axial symmetry (2D))
 - Diffusion (Chemical Engineering Module)

3. Constants

Name	Expression	Value	Description
k_d1	2.705671296e-8	2.705671e-8	Degradation rate of wild-type mtDNA
k_d2	2.705671296e-8	2.705671e-8	Degradation rate of mutant mtDNA
v_max	2.460856939e-14	2.460857e-14	Maximum replication rate of mtDNA
k_m	5.5e-7	5.5e-7	Mutation rate of mtDNA
Hill_Coeff	2	2	Hill coefficient (n)
Hill_Const	0.15	0.15	Hill Constant (K)
b_F	3	3	Replicative advantage factor
No_React_Rate	0	0	
Diff_Const	7e-16	7e-16	
Ave_No	6.023e23	6.023e23	
Cell_Vol	$\pi \cdot (32.5e-6^2) \cdot 25e-6$	8.295768e-14	
MolW	1345	1345	
MolM	1123	1123	
c_W0	2.00138451e-8	2.001385e-8	

4. Global Expressions

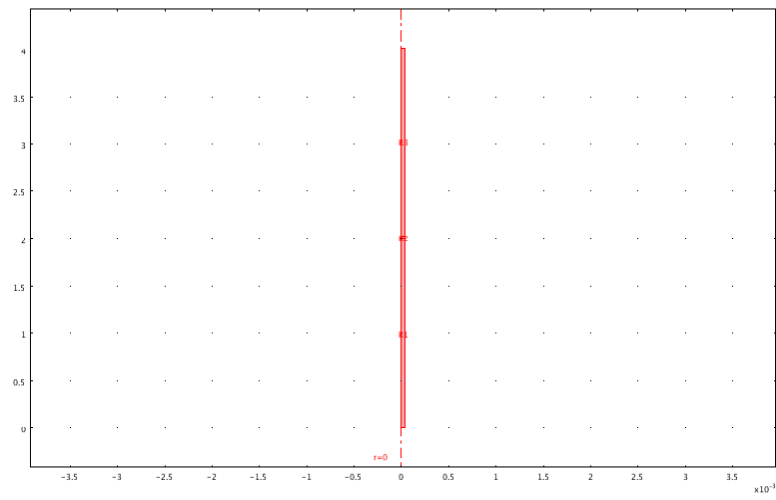
Name	Expression	Unit	Description
Ratio1	$c_W / (c_W + c_M)$	1	
Ratio2	1 - Ratio1	1	
WMRatio	Ratio1	1	
HillTerm	$1 - ((WMRatio^Hill_Coeff) / ((Hill_Const^Hill_Coeff) + (WMRatio^Hill_Coeff)))$	1	
k_R	$v_max \cdot HillTerm$	1	
MTRatio	1 - WMRatio	1	
K_1	k_d1		
K_2	k_d2		
K_3	k_R	1	
K_4	b_F * k_R	1	
K_5	k_m * k_R	1	
DiffConst_Value	Diff_Const		
Conc_W	$(MolW / Ave_No) / Cell_Vol$		
Conc_M	$(MolM / Ave_No) / Cell_Vol$		

5. Geometry

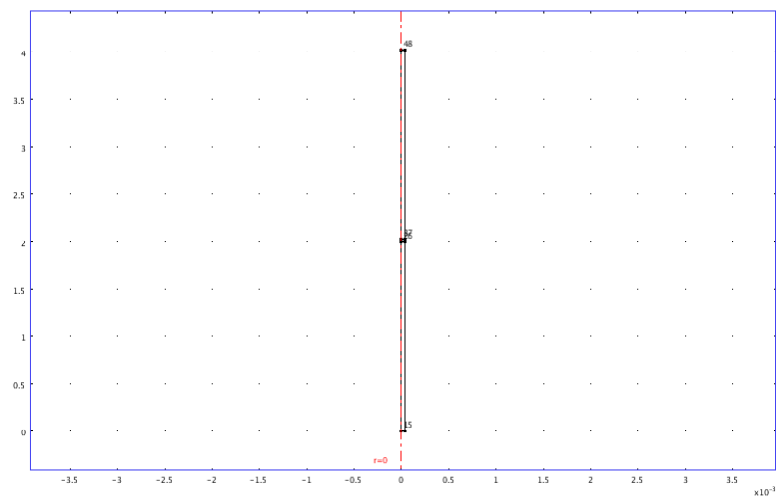
Number of geometries: 1

2011-02-22 system:/home/CPPCodes/MiceCodes/Compre...CRE_Sarcopenia_Module_New_working.html

#2

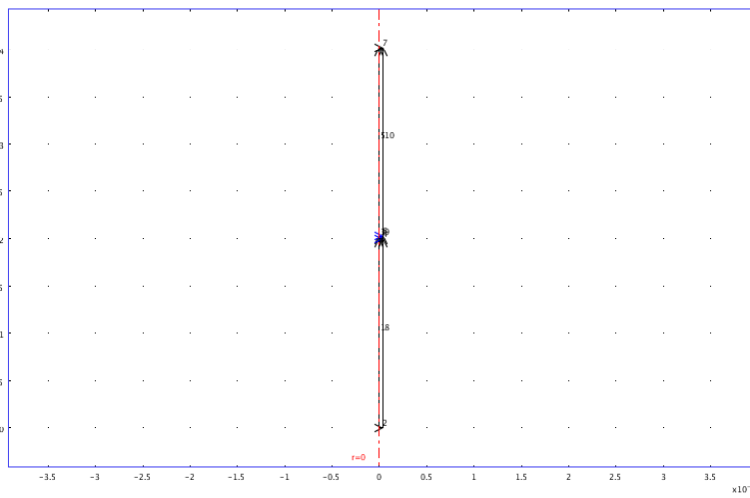


5.1.1. Point mode

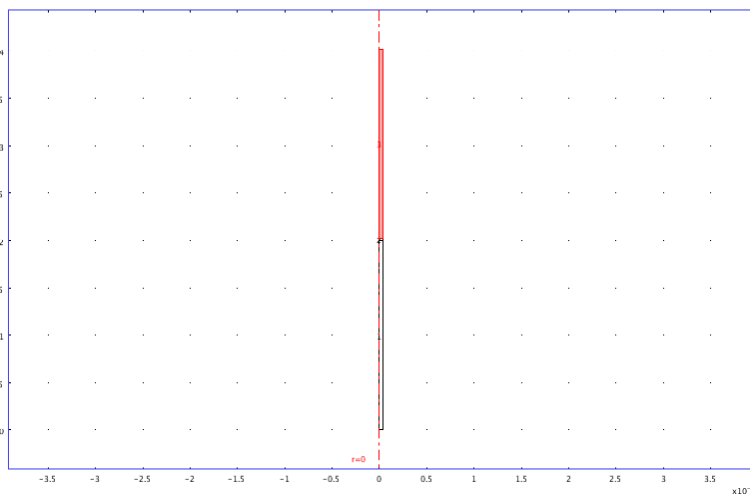


2011-02-22 system:/home/CPPCodes/MiceCodes/Compre...CRE_Sarcopenia_Module_New_working.html

#3



5.1.3. Subdomain mode



6. Geom1

Space dimensions: Axial symmetry (2D)

Independent variables: r, phi, z

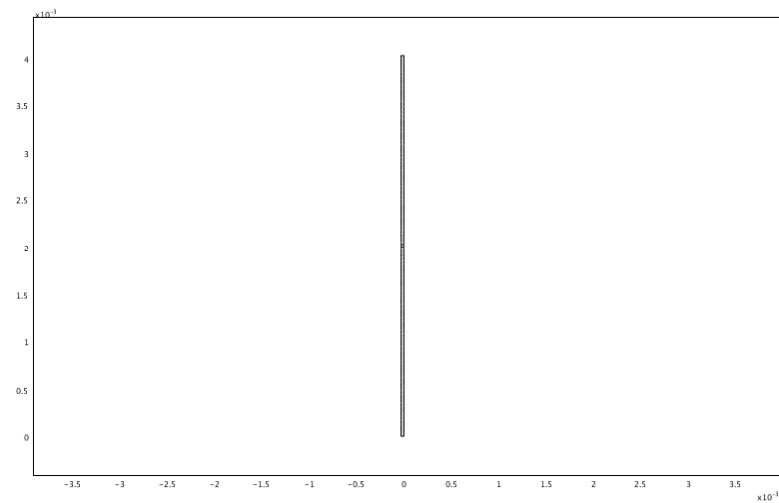
6.1. Mesh

6.1.1. Mesh Statistics

Number of degrees of freedom	1942
Number of mesh points	295
Number of elements	382
Triangular	382
Quadrilateral	0
Number of boundary elements	220
Number of vertex elements	8
Minimum element quality	0,812
Element area ratio	0,007

2011-02-22 system:/home/CPPCodes/MiceCodes/Compre...CRE_Sarcopenia_Module_New_working.html

#4



6.2. Application Mode: Diffusion (chdi)

Application mode type: Diffusion (Chemical Engineering Module)

Application mode name: chdi

6.2.1. Application Mode Properties

Property	Value
Default element type	Lagrange - Quadratic
Analysis type	Transient
Equilibrium assumption	Off
Frame	Frame (ref)
Weak constraints	Off
Constraint type	Ideal

6.2.2. Variables

Dependent variables: c_W, c_M

Shape functions: shlag(2,'c_W'), shlag(2,'c_M')

Interior boundaries not active

6.2.3. Boundary Settings

Boundary	1-3, 5, 7-10
Type	Insulation/Symmetry

6.2.4. Subdomain Settings

Subdomain		1, 3	2
Diffusion coefficient (D)	m ² /s	{DiffConst_Value;DiffConst_Value}	{DiffConst_Value;DiffConst_Value}
Reaction rate (R)	mol/(m ³ .s)	{-(K_1*c_W)+(K_3*Ratio1)-(K_5*Ratio1)}-(K_2*c_M)+(K_4*Ratio2)+(K_5*Ratio1)}	{-(K_1*c_W)+(K_3*Ratio1)-(K_5*Ratio1)}-(K_2*c_M)+(K_4*Ratio2)+(K_5*Ratio1)}

Subdomain initial value		1, 3	2
Concentration, c_W (c_W)	mol/m ³	c_W0	Conc_W
Concentration, c_M (c_M)	mol/m ³	0	Conc_M

7. Solver Settings

Solve using a script: off

Analysis type	Transient
Auto select solver	On
Solver	Time dependent
Solution form	Automatic
Symmetric	auto
Adaptive mesh refinement	Off
Optimization/Sensitivity	Off
Plot while solving	Off

7.1. Direct (UMFPACK)

Solver type: Linear system solver

Parameter	Value
Pivot threshold	0.1
Memory allocation factor	0.7

2011-02-22 system:/home/CPPCodes/MiceCodes/Compre...CRE_Sarcopenia_Module_New_working.html

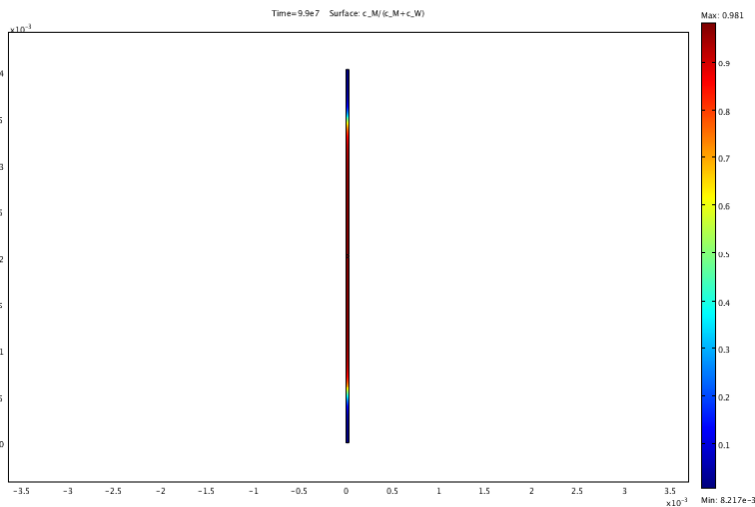
#5

Parameter	Value
Times	range(0,100000,99000000)
Relative tolerance	1e-6
Absolute tolerance	1e-8
Times to store in output	Specified times
Time steps taken by solver	Free
Maximum BDF order	5
Singular mass matrix	Maybe
Consistent initialization of DAE systems	Backward Euler
Error estimation strategy	Include algebraic
Allow complex numbers	Off

7.3. Advanced

Parameter	Value
Constraint handling method	Elimination
Null-space function	Automatic
Automatic assembly block size	On
Assembly block size	1000
Use Hermitian transpose of constraint matrix and in symmetry detection	Off
Use complex functions with real input	Off
Stop if error due to undefined operation	On
Store solution on file	Off
Type of scaling	Automatic
Manual scaling	On
Row equilibration	Off
Manual control of reassembly	On
Load constant	On
Constraint constant	On
Mass constant	On
Damping (mass) constant	On
Jacobian constant	On
Constraint Jacobian constant	On

8. Postprocessing



9. Variables

9.1. Boundary

Name	Description	Unit	Expression
ndflux_c_W_chdi	Normal diffusive flux, c_W	mol/(m ² *s)	nr_chdi * dflux_c_W_r_chdi+nz_chdi * dflux_c_W_z_chdi
ndflux_c_M_chdi	Normal diffusive flux, c_M	mol/(m ² *s)	nr_chdi * dflux_c_M_r_chdi+nz_chdi * dflux_c_M_z_chdi

9.2. Subdomain

Name	Description	Unit	Expression
grad_c_W_r_chdi	Concentration gradient, c_W, r component	mol/m ⁴	c_Wr
dflux_c_W_r_chdi	Diffusive flux, c_W, r component	mol/(m ² *s)	-Drr_c_W_chdi * c_Wr-Drz_c_W_chdi * c_Wz
grad_c_W_z_chdi	Concentration gradient, c_W, z component	mol/m ⁴	c_Wz
dflux_c_W_z_chdi	Diffusive flux, c_W, z component	mol/(m ² *s)	-Dzr_c_W_chdi * c_Wr-Dzz_c_W_chdi * c_Wz
grad_c_W_chdi	Concentration gradient, c_W	mol/m ⁴	sqrt(grad_c_W_r_chdi ² +grad_c_W_z_chdi ²)
dflux_c_W_chdi	Diffusive flux, c_W	mol/(m ² *s)	sqrt(dflux_c_W_r_chdi ² +dflux_c_W_z_chdi ²)
grad_c_M_r_chdi	Concentration gradient, c_M, r component	mol/m ⁴	c_Mr
dflux_c_M_r_chdi	Diffusive flux, c_M, r component	mol/(m ² *s)	-Drr_c_M_chdi * c_Mr-Drz_c_M_chdi * c_Mz
grad_c_M_z_chdi	Concentration gradient, c_M, z component	mol/m ⁴	c_Mz
dflux_c_M_z_chdi	Diffusive flux, c_M, z component	mol/(m ² *s)	-Dzr_c_M_chdi * c_Mr-Dzz_c_M_chdi * c_Mz
grad_c_M_chdi	Concentration gradient, c_M	mol/m ⁴	sqrt(grad_c_M_r_chdi ² +grad_c_M_z_chdi ²)
dflux_c_M_chdi	Diffusive flux, c_M	mol/(m ² *s)	sqrt(dflux_c_M_r_chdi ² +dflux_c_M_z_chdi ²)

List of Publications

Peer reviewed Journal Publications

- ★ **S. K. Poovathingal**, J. Gruber, B. Halliwell and R. Gunawan, (2009); **Stochastic Drift in Mitochondrial DNA Point Mutations: A Novel Perspective *ex silico***, *PLoS Computational Biology*, Vol. 5, No. 11, e1000572.
- ★ J. Gruber, N. L. Fang, **S. K. Poovathingal** and B. Halliwell, (2009); **Deceptively simple but simply deceptive - *Caenorhabditis elegans* life span studies: considerations for aging and antioxidant effects**, *FEBS letters*, Vol. 583, No. 21, pp. 3377-87.
- ★ J. Gruber, **S. K. Poovathingal**, N. L. Fang, R. Gunawan and B. Halliwell, (2010); ***Caenorhabditis elegans* lifespan studies: The challenge of maintaining synchronous cohorts**, *Rejuvenation Research*, Vol. 13, No. 2-3, pp. 347-349.
- ★ **S. K. Poovathingal** and R. Gunawan, (2010); **Global Parameter Estimation Methods for Stochastic Biochemical Systems**, *BMC Bioinformatics*, Vol. 11, pp. 414.
- ★ **S. K. Poovathingal**, J. Gruber, L. Lakshminarayanan, B. Halliwell and R. Gunawan; **Is mitochondrial DNA turnover slower than commonly assumed?**, *Aging Cell*, In review.
- ★ **S. K. Poovathingal**, J. Gruber, B. Halliwell and R. Gunawan; **Clonal expansion of mtDNA deletions - mechanisms revisited**, In Preparation.
- ★ **S. K. Poovathingal**, J. Gruber, B. Halliwell and R. Gunawan; **Stochastic dynamics of mtDNA deletion and its role in the etiology of sarcopenia in mouse muscle fabric**, In Preparation.
- ★ **S. K. Poovathingal**, J. Gruber, N. L. Fang, B. Halliwell and R. Gunawan; **Maximizing Signal-to-Noise Ratio in the Random Mutation Capture Assay**, In Preparation.

Peer reviewed Conference Publications

- ★ **S. K. Poovathingal** and R. Gunawan; **A Global Approach for Estimating the Kinetic Parameters of Stochastic Biological Systems**, *Foundations of Systems Biology and Engineering (FOSBE-2009)*, Denver, Colorado, USA, Aug 9-12, 2009.

Conferences

- ★ **S. K. Poovathingal**, J. Gruber, B. Halliwell and R. Gunawan; **Aging Studies: A Stochastic Approach**, *12th Annual International Conference on Research in Computational Molecular Biology (RECOMB-2008)*, Singapore, Mar 30-Apr 2, 2008.

-
- ★ **S. K. Poovathingal**, J. Gruber, B. Halliwell and R. Gunawan, **Aging Studies: A Stochastic Approach in point mutation dynamics in mouse model**, *AICHE Annual General Meeting (AICHE-2008)*, Philadelphia, Pennsylvania, USA, Nov 16-21, 2008.
 - ★ **S. K. Poovathingal**, J. Gruber, B. Halliwell and R. Gunawan, **Stochasticity in Mitochondrial DNA Point Mutations and Its Relevance in *Caenorhabditis elegans* Aging**, *10th International Conference on Systems Biology (ICSB-2009)*, Stanford, California, USA, Aug 30-Sep 4, 2009.
 - ★ J. Gruber, **S. K. Poovathingal**, N. L. Fang and B. Halliwell, **Deceptively simple - considerations regarding *Caenorhabditis elegans* lifespan, ageing and antioxidant studies**, *Strategies for Engineered Negligible Senescence (SENS-4)*, Cambridge, England, Sep 3-7, 2009.
 - ★ **S. K. Poovathingal**, L. N. Lakshmanan, J. Gruber, B. Halliwell and R. Gunawan, **Elucidating Mechanisms of Age-dependent Accumulation of Mitochondrial DNA deletions - An *in silico* Approach**, *Biology of Aging; Gordon Research Conferences*, Les Diablerets, Switzerland, Aug 22-27, 2010.
 - ★ **S. K. Poovathingal**, L. N. Lakshmanan, J. Gruber, B. Halliwell and R. Gunawan, **Stochastic Dynamics of Mitochondrial DNA Deletion Accumulation in Mammalian Musculature and its Relevance to Sarcopenia**, *11th International Conference on Systems Biology (ICSB-2010)*, Edinburgh, Scotland, UK, Oct 10-15, 2010.
 - ★ **S. K. Poovathingal**, J. Gruber, B. Halliwell and R. Gunawan, **Sarcopenia *in silico***, *AICHE Annual General Meeting (AICHE-2010)*, Salt Lake City, Utah, USA, Nov 7-12, 2010.

Vita

POOVATHINGAL, Suresh Kumar

Current Address: Blk 333, Clementi Avenue - 2, #10-100, SINGAPORE
- 120333.
contact: +65 9132 0771
e-mail: suresh.poovathingal@nus.edu.sg

Permanent Address: #3714, 10th Cross, 13th B Main, HAL II Stage, Indiranagar, Bangalore - 560 008, INDIA.
contact: +91 080 25273249
e-mail: suresh.tuhh@gmail.com

Research Summary

Extensive experience in systems and computational biology, with specialization in developing stochastic/statistical and deterministic models of biological systems. Experience in working with a team of international researchers from a multidisciplinary research areas. State-of-the-art knowledge of latest programming/parallel computing tools.

Research Interests

Computational modeling & simulation, Process Engineering, Biochemical Engineering, Optimization, systems, Data driven modeling (machine learning), Multivariate statistics and Metabolic engineering.

Research Experience

- ★ **Research Engineer;** Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore; July 2010 – Present.
 - » Reverse engineering of biological systems using systems approach to gain insights towards genetic and pharmacological interventions to prevent, treat or retard age-related pathologies.
 - » Developed parsimonious stochastic models based on Monte Carlo approach to characterize the underlying variability in aging / age-related diseases.
 - » Developed statistical modeling-based approach to reduce the measurement-related variability.
- ★ **Graduate Researcher;** Fraunhofer institute for Ceramic Technologies and Systems (IKTS), Dresden, Germany; Aug 2005 – May 2006.
 - » Developed a fluid dynamics model to characterize the design and optimization of flow channels related to μ -Direct Methanol Fuel Cell (μ -DMFC).

-
- » Implemented a synergistic modeling platform (based on COMSOL Multiphysics) to aid the experimental group of IKTS to test their future flow channel designs.
 - ★ **Graduate Researcher;** Department of Plant Systems, Technische Universität Hamburg - Harburg, Hamburg, Germany; Oct 2004 – Jul 2005.
 - » Involved in conceiving project involving computer-aided design of metal hydride (MH) storage canister for hydrogen storage (Hydrogen economy).
 - » Developed a novel mass transport model for hydrogen transport in MH matrix.
 - ★ **Graduate Research Internship;** AQUABIOTEC GmbH, Lübeck, Germany; Sep 2005 – May 2006.
 - » Kinetic characterization and setting-up of lab-scale high voltage oxidative degradation apparatus for treating industrial high COD effluent.

Non-academic Experience

- ★ **Group Leader;** Plant Engineering Division, Toyota Kirloskar Motor Limited; Mar 2002 – Aug 2003.
 - ‡ Supervision of O&M of utilities department (Power, Water and Air). ‡ Spear headed the "Zero Stoppage Campaign" by implementing and maintaining "Kanban" system for the O&M of the utilities apparels. ‡ Overseeing the operation and performance of the team members (2 team leader and 27 team members). ‡ Training team members on O&M, Safety, Environment and Quality improvement.
- ★ **Process Engineer;** Production Engineering Department, Toyota Kirloskar Motor Limited; Mar 2000 – Mar 2002.
 - ‡ Responsible for projects in environmental division (Water and Wastewater treatment facility). ‡ Training team members on O&M of these plants. ‡ Liaison with pollution control board for plant O&M approvals. ‡ Setting up in-house laboratory for monitoring performance of the treatment plants. ‡ Core member of the ISO 14001 Environmental management system team.

Educational Details

- ★ **Doctor of Philosophy;** Chemical and Biomolecular Engineering, National University of Singapore, Singapore; 2006 – 2011 (Exp).
Specialization: Computational Biology & Systems Biology.
- ★ **Master of Science;** Process Engineering, Technische Universität Hamburg-Harburg (TUHH), Hamburg, Germany; 2003 – 2006.
Specialization: Process Engineering.
- ★ **Bachelor of Engineering;** Chemical Engineering, Bangalore University, Bangalore; 1995 – 1999.

Leadership & Communication

- ★ 8 journal papers in peer-reviewed leading scientific journals (4 published).
- ★ 2 podium presentations and 6 poster presentations at international scientific research conferences.

-
- ★ Supervised final year projects of 4 undergraduate students, NUS.
 - ★ Teaching assistant for several undergraduate courses, NUS.
 - ★ Organizing Committee member, Graduate Conference in Biological and Chemical Engineering, NUS; 2007.
 - ★ Committee member of Indian Students Association, TU-Hamburg; 2005
 - ★ National Cadet Corps (NCC) 'A' certificate. NCC is a unit of Indian Army which imparts military training to students.
 - ★ Presented Reform Activities for the Optimization of the Effluent Treatment Plant at National leadership award for best industrial safety, health and environmental Practices, 2003, organized by Confederation of Indian Industries (CII).

Honors & Awards

- ★ Research Scholarship, NUS offered by Ministry of Education, Singapore.
- ★ Fellowship for outstanding performance (2004, 05) offered by TU-Hamburg, a competitive fellowship offered to two graduate students annually.
- ★ Annual scholarship for exceptional performance (1997, 98 and 99). Competitive scholarship for undergraduate studies, offered by Department of Telecommunication (BSNL), Govt. of India.

Training & Certifications

- ★ Trained in Toyota Motor Corporation, Motomachi, Japan on:
 - O&M of utilities (emphasis on Water and Effluent treatment plant).
 - ISO 14001 elements of Water, Chemical and Energy management.
 - Elements of Toyota Production System (TPS).
- ★ Basic Guidelines of ISO 14001 and Internal Auditors certification training by Aglo-Japanese American Registrars, Thailand and ERM (Environmental Resource Management), India.
- ★ Human Resource Development and techniques to improve industrial relation - Nagaraj Management consultant, India.
- ★ Effective communication methods - Vikas Management Consultant, India.

Skill Sets

Operating System:	Windows; Unix and Linux
Languages:	C; C++; Java (basics); Fortran - 77; 90
Cluster tools:	OpenMP - 2.5; MPI - 1, 2; CUDA - NVIDIA, (2.0 - 2.3) MATLAB (R2006 - R2009); MAPLE, HYSYS; R;
Other Softwares:	COMSOL Multiphysics; Adobe CS3; Corel DRAW - 12; LATEX

References

- ★ **Dr. Rudiyanto Gunawan** (Advisor, NUS)
Asst. Prof., Dept. of Chemical and Biomolecular Engg, NUS, Singapore.
Tel: +65 6516 6619; Email: chegr@nus.edu.sg.
- ★ **Dr. Jan Gruber** (Co-Advisor, NUS)
Research Fellow, Dept. of Biochemistry, NUS, Singapore.
Tel.: +65 6516 4638 Email: jan_gruber@nus.edu.sg
- ★ **Prof. Barry Halliwell** (Co-Advisor, NUS)
Tan Chin Tuan Centennial Professor
Deputy President (Research and Technology)
Prof. of Biochemistry, NUS, Singapore.
Tel.: +65 6516 3247; Email: bchbh@nus.edu.sg.

This page is left blank intentionally