Uniparental inheritance of cytoplasmic genomes: its evolution and consequences

Joshua Christie
Faculty of Science
The University of Sydney

A thesis submitted in fulfilment of the requirements of the degree of

Doctor of Philosophy

August 2016
Statement of Originality

This is to certify that, to the best of my knowledge, the content of this thesis is my own work, except where specifically acknowledged. The work in this thesis has not been previously submitted for a degree at The University of Sydney or any other institution.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Joshua Christie
August 2016
Acknowledgements

When I started my PhD in 2013, I intended to empirically test a hypothesis for the evolution of uniparental inheritance using the slime mold Physarum Polycephalum; indeed, for much of the first 18 months this is what I worked on. My venture into the theory behind mitochondrial inheritance started when I was fortunate to receive a scholarship that encouraged collaboration between biologists and mathematicians. At that time, I wanted to supplement the experimental work with theory exploring whether biparental inheritance in P. Polycephalum could be an adaptive trait. I tried, without success, to outsource the modeling to a mathematician—thinking, in my infinite wisdom, that the modeling was the easy part. My inability to outsource the project was, in hindsight, incredibly fortuitous because it forced me to learn how to build mathematical models. That project eventually morphed into chapter 2, and it was so enjoyable that I decided to switch from an empirical to a theoretical PhD. A number of people have helped me before, during, and after this transition.

First, I’d like to thank my adviser Madeleine Beekman. I am very lucky to have had an adviser who considers student supervision to be her top priority. I rarely had to wait more than a day for a draft to be returned with comments. I think Madeleine would simply see this as a supervisor’s duty, but I know full well that few students receive support of that quality. Thank you for your support when I decided to switch from experimental to theoretical work, even though it was in your best interests that I continue with the empirical work.

Thank you to Tim Schaerf for jumping in as an assistant supervisor once I started working on theory. Knowing that there was somebody to help me if/when I got hopelessly stuck made the transition a lot easier. Thank you to Tanya (slime mold extraordinaire), Julie, and Emily for their help with lab work in the first 18 months—sorry I abandoned the slime molds, but they are in good hands!

I’d like to thank the many past and present members of the Behaviour and Genetics of Social Insects Lab for creating a fun and supportive environment: James, Michael, Francis, Isobel, Nick, Jules, Kath, Rachael, Kait, Nik, Jordan, Matt, Danya, Julia, Tom, Becky, Miles, Ruby, Manlika (Eve), Benjamin, Guiling, Tulio, Julie, Gabi, Emily, Nadine, Ros, Vanina, Guenael, Boris, Tim, Tanya, Ben, and Madeleine (sorry
if I missed anybody). Aside from making the lab a fun place to work, these people have helped out a lot with comments on conference talks and draft manuscripts.

Special thanks to the contemporary PhD students in the lab who shared the experience: Michael for imparting his knowledge about PhD life and for dealing with Jasper (sorry!); James for lessons on obtaining free food and for stories of bee swarms running amok inside of buildings; Francis for opening my eyes to the myriad of uses for dog hair; Isobel for being my personal EvolDir and Web of Science filter; Nick for getting me into Emacs and \LaTeX, but most importantly for “Hamilroyd” and the “Barnaby incident”; and Jules for working on slime mold, so that I could finally be rid of the “slime mold guy” moniker.

I’d also like to thank the wider postgrad community at The University of Sydney (of which there are too many to name) for fun Warrah weekends and other events, as well as weekly frisbee and soccer games. Thanks to the biology postgrad committee presidents (Martyna, Lizzy, Steve, and Charles) for keeping things running, as it is a time-consuming and often thankless job. I’d also like to thank Isobel and Jun for helping me run the grant-writing workshop last year, as well as the speakers who generously volunteered their time to speak to us.

I would like to acknowledge financial support from the following organizations: Department of Education and Training (Australian Government), Winton Foundation, The University of Sydney, European Molecular Biology Laboratory Australia, Australia and New Zealand Industrial and Applied Mathematics, Society for Experimental Biology, Company of Biologists, Society for Mathematical Biology, European Society for Mathematical and Theoretical Biology, Intersect Australia/National Computational Infrastructure, and Society for Molecular Biology and Evolution. Their support has given me access to infrastructure and knowledge, and it has allowed me to attend more than my fair share of conferences.

I thank my parents and siblings for their love and support over the years. The sacrifices you made, and the values you taught me, have given me many opportunities and for that I am very grateful. Lastly, I would like to thank Danica for her patience, love, and support during this thesis. She has been subjected to countless hours of me blathering on about mitochondria, papers, academia, and other equally uninteresting topics. I am forever grateful for your support.
Abstract

Many eons ago, a proto-eukaryote engulfed a prokaryote, giving rise to the most enduring symbiotic partnership in the history of life. That bacterium evolved into the mitochondrion and with it evolved an array of innovations. Mitochondria play a crucial role in the energy production of the cell, a function that has released energy constraints on the eukaryotic cell and enabled the evolution of complex life. Mitochondria have retained their own genome, although its size has been greatly reduced in many organisms, particularly in animals. As a result, eukaryotic cells contain both a nuclear and a mitochondrial genome. In addition to mitochondria, eukaryotes can carry other cytoplasmic genomes: chloroplasts and bacterial endosymbionts. Together, these cytoplasmic genomes share a number of common features: all exist within their host’s cytoplasm and are generally inherited via a single parent. The evolutionary reasons behind uniparental inheritance are not well understood. We do know that many organisms go to great pains to actively avoid the biparental transmission of cytoplasmic genomes. Clearly, uniparental inheritance is important, but why? The most widely accepted explanation for the evolution of uniparental inheritance is conflict between the nuclear and cytoplasmic genomes. According to this hypothesis, uniparental inheritance evolved to protect hosts against “selfish” cytoplasmic genomes—those that invest in their own replication to the detriment of the host. In the first part of this thesis, I challenge this hypothesis, arguing that it requires unrealistic biological conditions. Instead, I propose two alternative hypotheses for the evolution of uniparental inheritance: (1) avoidance of costly mixing of different cytoplasmic genomes within hosts; and (2) selection for the accumulation of beneficial cytoplasmic mutations within hosts. I conclude that the need to avoid costs associated with the mixing of cytoplasmic genomes has the strongest support of any existing hypothesis. Irrespective of the evolutionary reasons behind uniparental inheritance, this mode of inheritance has implications for the spread and evolution of cytoplasmic genomes, which is the focus of the remainder of my thesis. Cytoplasmic genomes are asexual and generally lack recombination. Both theoretical and empirical work have shown that the absence of sexual reproduction and recombination should impair adaptive evolution. In fact, asexual genomes should suffer from
irreparable mutational meltdown in a process known as Muller’s ratchet. Increasingly, empirical evidence suggests that the mitochondrial genome, particularly that of animals, shows pervasive signatures of adaptive evolution despite lacking sex and recombination. In the second part of this thesis, I show that uniparental inheritance dramatically alters the evolutionary dynamics of cytoplasmic genomes, explaining why these genomes have higher levels of adaptive evolution than predicted by existing theory. I then move on to investigate the consequences of uniparental inheritance on the bacterial endosymbionts of arthropods. Many bacterial endosymbionts manipulate the reproduction of their host to promote their own spread. I show that uniparental inheritance of cytoplasm protects arthropods from invasion by harmful bacteria. This places evolutionary pressure on endosymbionts to evolve mechanisms to manipulate their host’s reproduction, explaining the pervasiveness of reproductive manipulation by the endosymbionts of arthropods.
## Contents

1 Introduction  
1.1 Cytoplasmic genomes  
1.2 Evolution of the mitochondrial genome  
1.3 Thesis outline

2 Selection against heteroplasmy explains the evolution of uniparental inheritance of mitochondria  
2.1 Abstract  
2.2 Introduction  
2.3 Basic model description  
2.4 Results  
2.5 Discussion  
2.6 Conclusion

3 Uniparental inheritance promotes adaptive evolution in cytoplasmic genomes  
3.1 Abstract  
3.2 Introduction  
3.3 Model  
3.4 Results  
3.5 Discussion

4 Can selection for beneficial mitochondrial mutations explain the evolution of uniparental inheritance?  
4.1 Abstract  
4.2 Introduction  
4.3 Model  
4.4 Results
C Chapter 4: Supplementary Material

C.1 Supplementary tables ........................................ 230
C.2 Single-locus model assuming mating types .................. 238
C.3 Two-locus model ............................................. 245
C.4 No mating types model ....................................... 251
C.5 Fluctuating environment model ................................. 254

D Chapter 5: Supplementary Material

D.1 Supplementary figures ........................................ 257
D.2 Supplementary tables ......................................... 261
D.3 Variation in endosymbiont load slows growth ................. 263
D.4 Removing oogamy, multicellularity, and
    germline .................................................... 265
D.5 Protist model .................................................. 266
D.6 Arthropod model .............................................. 269
List of Figures

2.1 Uniparental inheritance replaces biparental inheritance for all tested parameter values ........................................ 26
2.2 Fitness and distribution of cell types ........................................ 28
2.3 Fitness and distribution of gamete types ................................... 30
2.4 Recombination and no mating types scenarios ............................... 34
3.1 Fitness functions ................................................................. 52
3.2 Dynamics in the accumulation of beneficial substitutions .................... 54
3.3 Uniparental inheritance reduces clonal interference ............................. 56
3.4 Accumulation of deleterious substitutions in the absence of beneficial mutations ....................................................... 58
3.5 Genetic hitchhiking ............................................................................. 59
3.6 Inheritance mode and the distribution of genetic hitchhiking ................. 60
3.7 Uniparental inheritance promotes adaptive evolution ............................ 61
4.1 The three fitness functions ................................................................. 76
4.2 Frequency of uniparental inheritance at equilibrium ............................... 77
4.3 Relative fitness of alleles and their change in frequency per generation ......... 79
4.4 Comparing single-locus and two-locus models ..................................... 81
4.5 Frequency of uniparental inheritance at equilibrium assuming no mating types ...................................................... 82
4.6 Frequency of uniparental inheritance in a fluctuating environment .......... 84
5.1 Endosymbionts invade less readily in the arthropod model ..................... 100
5.2 Beneficial endosymbionts become fixed less frequently in the arthropod model ....................................................... 101
5.3 The arthropod model increases variation in endosymbiont load between hosts, promoting selection on hosts but reducing endosymbiont growth within hosts .......................................................... 102
5.4 The arthropod model increases selection on hosts but decreases endosymbiont growth .......................................................... 104
5.5 Oogamy, multicellularity, and a germline promote selection on hosts at the expense of endosymbiont growth .......................................................... 106
5.6 Negative covariance in endosymbionts carried by soma and germline . 108
5.7 By manipulating reproduction, harmful endosymbionts can invade arthropods .......................................................... 110
Eons ago, a proto-eukaryote engulfed an α-proteobacterium. This symbiosis led to modern-day eukaryotes, and that α-proteobacterium evolved into one of most important eukaryote organelles: the mitochondrion [1, 2]. Similarly, in plants and algae, an ancient symbiosis between an early eukaryote and a cyanobacterium led to the evolution of chloroplasts [1, 2]. These symbioses have played crucial roles in the evolution of key eukaryote innovations [3]. Mitochondria and chloroplasts are integral components of the eukaryote metabolic system, producing the majority of energy for these organisms. Mitochondrial genes encode polypeptides necessary for functioning of the electron transport chain, which is used by cells to produce energy [4]. Chloroplasts use sunlight to produce energy and store it in carbohydrates for use by plant or algae cells [5]. In addition to these organelles, other symbioses between eukaryotes and bacteria have played, and continue to play, key roles in eukaryote evolution. Many eukaryotes carry obligate endosymbionts that are important for metabolism and nutrition. For example, the protist Paulinella chromatophora contains obligate photosynthetic bacteria called chromatophores [6]. Similarly, all aphids require the obligate endosymbiont Buchnera for proper metabolic function and survival [7]. Although these endosymbionts have had disparate evolutionary histories and differ from each other in details, they share many key features.

1.1 Cytoplasmic genomes

Mitochondria, chloroplasts, and obligate endosymbionts share a number of important traits. All are nested within the cytoplasm of a host, and each contains a genome
1.1. Cytoplasmic genomes

derived from its bacterial ancestor (indeed, I will refer to them generically as cytoplasmic genomes) [8, 9]. These genomes have all degenerated compared to their free-living relatives—as genes are exported to the nucleus or lost [10, 11]—and play crucial metabolic roles in their hosts.

1.1.1 Levels of selection

Each eukaryote cell carries multiple cytoplasmic genomes. In turn, each of these genomes can itself carry multiple cytoplasmic DNA molecules. In mitochondria and chloroplasts, these DNA molecules are packaged into structures called nucleoids [4, 12], while obligate endosymbionts are found in specialized structures called bacteriocytes in some insects [7]. For a de-novo mutation in a cytoplasmic genome to become fixed in a population, it must therefore traverse three levels: the nucleoid or bacteriocyte, the host, and the population [13]. Because of their multiple copy number and their sub-structuring within hosts, cytoplasmic genomes may experience selection at two important levels: within-host and between-host. While between-host selection is dictated by a cytoplasmic genome’s effect on host fitness, within-host selection is dictated by its replication rate [13]. The presence of two types of genomes within the single cell—the host’s nuclear genome and multiple cytoplasmic genomes—creates conditions for cooperation and competition. Within-host and between-host selection can act in the same direction. Such a scenario would occur if a mutation in a cytoplasmic genome were to increase both the replication rate of the cytoplasmic genome and the fitness of its host, or alternatively, to decrease both [14]. But the fitness of the cytoplasmic genome need not align with that of its host. Indeed, selection for mutations that improve the cytoplasmic genome’s replication, but lower the fitness of the host, can lead to cytoplasmic genomes that act “selfishly”.

1.1.2 Control over replication

The strength of within-host selection experienced by a cytoplasmic genome depends primarily on how much control the genome has over its own replication. Exactly where the balance between host and endosymbiont control over replication lies is difficult to determine and depends on the type of cytoplasmic genome. From the host’s perspective, minimizing within-host selection can prevent cytoplasmic genomes from acting selfishly. By controlling the replication of a cytoplasmic genome, a host can align the fitness of its cytoplasmic genome with its own fitness [11]. Indeed, the older the association between host and cytoplasmic genome, the more control the
host has over the genome’s replication. In the most recent symbioses, such as the facultative endosymbionts of arthropods, hosts appear to have little control over the replication of the endosymbiont [15, 16]. In older symbioses, such as the obligate Buchnera, endosymbionts have lost some of the genes required for regulation of cell division and growth and hosts seem to control endosymbiont titre [17–19]. Organelles, such as the mitochondrion, have taken this to the extreme, losing all but a handful of the many genes required for their own biosynthesis [4]. As a result, regulation of mitochondrial copy number and division appears to be primarily under the control of the host [20, 21], although the exact mechanisms are not well understood. There are, however, some known examples of “selfish” mitochondrial genomes that can subvert host control over mitochondrial replication [22–25]. In fact, conflict between within- and between-host selection has inspired the dominant explanation for the evolution of arguably the most striking trait of cytoplasmic genomes: uniparental inheritance.

1.1.3 Uniparental inheritance

Unlike the nuclear genome, cytoplasmic genomes are generally inherited from a single parent. In animals, the mother almost always transmits cytoplasmic genomes so this trait is known as maternal inheritance. But, as this form of inheritance also occurs in organisms that lack sexes, the general pattern is known as uniparental inheritance [8]. Four decades ago, Grun proposed that the uniparental inheritance of cytoplasmic organelles could have evolved because it promotes selection against “selfish” organelles that harm their host [26]. While it was initially formulated as a “for the good of the population” argument [27, 28], this idea was reframed in evolutionary terms in a series of mathematical modelling papers in the early 1990s [27–31]. These models assumed that the ancestral population transmitted cytoplasmic genomes via both parents. They then tested whether a trait that forced inheritance through a single parent could spread in this ancestral population. While some of these models focused exclusively on the evolution of uniparental inheritance [27–29], others jointly examined the evolution of uniparental inheritance and the evolution of mating types or sexes [30, 31]. These papers concluded that selfish cytoplasmic genomes could indeed have driven the evolution the uniparental inheritance. Over the next two decades, aside from a couple of exceptions that came to similar conclusions as previous work [32, 33], the modeling stopped. The issue was settled: the evolution of uniparental inheritance had been driven by selfish cytoplasmic genomes.

But had it? With the exception of one study [29], for simplicity all models had ignored the possibility that cells could carry more than one type of cytoplasmic genome
(a condition known as heteroplasmy) [27, 28, 32, 33]. Rather than explicitly model heteroplasmy, these models assumed that heteroplasmic cells always produced homoplasmic offspring (i.e. those carrying a single type of cytoplasmic genome). Some studies went further, both ignoring heteroplasmy and assuming that biparental inheritance had a fixed cost relative to uniparental inheritance [30, 31]. Only one study, that of Hastings [29], explicitly allowed cells and gametes to carry multiple types of cytoplasmic genomes. The more realistic model of Hastings produced substantially different dynamics than the other models. Hastings showed that, by ignoring heteroplasmy, the simplified models had obscured the dominant dynamic underlying the spread of a trait for uniparental inheritance due to conflict with selfish genomes. Uniparental inheritance was, in fact, so efficient at purging selfish cytoplasmic genomes that it experienced strong negative frequency-dependent selection [29]. When uniparental inheritance reached a moderate frequency in the population, selfish mutants were purged from the population and biparental inheritance was no longer costly (recall that biparental inheritance is assumed to be the ancestral state). Hasting’s work did not show that uniparental inheritance replaced biparental inheritance, but rather that it coexisted with biparental inheritance at equilibrium; in fact, the allele for uniparental inheritance never exceeded a frequency of 20% in the population, far from the strict uniparental inheritance seen in most organisms [29].

It was not until 2013 before Hadjivasiliou and colleagues picked up the initial problems identified by Hastings with respect to the selfish cytoplasmic genome hypothesis [34]. In particular, they pointed out that Hasting’s findings [29] implied that the assumption of a fixed cost of biparental inheritance made by previous authors was unrealistic. If uniparental inheritance is subject to negative frequency-dependent selection, then it cannot have a fixed advantage over biparental inheritance. Hadjivasiliou and colleagues confirmed Hasting’s finding [29] that uniparental inheritance was subject to negative frequency-dependent selection. Furthermore, they showed that uniparental inheritance could replace the ancestral condition of biparental inheritance but only when four conditions were met: (1) selfish genomes arise extremely frequently; (2) selfish genomes replicate much faster than wild type genomes; (3) hosts die when they carry only selfish genomes; and (4) hosts carry a large number of cytoplasmic genomes [34]. These conditions are required to override the negative frequency-dependent selection against uniparental inheritance. The most likely candidate to satisfy the first three conditions is the petite mutation in the mitochondria of the yeast *Saccharomyces cerevisiae*, which prevents yeast from undergoing aerobic respiration. The petite mutation arises frequently, mitochondria that carry the
petite mutation can out-compete wild type mitochondria in the same host, and the mutation knocks out mitochondrial function [22, 35]. But yeast only carry 1–10 mitochondria [36, 37], violating the fourth assumption, which requires > 50 mitochondria per cell. Overall, these conditions are so restrictive that it is unlikely any extant organism satisfies these criteria.

Although the selfish, or conflict, hypothesis has received the most attention as an evolutionary explanation for uniparental inheritance, other hypotheses exist. One proposes that uniparental inheritance facilitates selection against deleterious mutations in cytoplasmic genomes, thereby improving a host’s fitness due to its association with healthy cytoplasmic genomes [34]. Another proposes that uniparental inheritance facilitates coadaptation between the host’s genome and the genome of its cytoplasmic organelles, as both genomes are required for metabolism [38, 39]. When tested using mathematical models, however, these alternative hypotheses require even more restrictive conditions than the selfish genome hypothesis [34]. Since deleterious mutations in cytoplasmic genomes do not have the within-host replication advantage of selfish mutations, they require an extraordinarily high mutation rate to drive the evolution of uniparental inheritance [34]. Coadaptation between the genomes of the host and cytoplasmic organelles is unable to drive the evolution of uniparental inheritance under any set of assumptions [34]. Thus, the forces driving the evolution of uniparental inheritance are still very much unknown. The widespread occurrence of uniparental inheritance suggests that this trait evolved for an adaptive purpose [8]; yet so far we lack a satisfying explanation for what that adaptive purpose might be.

1.2 Evolution of the mitochondrial genome

The parallels between mitochondria, chloroplasts, and obligate endosymbionts discussed in section 1.1 mean that theory developed for one cytoplasmic genome will often broadly apply to the other cytoplasmic genomes. Of all cytoplasmic genomes, the mitochondrial genome is the most well-studied. Since models are most informative when their assumptions and parameters are guided by biological evidence, I will focus on the mitochondrial genome to illustrate general principles of cytoplasmic genomes.

1.2.1 Mitochondrial genome structure

The mitochondrial genome (mtDNA) encodes proteins that, along with a subset of nuclear proteins, form a series of complexes that the cell uses to produce energy
1.2. Evolution of the mitochondrial genome

via oxidative phosphorylation [4]. The mitochondrial genome generally consists of a single, circular molecule [4], although some species contain linear mtDNA molecules or even multiple, fragmented mtDNA molecules [40, 41]. The size and structure of the mitochondrial genome differs substantially between plants, protists, fungi, and animals. Plant mtDNA genomes vary dramatically in size, ranging from about 200 kb in white mustard [42] to about 11 Mb in the angiosperm *Silene* [43]). Plant mtDNA genomes contain introns and extensive intergenic sequences, experience horizontal transfer [44], and undergo substantial recombination [45]. The mtDNA genomes of protists and fungi vary widely between species [46], but they contain smaller mtDNA genomes than found in plants. Yeast have mtDNA genomes around 25–85 kb [47] while a few fungi have mtDNA genomes that exceed 150 kb [48]. The mtDNA of the protist *Reclinomonas americana* is gene-rich, containing almost 100 genes over its 70 kb length [46].

Animal mtDNA genomes, however, are small and streamlined (15–20 kb) and are remarkably similar in structure [49], considering the diversity found in plants, protists, and fungi. Animal mtDNA genomes do not appear to undergo horizontal transfer, and rarely or never undergo recombination [50]. In addition, these genomes have much higher mutation rates than seen in plants [43, 51]. Human mtDNA, for example, is about 16.6 kb in length and contains 37 genes—13 polypeptides, 2 ribosomal RNAs, and 22 transfer RNAs (tRNAs) [4]. Animal mtDNA also contains a non-coding region called the D-loop that is involved in the regulation of mitochondrial replication and transcription [4].

Unfortunately, no single body of literature can account for the vastly different evolutionary trajectories followed by the mitochondrial genomes of animals, plants, fungi, and protists. For the parts of this thesis concerned with genome evolution, I will focus primarily on the trajectory followed by animal mtDNA, although some of the insights will no doubt apply to other groups of organisms. The remarkably similar structure of mtDNA between animal species means that theory will apply broadly to many well-studied organisms. The unexpected evolutionary trajectory of the animal mitochondrial genome also makes it a very interesting case to examine.

1.2.2 Evolutionary pressures on the mitochondrial genome

Like its bacterial ancestor, the animal mitochondrial genome is asexual and undergoes little to no recombination [8, 50]. This has important implications for the evolution of the mitochondrial genome. A large body of theoretical [52–57] and empirical [58–61] literature shows that asexual genomes experience lower rates of adaptive evolution
than sexual genomes in populations of a finite size. Compared to sexual genomes, asexual genomes are less able to accumulate beneficial substitutions and less able to purge deleterious substitutions [52–55].

The lack of recombination means that independent beneficial substitutions on different genomes will compete with one another for fixation. This process, known as clonal interference, leads to a loss of beneficial mutations [52–55, 58, 59, 62, 63]. Without recombination, beneficial substitutions also cannot be extracted and “rescued” from genomes that carry multiple linked deleterious substitutions (the latter idea is known as “a ruby in the rubbish”) [64].

Similarly, the lack of recombination means that asexual genomes struggle to purge deleterious substitutions due to a process known as Muller’s ratchet [55, 65]. In small populations, selection on slightly deleterious substitutions can be overwhelmed by the influence of genetic drift. Drift can cause the class of genome carrying the fewest deleterious substitutions to be randomly lost from the population. When this occurs, that class of genome can only be rescued through a back mutation, or through a compensatory mutation that restores function. As both back and compensatory mutations are rare, asexual genomes should accumulate slightly deleterious mutations over time [55, 65].

Mitochondria, as ancient asexual genomes, should suffer the above limitations of asexual reproduction. In fact, there are reasons to believe that mitochondrial genomes should be even more susceptible to clonal interference and Muller’s ratchet. Free-living asexual organisms, such as bacteria, typically have massive population sizes [66] and occasional sexual exchange [67], alleviating many of the limitations of asexual reproduction [55]. But mitochondria are nested within eukaryote hosts, which have effective population sizes (\(N_e\)) much lower than free-living bacteria [66, 68].

To make matters worse, the mtDNA is uniparentally inherited and often undergoes tight transmission bottlenecks during gametogenesis [8, 69]. In the latter case, the effective numbers of mitochondria in a cell can be reduced from thousands to tens or hundreds [69, 70]. Both uniparental inheritance and a transmission bottleneck further reduce the \(N_e\) of mitochondria [71]. The lower the \(N_e\), the higher the levels of genetic drift; the higher the levels of genetic drift, the lower the efficacy of selection on hosts [72, 73]. Thus, theory predicts that mtDNA will carry a high load of deleterious substitutions and few beneficial substitutions.
1.2.3 Accumulation of deleterious mtDNA substitutions

Some studies have indeed indicated that animal mitochondria suffer from Muller’s ratchet, leading to high levels of deleterious substitutions. The mean between-species divergence in synonymous sites in mtDNA (not altering the amino acid) is 5–50 times higher than that of nuclear DNA [74]. A similar pattern is seen in mitochondrial tRNAs, where the substitution rate is 5–20 times that found in comparative nuclear tRNAs [75, 76]. The elevated substitution rate reduces function. Mitochondria tRNA stems have an average binding stability less than half that found in their nuclear counterparts, indicating that the substitutions are largely deleterious [75, 76]. Given that the mutation rate in animal mtDNA is at least an order of magnitude higher than that of nuclear DNA [77], and that mitochondria should suffer the limitations of asexual reproduction, these observations fit nicely with theory.

Yet, a completely different story emerges when one looks at the protein-coding mtDNA genes. Mitochondrial genes coding for polypeptides involved in oxidative phosphorylation (OXPHOS) are more conserved than orthologous genes in free-living bacteria, despite the fact that bacteria have a higher $N_e$ and non-zero recombination rates [66]. Furthermore, mitochondrial OXPHOS genes are more conserved than the genes for nuclear OXPHOS polypeptides with which they interact [77]. In fact, it is estimated that mutations in mtDNA OXPHOS genes are effectively five-fold more harmful than analogous nuclear mutations [77]. This is in stark contrast to theory, which predicts that the ability of the mitochondrial genome to undergo purifying selection should be markedly reduced.

1.2.4 Accumulation of beneficial mtDNA substitutions

The search for the existence of beneficial substitutions in mtDNA was for many years eschewed in favour of deleterious substitutions, possibly due to the belief that Muller’s ratchet was the main force behind mtDNA evolution. If mtDNA cannot undergo adaptive evolution, then why look for it? But within the past decade, it has become clear that adaptive evolution in mtDNA is not only possible—it is pervasive. Anecdotal examples of adaptive evolution in animal mtDNA abound [78–86]. A decade ago, Bazin and colleagues proposed that the patterns of diversity seen in a large dataset of animal mitochondria was best explained by frequent adaptive sweeps [87]. When that dataset was reanalyzed by using a method to control for the presence of deleterious mutations, it was estimated that a quarter of non-synonymous substitutions (those altering the amino acid) were due to adaptive evolution [88]. Thus, while theory
predicts that adaptive evolution should be impaired in the mitochondrial genome, empirical evidence clearly points in the opposite direction.

## 1.3 Thesis outline

This thesis is conceptually divided into two parts: (1) the evolution of uniparental inheritance (chapters 2 and 4); and (2) the consequences of uniparental inheritance on the spread and evolution of cytoplasmic genomes (chapters 3 and 5).

In chapter 2, I propose a new hypothesis to explain the evolution of uniparental inheritance: selection against hosts that carry multiple different mitochondrial types (heteroplasmy). Like previous studies, I assume that the ancestral population inherits mitochondria from both parents. I show that uniparental inheritance can invade and completely replace biparental inheritance. Unlike the hypothesis for selfish genomes, my hypothesis results in the evolution of uniparental inheritance under an extremely wide range of realistic conditions. This hypothesis is also able to explain why some organisms do not exhibit strict uniparental inheritance. Importantly, the central assumption—that heteroplasmy imposes a cost on the organism—is backed up by strong empirical evidence in mice [89] and nematodes [90].

To ensure that my findings in chapter 2 do not depend on mitochondrial haplotypes having no effect on the host (other than a cost when mixed with an incompatible haplotype), I let mitochondrial haplotypes have an underlying detrimental, neutral, or beneficial effect on their host. While analyzing simulations that involved haplotypes with beneficial effects, I noticed that beneficial mutations increased the rate at which the trait for uniparental inheritance spread. This indicated that the mode of inheritance could alter the speed of selective sweeps—something that should have profound implications for the evolution of the genome.

Thus, in chapter 3, I examine how uniparental inheritance affected adaptive evolution in cytoplasmic genomes. In particular, I focused on three limitations of asexual reproduction: (1) inefficient spread of beneficial substitutions; (2) inability to purge deleterious substitutions; and (3) genetic hitchhiking of deleterious substitutions that are linked to a beneficial substitution that sweeps through the population. I show that uniparental inheritance increases the efficacy of selection, acting strongly in favor of hosts with beneficial substitutions and against hosts with deleterious substitutions. By improving selection on hosts, uniparental inheritance leads to rapid selective sweeps compared to selective sweeps under biparental inheritance of cytoplasmic genomes or in comparable free-living asexual genomes. As a result, uniparental
inheritance promotes the spread of beneficial substitutions, acts to purge deleterious substitutions, and reduces the hitchhiking of deleterious substitutions on the back of selective sweeps. My work in chapter 3 thus explains why mitochondrial genomes have much higher levels of adaptive evolution than predicted by existing theory.

In chapter 3, I show that uniparental inheritance increases the rate at which beneficial substitutions accumulate within hosts. In doing so, uniparental inheritance should increase the fitness of hosts that carry the trait. This raises the intriguing possibility that uniparental inheritance’s ability to promote the spread and accumulation of beneficial substitutions could have played a key role in its evolution.

So, in chapter 4, I revisit the evolution of uniparental inheritance. I explore whether selection for the accumulation of beneficial substitutions is sufficient to explain the evolution of uniparental inheritance. This hypothesis has two main strengths: (1) there is abundant empirical evidence for adaptive evolution in the mitochondrial genome, providing a strong foundation for the fundamental tenet; and (2) unlike hypotheses for selfish or deleterious mutations, which require unrealistic driving forces (e.g. mutation rates), beneficial mutations can drive themselves by providing a selective advantage for their host. I show that uniparental inheritance can indeed invade a population in which the ancestral model of inheritance is biparental. But surprisingly, despite the benefits that uniparental inheritance has on hosts, it can only replace biparental inheritance when the mtDNA haplotype that is under positive selection constantly changes due to a variable environment (e.g. due to fluctuating climatic conditions). While this hypothesis compares favourably to the selfish genome hypothesis, it falls short of selection against heteroplasmy as a general explanation for the evolution of uniparental inheritance.

In chapter 5, I focus on factors that affect the initial uptake and spread of an endosymbiont. Symbioses between eukaryotes and bacteria are widespread. Some of the facultative symbioses that occur today will undoubtedly become the obligate symbioses of tomorrow and the novel organelles of the future. Understanding the initial uptake of endosymbionts is thus crucial to explaining how obligate endosymbionts and organelles evolve. One particularly interesting trait that has evolved in some endosymbionts is reproductive manipulation of their hosts. Endosymbionts that manipulate their host’s reproduction are common in arthropods but appear to be rare in more primitive eukaryotes, such as protists. In chapter 3, I show that uniparental inheritance strengthens selection on hosts. While chapter 3 was concerned with genome evolution, the efficacy of selection will also affect the spread of a novel endosymbiont. In chapter 5, I argue that reproductive manipulation is especially
common in arthropods because of the combined effects of uniparental inheritance of endosymbionts (oogamy), multicellularity, and the soma-germline separation. I show that beneficial endosymbionts can easily invade and spread in protists and arthropods. While harmful endosymbionts easily invade protists, such endosymbionts are completely unable to invade arthropods unless they evolve traits to manipulate their host’s reproduction.

Finally, in chapter 6 I synthesize my findings and suggest directions for future research. Chapters 2–5 are formatted for submission to a journal, and these chapters are either published (chapter 2), accepted pending satisfactory revision (chapter 3), submitted (chapter 4), or soon to be submitted (chapter 5).
1.3. Thesis outline

1.3.1 Publication status and author contributions

Chapter 2

Conceived and designed the experiments: JRC, MB. Performed the experiments: JRC. Analyzed the data: JRC. Wrote the paper: JRC, MB. Developed the model: JRC, TMS. Wrote the model description: JRC, TMS.

Chapter 3


Conceived and designed the experiments: JRC. Performed the experiments: JRC. Analyzed the data: JRC. Wrote the paper: JRC, MB.

Chapter 4

Conceived and designed the experiments: JRC. Performed the experiments: JRC. Analyzed the data: JRC. Wrote the paper: JRC, MB.

Chapter 5
Joshua R. Christie, Madeleine Beekman. Oogamy, multicellularity, and germline protect arthropods against harmful endosymbionts. *To be submitted.*

Conceived and designed the experiments: JRC. Performed the experiments: JRC. Analyzed the data: JRC. Wrote the paper: JRC, MB.

1.3.1.1 Other work produced during thesis period

Bibliography


[48] Salavirta H et al. (2014) Mitochondrial genome of *Phlebia radiata* is the second largest (156 kbp) among fungi and features signs of genome flexibility and recent recombination events. *PLoS ONE* 9(5):e97141. [1.2.1]


[65] Muller HJ (1964) The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 1(1):2–9. [1.2.2, 2.6, 3.2, 3.4.6, 4.2]


[86] Rollins LA et al. (2016) Selection on mitochondrial variants occurs between and within individuals in an expanding invasion. *Molecular Biology and Evolution*. [1.2.4, 6.2]


Chapter 2

Selection against heteroplasmy explains the evolution of uniparental inheritance of mitochondria

2.1 Abstract

Why are mitochondria almost always inherited from one parent during sexual reproduction? Current explanations for this evolutionary mystery include conflict avoidance between the nuclear and mitochondrial genomes, clearing of deleterious mutations, and optimization of mitochondrial-nuclear coadaptation. Mathematical models, however, fail to show that uniparental inheritance can replace biparental inheritance under any existing hypothesis. Recent empirical evidence indicates that mixing two different but normal mitochondrial haplotypes within a cell (heteroplasmy) can cause cell and organism dysfunction. Using a mathematical model, we test if selection against heteroplasmy can lead to the evolution of uniparental inheritance. When we assume selection against heteroplasmy and mutations are neither advantageous nor deleterious (neutral mutations), uniparental inheritance replaces biparental inheritance for all tested parameter values. When heteroplasmy involves mutations that are advantageous or deleterious (non-neutral mutations), uniparental inheritance can still replace biparental inheritance. We show that uniparental inheritance can evolve with or without pre-existing mating types. Finally, we show that selection against heteroplasmy can explain why some organisms deviate from strict uniparental inheritance. Thus, we suggest that selection against heteroplasmy explains the evolution
2.2 Introduction

During sexual reproduction, offspring receive two genomes: nuclear genomes from both parents and haploid cytoplasmic genomes, contained in mitochondria and chloroplasts (in plants and algae), usually from one parent. Although uniparental inheritance is nearly ubiquitous, the reasons behind its evolution remain unresolved [1, 2]. Cells contain multiple mitochondria, and the mitochondrial genome (mtDNA) encodes polypeptide subunits of the electron transport chain, which the cell uses to generate ATP via oxidative phosphorylation [2]. If mutations increase mtDNA replication rate but simultaneously decrease respiration, then increased mtDNA fitness comes at the expense of cell and organism fitness [3–5]. Nuclear and mitochondrial genomes are thus potentially in conflict. The genomic (or selfish) conflict theory argues that uniparental inheritance evolved because biparental inheritance facilitates the spread of such selfish mitochondria [1, 3–6]. Although the conflict theory has been the predominant explanation for uniparental inheritance for over three decades [3, 4], other explanations exist. A second theory suggests that uniparental inheritance facilitates the removal of deleterious mutations. Uniparental inheritance decreases variation of mtDNA within cells, but increases variation between cells, allowing purifying selection against cells with increased mutation load [1, 7]. A third hypothesis argues that because the oxidative phosphorylation pathway is composed of interacting nuclear- and mitochondrial-encoded polypeptides, uniparental inheritance optimizes mitochondrial-nuclear coadaptation by maintaining coevolved mitochondrial-nuclear combinations [1, 8]. While uniparental inheritance spreads in mathematical models of the above hypotheses [1, 5, 6], it cannot replace biparental inheritance under realistic assumptions and parameter values [1, 5]. Thus, despite decades of theoretical work, we still lack a convincing explanation for why uniparental inheritance is widespread amongst extant organisms [1, 2].

Although uniparental inheritance is the general rule in eukaryotes, there are a few exceptions. Probably the best-known exception is baker's yeast (Saccharomyces cerevisiae) in which both parents contribute mitochondria to offspring [9, 10]. However, the repeated division of cells that contain two mitochondrial lineages (heteroplasmy) leads to cells that contain a single type of mitochondria (homoplasmy) [9, 10]. Another example is the male bivalve (Mytilus), which also inherits mitochondria from both parents. But in this case maternal and paternal mitochondria do not mix within
single cells, as maternal mitochondria segregate to the soma while paternal mitochondria segregate to the gonads [11]. Thus, even when mitochondria are inherited from both parents, heteroplasmy is avoided. Recent experimental evidence suggests that this is because heteroplasmy imposes a cost on the organism. A study on mice found that the mere mixing of different, but phenotypically normal, mitochondria within a cell leads to physiological and behavioral abnormalities [12]. Could uniparental inheritance have evolved simply because carrying multiple mitochondrial types imposes a cost on the organism? Here we use a mathematical model to explore whether selection against heteroplasmy could have led to the evolution of uniparental inheritance.

2.3 Basic model description

Our model is based on an idealized life cycle of a single-cell diploid eukaryotic organism, such as the algae *Chlamydomonas reinhardtii*. Diploid cells contain *n* mitochondria and haploid cells have *n/2* mitochondria. All mitochondria are initially wild type but mitochondria can mutate from wild type to mutant (and vice versa). The starting population contains haploid gametes with a nuclear allele regulating biparental inheritance (*B*). Gametes are evenly split between two nuclear self-incompatible mating types (*B*1 and *B*2). In the basic model, we assume no recombination between the mitochondrial inheritance and mating type loci because these are tightly linked in many isogamous organisms [9] (later we explore recombination and no mating types). Cell types are characterized by the proportion of wild type and mutant mitochondria that they carry and their nuclear allele (haploid) or genotype (diploid).

Our life cycle has four discrete stages and is similar to the life cycles used in previous models [1, 5, 8]. Since we begin with a population of gametes, the first stage is random mating. Here, gametes randomly mate with the opposite mating type to produce diploid cells. Matings are controlled by the nuclear allele in gametes. In biparental inheritance (between *B*1 and *B*2 gametes), both gametes contribute mitochondria to the *B*1*B*2 diploid cells (see later for uniparental inheritance). The second stage is mutation. Each mitochondrion can mutate to the other haplotype with probability *µ*. The third stage is selection. Here, diploid cells have a relative fitness based on the proportion of each haplotype in the cell. We assume that fitness decreases as the level of heteroplasmy increases (here we consider the highest level of heteroplasmy to be when cells contain 50% mutant mitochondria). The fourth stage is meiosis, where diploid cells produce gametes that contain a single nuclear allele and
n/2 mitochondria. As mitochondria are stochastically partitioned into gametes [9], diploid heteroplasmic cells produce gametes with varying degrees of heteroplasmy.

First, we let the population of $B_1$ and $B_2$ gametes reach mutation-selection equilibrium. We then simulate a mutation leading to uniparental inheritance of mitochondria by converting a small proportion ($10^{-2}$) of $B_1$ gametes to $U_1$ gametes. We assume no further mutations between $B$ and $U$ alleles. Matings between $U_1$ and $B_2$ gametes result in uniparental inheritance, in which the cell inherits mitochondria from $U_1$ alone. (Matings between $U_1$ and and $B_1$ are not possible as they are the same mating type.) The population now consists of three alleles ($U_1$, $B_1$, and $B_2$) and two genotypes ($U_1B_2$ and $B_1B_2$). The model tracks the proportion of each cell type at each stage of the life cycle. $U_1$ spreads at the expense of $B_1$ when uniparental inheritance is more advantageous than biparental inheritance (the frequency of $B_2$ always remains at 0.5), and the simulation ends when the alleles reach equilibrium (see section A.5–section A.11 for details of the model).

To explore whether a cost to heteroplasmy could have led to the evolution of uniparental inheritance, we study several scenarios. We first examine the simplest case, where mutations in mitochondria are neither advantageous nor disadvantageous (neutral mutations), but heteroplasmic cells incur a fitness cost proportional to the degree of heteroplasmy. Because no empirical data relate fitness to the degree of heteroplasmy, we consider three forms of fitness function to describe selection against heteroplasmy: concave, linear and convex (Figure 2.1A). For each fitness function, we vary the cost of heteroplasmy ($c_h$), given by $c_h = 1 - h$ where $h$ is the fitness of the most heteroplasmic cell in the population, to see how this affects the spread of $U_1$. We generate the concave fitness function by

$$
\omega(i) = \begin{cases} 
1 - c_h \left( \frac{i}{n/2} \right)^2 & \text{for } 0 \leq i < n/2, \\
1 - c_h \left( \frac{n-i}{n/2} \right)^2 & \text{for } n/2 \leq i \leq n,
\end{cases}
$$

the linear function by

$$
\omega(i) = \begin{cases} 
1 - c_h \left( \frac{i}{n/2} \right) & \text{for } 0 \leq i < n/2, \\
1 - c_h \left( \frac{n-i}{n/2} \right) & \text{for } n/2 \leq i \leq n,
\end{cases}
$$

and the convex function by

$$
\omega(i) = \begin{cases} 
1 - c_h \sqrt{\frac{i}{n/2}} & \text{for } 0 \leq i < n/2, \\
1 - c_h \sqrt{\frac{n-i}{n/2}} & \text{for } n/2 \leq i \leq n,
\end{cases}
$$
where $i$ represents the number of mutant mitochondria.

We also vary $\mu$ (mutation rate) and $n$ (number of mitochondria) to ensure that our findings are robust. Second, we explore the effect of advantageous or deleterious mutations (non-neutral mutations) on the spread of $U_1$. Third, we relax the assumption of tight linkage between mating type and inheritance loci by exploring two cases: recombination between mating types and the absence of mating types altogether. Finally, we examine whether selection against heteroplasmy can explain the rare, but nevertheless important, exceptions to uniparental inheritance. To ensure that our results generalize to more than two mitochondrial types, we developed a second model that considers three mitochondrial types (section A.11).

![Figure 2.1](image-url)  
Figure 2.1: Uniparental inheritance replaces biparental inheritance for all tested parameter values. A. The three fitness functions when $c_h = 1$. Unless indicated otherwise, the parameters for B–F are $n = 20$, $\mu = 10^{-7}$, $c_h = 0.2$, and concave fitness. B. $U_1$ replaces $B_1$. C. $U_1$ takes longer to replace $B_1$ as $n$ increases. D. $U_1$ takes longer to replace $B_1$ as $\mu$ decreases. E. $U_1$ replaces $B_1$ under all three fitness functions. F. Number of generations for $U_1$ to replace $B_1$ across a range of costs of heteroplasmy. $U_1$ replaces $B_1$ even if the cost of heteroplasmy is extremely low.

### 2.4 Results

#### 2.4.1 When both mitochondrial haplotypes are neutral

We find that $U_1$ always replaces $B_1$, resulting in complete uniparental inheritance in the population (Figure 2.1B). These findings are independent of the number of mitochondria per cell (Figure 2.1C), mutation rate (Figure 2.1D), fitness function (Figure 2.1E), and cost of heteroplasmy (Figure 2.1F) (see Table A.1–Table A.10 for
more parameter combinations). We find the same results when we generalize the model to three mitochondrial haplotypes (Figure A.1).
Figure 2.2: Fitness and distribution of cell types. Parameters: $n = 20$, $\mu = 10^{-4}$, $c_h = 0.2$, and concave fitness. $U_1B_2$ cells appear at generation 0, which is the point at which the $B_1$ and $B_2$ gametes reach mutation-selection equilibrium. A. Relative advantage of each genotype through time (see section A.5.6 for details). For B–E, the relative proportion is the sum of a particular cell type divided by the sum of all cells that carry the same genotype. The heteroplasmic category includes all cells with any level of heteroplasmy. B–C shows the distribution of cells carrying the $U_1B_2$ genotype (B) and the $B_1B_2$ genotype (C). D–E show a more detailed distribution of cell types carrying the $B_1B_2$ genotype at generation 1350 (D) and at generation 1820 (E). The decrease in heteroplasmy in $B_1B_2$ cells between generations 0-100 is an artifact of introducing $U_1$ at a frequency of 0.01 (the influx of $U_1$ gametes homoplasmic for the wild type haplotype converts some heteroplasmic $B_1$ and $B_2$ gametes into homoplasmic gametes, which increases the proportion of homoplasmic $B_1B_2$ cells). From generations 1350-1820, the proportion of heteroplasmic $B_1B_2$ cells decreases (C) but the level of heteroplasmy increases (compare D with E). This more than offsets the decrease in the proportion of heteroplasmic cells and $\bar{w}_{B_1B_2}$ continues to decrease (A).
2.4.2 General patterns

In our model, heteroplasmic cells are generated by mutation. During meiosis, heteroplasmic cells produce gametes with varying levels of heteroplasmy, including homoplasmic gametes. Uniparental inheritance maintains this variation created by meiosis, which leads to homoplasmic \( U_1B_2 \) cells (Figure 2.2A–B and Figure A.2A–B). Mutants that arise in \( U_1B_2 \) cells quickly segregate into \( U_1 \) gametes that carry mutant haplotypes only (Figure 2.3A–B and Figure A.3A–B), which leads to \( U_1B_2 \) cells that are homoplasmic for mutant mitochondria (Figure 2.2B and Figure A.2B). Since we assume that mutations are neutral, cells homoplasmic for mutant mitochondria suffer no fitness costs.

\( U_1B_2 \) cells carrying mutant mitochondria produce \( B_2 \) gametes that also carry mutant mitochondria (Figure 2.3D and Figure A.3D). When these \( B_2 \) gametes mate with \( B_1 \) gametes carrying wild type mitochondria, the resulting \( B_1B_2 \) cells are highly heteroplasmic (Figure 2.2C–E and Figure A.2C). As \( U_1 \) spreads, matings between \( U_1 \) and \( B_2 \) become more likely, increasing the level of heteroplasmy in both \( B_1B_2 \) cells and in \( B_1 \) and \( B_2 \) gametes (Figure 2.2C–E, Figure 2.3C–F, Figure A.2C, and Figure A.3C–D). Increased levels of heteroplasmy reduce the fitness of both \( B_1 \) and \( B_2 \) gametes (\( \bar{\omega}_{B_1}, \bar{\omega}_{B_2} \) in Figure 2.3A and Figure A.3A) and \( B_1B_2 \) cells (\( \bar{\omega}_{B_1B_2} \) in Figure 2.2A and Figure A.2A). The difference in fitness between \( B_1 \) and \( B_2 \) becomes stronger (Figure 2.3A and Figure A.3A) as more \( B_2 \) gametes that carry mutant mitochondria are produced (Figure 2.3D and Figure A.3D). As a result \( U_1 \) spreads at the expense of \( B_1 \).

In the above description (Figure 2.2 and Figure 2.3), the mutation from \( B_1 \) to \( U_1 \) occurred in gametes homoplasmic for wild type mitochondria. When \( U_1 \) is introduced into heteroplasmic gametes, it takes fewer generations to reach equilibrium because \( B_2 \) gametes homoplasmic for mutant mitochondria are produced more quickly (Figure A.4). Our results are robust to changes in the frequency at which \( U_1 \) gametes are introduced (Figure A.5). For more detailed model dynamics, see section A.3 and S1–S2 Videos.
Figure 2.3: **Fitness and distribution of gamete types.** Parameters: $n = 20$, $\mu = 10^{-4}$, $c_h = 0.2$, and concave fitness. $U_1$ gametes appear at generation 0, which is the point at which the $B_1$ and $B_2$ gametes reach mutation-selection equilibrium. **A.** Relative advantage of each gamete through time (see section A.5.7 for details). For **B–F,** the relative proportion is the sum of a particular gamete type (e.g. a homoplasmic wild type $U_1$ gamete) divided by the sum of all cells carrying that allele (all gametes carrying the $U_1$ allele). Thus, the relative proportion describes how an allele is distributed across different gamete types but it does not show their actual frequencies in the population. The heteroplasmic category combines all gametes with any level of heteroplasmy. **B–D** show the distribution of gametes carrying the $U_1$ allele ($B$), $B_1$ allele ($C$) and the $B_2$ allele ($D$). **E–F** show a more detailed distribution of gametes carrying the $B_1$ allele at generation 1350 (E) and generation 1820 (F). The decrease in heteroplasmy in $B_1$ and $B_2$ gametes between generations 0–100 is an artifact of introducing $U_1$ at a frequency of 0.01 (the influx of $U_1$ gametes homoplasmic for the wild type haplotype converts some heteroplasmic $B_1$ and $B_2$ gametes into homoplasmic gametes). From generations 1350–1820, the proportion of heteroplasmic $B_1$ and $B_2$ gametes decreases ($C$ and $D$) but the level of heteroplasmy increases (compare $E$ with $F$). This more than offsets the decrease in the proportion of heteroplasmic cells and $\bar{\omega}_{B_1}$ continues to decrease ($A$). Around generation 1350, $B_2$ gametes homoplasmic for mutant mitochondria begin to appear, which causes $\bar{\omega}_{B_2}$ to increase and eventually converge with $\bar{\omega}_{U_1}$. 
2.4.3 The effect of varying parameters

$U_1$ spreads more slowly when mutation rate ($\mu$) is lower (Figure 2.1D) and number of mitochondria ($n$) is higher (Figure 2.1C). Reducing $\mu$ slows the spread of $U_1$ because mutant mitochondria are produced more slowly, slowing the generation of $B_2$ gametes that only carry the mutant haplotype. Increasing $n$ has the same effect.

While varying the cost of heteroplasmy does not change the qualitative behavior of the model, it does affect the number of generations required for $U_1$ to replace $B_1$ (Figure 2.1F). In general, $U_1$ spreads more quickly when the cost of heteroplasmy is low for all three fitness functions (Figure 2.1F). Strong selection against heteroplasmy (e.g. $c_h = 1$) slows the production of $B_2$ gametes homoplasmic for the mutant haplotype because a transition via heteroplasmy is needed to lead to $U_1B_2$ cells homoplasmic for mutant mitochondria. Heteroplasmy levels thus remain low in $B_1B_2$ cells, and $U_1$ takes longer to replace $B_1$ (Figure A.6A,D). At lower costs of heteroplasmy (e.g. $c_h = 0.2$), more $B_2$ gametes that are homoplasmic for the mutant haplotype are produced and levels of heteroplasmy in $B_1B_2$ cells increase, leading to a faster spread of $U_1$ (Figure A.6B,E). Although levels of heteroplasmy in $B_1B_2$ cells increase even further as the cost of heteroplasmy approaches 0 (e.g. $c_h = 0.01$), selection against heteroplasmy is now very weak, which slows the spread of $U_1$ compared with $c_h = 0.2$ (Figure A.6C,F). When the number of mitochondria is higher, $U_1$ spreads more quickly when the cost of heteroplasmy is low. This is because $B_2$ gametes homoplasmic for mutant mitochondria are produced more slowly at higher values of $n$ and strong selection against heteroplasmy compounds this problem (Figure A.7). A similar logic can be applied to understand the differences between the three fitness functions. Since heteroplastic cells are under weaker selection when fitness is concave (followed by linear and convex respectively) (Figure 2.1A), the level of heteroplasmy is highest using a concave function (Figure A.8). Thus, $U_1$ spreads more quickly using a concave function (followed by linear and convex respectively) when the cost of heteroplasmy is high because it is easier to generate heteroplastic cells, and thus easier to generate $B_2$ gametes homoplasmic for mutant mitochondria, when selection against heteroplastic cells is weaker (Figure 2.1F and Figure A.8). As the cost of heteroplasmy decreases, the number of generations for $U_1$ to spread under the three fitness functions converges because it becomes easier to generate $B_2$ gametes homoplasmic for mutant mitochondria (Figure 2.1F).
2.4.4 When mutations are deleterious

We next investigate how the $U_1$ allele spreads when mutations are non-neutral, as is the case for most mtDNA mutations [13]. We start by assuming that mutations are deleterious so that cells carrying mutant mitochondria are more strongly selected against than cells that carry wild type mitochondria. We assume that a mutation from wild type to mutant haplotype is more common than the reverse [14]. We let the probability of a mutation from mutant to wild type haplotype be $\mu_b = \mu/100$. We vary the selection coefficient of the mutant haplotype to see how this affects the spread of the $U_1$ allele (the fitness of a cell homoplasmic for the mutant haplotype is $1 - s_d$, where $s_d$ is the selection coefficient of the mutant haplotype). Essentially there are now two fitness functions: one governing the effect of mitochondria on cell fitness (where the selection coefficient determines the magnitude of the effect) and one governing the cost of heteroplasmy. For deleterious mutations, we assume that fitness decreases as a concave function of the number of mutants, as this relationship is experimentally established [15]. We examine both concave and convex fitness functions for selection against heteroplasmy (yielding two combinations).

Again, $U_1$ replaces $B_1$ unless the fitness of heteroplasmic cells and the fitness of deleterious mutants are governed by a concave function and the selection coefficient is sufficiently large (Figure A.9, Table A.11, and Table A.12). $U_1$ generally spreads more slowly as $s_d$ increases and it always spreads more slowly compared to when mutations are neutral (Table A.11 and Table A.12). Stronger selection against mutant haplotypes leads to fewer $B_2$ gametes homoplasmic for mutant mitochondria, which slows the spread of $U_1$ (Figure A.10).

2.4.5 When mutations are advantageous

Next we explore the effect of advantageous mutations on the spread of $U_1$. In this case, cells that carry mutant haplotypes have an advantage over those carrying wild type haplotypes (the fitness of a cell homoplasmic for the wild type haplotype is $1 - s_a$, where $s_a$ is the selection coefficient of the mutant haplotype). We account for the rarity of advantageous mutations by setting $\mu_b = 100\mu$. Because it is unknown how fitness relates to the accumulation of advantageous mtDNA mutations, we model this relationship with both a concave and convex function. As in the deleterious case, we model selection against heteroplasmy by testing both concave and convex fitness functions (giving four combinations).
2.4. Results

\[ U_1 \] always replaces \( B_1 \) unless mutations are highly advantageous (\( s_a = 0.1 \)) and both the fitness of heteroplasmic cells and the fitness of advantageous mutants are governed by a concave function (Figure A.9, Table A.13, and Table A.14). \( U_1 \) spreads more quickly when \( s_a = 0.001 \) and \( s_a = 0.01 \) because \( B_2 \) gametes homoplasmic for mutant haplotypes now have a fitness advantage and are produced more quickly (Figure A.10). In contrast, \( U_1 \) spreads more slowly when \( s_a = 0.1 \) because the mutant haplotype quickly replaces the wild type as the dominant haplotype before \( U_1 \) has replaced \( B_1 \). Once \( B_1 \) gametes carry mostly mutant haplotypes, \( B_1 \times B_2 \) matings are less costly because they predominantly involve mutant haplotypes. We find the same patterns for non-neutral mutations when we generalize our model to three mitochondrial types (Table A.15).

2.4.6 Recombination between mating type and inheritance loci

Previously, \( U \times U \) matings were not possible because we assumed tight linkage between mating type and inheritance loci. But if we allow recombination to occur between these loci, \( U_1 \times U_2 \) matings become possible. In this scenario, the number of gametes increases to four (\( B_1, B_2, U_1, \) and \( U_2 \)), as does the number of genotypes (\( B_1B_2, U_1B_2, U_1U_2, \) and \( U_2B_1 \)). There are three main ways in which mitochondrial inheritance could be regulated in \( U_1 \times U_2 \) matings. (1) One \( U \) allele is dominant to the other, leading to uniparental inheritance; (2) each \( U \) allele ensures inheritance of its mitochondria, resulting in biparental inheritance; or (3) inheritance is more or less random so that some matings result in uniparental inheritance and some in biparental inheritance. We model all three cases.

When \( U_1 \times U_2 \) matings lead to uniparental inheritance, the \( U_1U_2 \) genotype always spreads until it is fixed in the population, leading to complete uniparental inheritance (Figure 2.4A and Tables A.16–A.18). When \( U_1 \times U_2 \) matings lead to biparental inheritance, however, uniparental inheritance does not become fixed and the population reaches a polymorphic equilibrium (Figure 2.4B-C). Under these conditions, the frequency of uniparental inheritance at equilibrium is 0.5 (Table A.19–Table A.21). Uniparental inheritance cannot exceed 0.5 because increasing the frequency of \( U_1 \) or \( U_2 \) simply increases the proportion of biparental \( U_1 \times U_2 \) matings. The frequency of uniparental inheritance remains very low when we assume a concave fitness function (Figure 2.4B), but reaches its maximum (0.5) when we assume a linear or convex fitness function (Figure 2.4C) (see Figure A.12 and Figure A.13 for an explanation).
Figure 2.4: Recombination and no mating types scenarios. Parameters: $n = 20$, $\mu = 10^{-4}$, $c_h = 0.2$. A. As the $U$ allele initially spreads (generations 0–1700), the $U_1B_2/U_2B_1$ genotypes increase in frequency. But, because $U_1B_2$ and $U_2B_1$ cells lead to $B_1B_2$ cells through meiosis and random mating, the $U_1U_2$ genotype soon takes over and uniparental inheritance becomes fixed. Additional parameters: $P_r = 0.5$ and concave fitness. B. Biparental inheritance dominates when $U \times U$ matings are biparental and fitness is concave. C. Uniparental inheritance invades to its maximum value (0.5) when $U \times U$ matings are biparental and fitness is linear or convex. (The frequency of uniparental inheritance is the sum of $U_1U_2$ and $U_2B_1$.) Additional parameters: linear fitness. D. $U \times U$ matings have a mixture of uniparental and biparental inheritance. Unlike in B, $U_1U_2$ no longer becomes fixed because some $U \times U$ matings now have biparental inheritance and further increasing $U_1U_2$ would only increase the overall level of biparental inheritance. Additional parameters: $P_b = 0.1$ and linear fitness. E. Lines represent the frequency of uniparental inheritance in separate simulations with linear fitness and varying probabilities of biparental inheritance ($P_b$) when $U \times U$ matings have a mixture of uniparental and biparental inheritance. As $P_b$ increases, $U \times U$ matings are more likely to lead to biparental inheritance, which decreases the frequency of uniparental inheritance at equilibrium. F. No mating types scenario under concave fitness. F is identical to A except that the frequency of $UB$ in F is the sum of the $U_1B_2$ and $U_2B_1$ frequencies in A.
When the probability of recombination ($P_r$) is sufficiently high ($10^{-4} \leq P_r \leq 0.5$ in Figure A.11), the $U_1B_2$ and $U_2B_1$ genotypes have the same frequency at equilibrium (Figure A.11B-D). Now uniparental inheritance is no longer associated with a single mating type but is evenly split between the two mating types (Table A.19–Table A.21). When $P_r$ is sufficiently small ($P_r = 10^{-5}$ in Figure A.11), the recombination rate is so low that the mating type and inheritance loci are essentially linked and the $U_1B_2$ genotype becomes fixed (as in the general model) (Figure A.11A).

When we assume a mixture of uniparental inheritance and biparental inheritance, we let $U_1 \times U_2$ matings lead to biparental inheritance with probability $P_b$ and to uniparental inheritance with probability $1 - P_b$. Lowering $P_b$ increases the frequency of uniparental inheritance, and uniparental inheritance becomes fixed when $P_b = 0$ (Figure 2.4A,E). Under linear and convex fitness functions, the equilibrium always maximizes the level of uniparental inheritance (Table A.22 and Table A.23). Under concave fitness, however, uniparental inheritance is only maximized for particular values of $P_b$ (roughly $P_b \leq 0.2$ for the parameter values we considered) (Table A.22; rows 2–3). (See section A.10 for how we determine when uniparental inheritance is maximized.)

We also find that uniparental inheritance can evolve in the complete absence of mating types. The no mating types scenario differs from the recombination case in that $UB$ equals the sum of $U_1B_2$ and $U_2B_1$ at equilibrium (Figure 2.4A,F) (see section A.4 for more details).

**2.4.7 Can selection against heteroplasm explain the exceptions to uniparental inheritance?**

In this section, we explore whether relaxing some of the assumptions in our general model can lead to mitochondrial inheritance patterns that resemble some of the known exceptions to uniparental inheritance. Exceptions to uniparental inheritance fall in three main categories: organisms that (1) regularly inherit mitochondria from both parents; (2) normally inherit mitochondria from one of the two parents but on occasion inherit mitochondria from both; and (3) inherit mitochondria from either or both parents.

Baker’s yeast, *Saccharomyces cerevisiae*, regularly inherits mitochondria from both parents (though uniparental inheritance also occurs), but heteroplasm is transient because the diploid cell has only a few mitochondria [16] and divides repeatedly, which separates heteroplasmic cells into cells homoplastic for either mitochondrial type (vegetative segregation) [9, 10]. Vegetative segregation is usually completed within.
twenty generations, but up to 50% of zygotes may be homoplasmic after the first division ([10] and references therein). Thus, *Saccharomyces* may restore homoplasmy as quickly as organisms that actively destroy one mitochondrial lineage [17]. Similarly, the geranium *Pelargonium zonale* often inherits cytoplasmic organelles from both parents (chloroplasts in this case). As with *Saccharomyces*, heteroplasmy is transient in *Pelargonium* because of rapid vegetative segregation of heteroplasmic cells shortly after syngamy [9]. We added mitotic divisions to our model to test whether vegetative segregation could maintain biparental inheritance under selection against heteroplasmy. When we include mitosis before selection (which assumes that vegetative segregation occurs swiftly, before selection has time to act), uniparental inheritance does not spread, provided that the number of mitochondria is low \((n = 4)\) and the number of divisions is high (Table A.24; rows 7–8). Under these conditions, biparental inheritance is stable because heteroplasmic cells resulting from biparental inheritance segregate into homoplasmic cells before selection acts. If there are insufficient mitotic divisions, or if selection acts before vegetative segregation is complete, then uniparental inheritance replaces biparental inheritance, although it spreads much more slowly than when there are no mitotic divisions (Table A.24 (rows 3–6) and Table A.25). When there are more mitochondria per cell (e.g. \(n = 8\)), biparental inheritance is only stable if the number of cell divisions increases to compensate (Table A.24; rows 9–10). Thus, biparental inheritance can be stable under selection against heteroplasmy but only under a narrow set of conditions, explaining why this form of inheritance is rare.

In other isogamous organisms, including the acellular slime molds *Physarum polycephalum* and *Didymium iridis* and the algae *Chlamydomonas reinhardtii*, mitochondria from both gametes mix before one mitochondrial lineage is destroyed post-fertilization, often by nucleases [18–20]. This mechanism is not perfect and these organisms sometimes deviate from strict uniparental inheritance [9, 18–20]. While uniparental inheritance is the norm in the slime mold *P. polycephalum*, sometimes both mitochondrial lineages survive, leading to varying degrees of biparental inheritance [18]. Could uniparental inheritance still spread under such conditions? Since mating types and inheritance loci are tightly linked in *Physarum* [18], we explore this question using our general model that assumes linkage. Now, \(U_1 \times B_2\) matings lead to biparental inheritance with probability \(P_b\) and to uniparental inheritance with probability \(1 - P_b\). For the parameter values that we examined, the \(U_1B_2\) genotype always goes to fixation when \(P_b < 1\) and the fitness function is linear or convex (Table A.26). (When fitness is concave, \(P_b\) must be roughly < 0.05 for the genotype
to become fixed.) Under these conditions, the frequency of biparental inheritance at equilibrium is equal to $P_b$ (Table A.26). In this scenario, the level of biparental inheritance in the population simply reflects the likelihood that an individual mating results in biparental inheritance.

*Chlamydomonas reinhardtii* and *Didymium iridis* can inherit mitochondria from either or both parents [19, 20]. *Chlamydomonas* normally inherits mitochondria from the *mt−* parent and chloroplasts from the *mt+* parent, but under some circumstances it can inherit mitochondria from *mt+* and chloroplasts from *mt−* or mitochondria and chloroplasts from both [20]. *Didymium iridis* has random, biased, or dominant patterns of uniparental inheritance. Under random uniparental inheritance, either parental strain is equally likely to be the mitochondrial donor while, under biased inheritance, one strain is more likely to be the mitochondrial donor [19]. Under dominant inheritance, one strain is always the donor. *Didymium* also has low levels of biparental inheritance [19]. In this scenario, we test whether selection against heteroplasmy could lead to the evolution of a system with a mixture of uniparental inheritance (from either parent) and biparental inheritance. We assume that mating types can recombine and that $U_1 \times U_2$ matings can lead to mitochondria being inherited from $U_1$, $U_2$, or both. Mitochondria are inherited from $U_1$ with probability $P_{U_1}$, from $U_2$ with probability $P_{U_2}$ and from both parents with probability $P_b$ (where $P_{U_1} + P_{U_2} + P_b = 1$). Now, uniparental inheritance comes from $U_1 \times B_2$ matings, $U_2 \times B_1$ matings and those $U_1 \times U_2$ matings with uniparental inheritance. Irrespective of the values of $P_{U_1}$ and $P_{U_2}$, we find the same results as with our earlier model in which $U_1 \times U_2$ matings led to a mixture of uniparental and biparental inheritance (Table A.22 and Table A.23). This is because equilibrium depends only on the value of $P_b$. (Since uniparental inheritance quickly eliminates most heteroplasmic cells, $U_1U_2$ cells are almost entirely homoplasmic regardless of which gamete donates mitochondria.) Consequently, different probabilities of inheriting mitochondria biparentally ($P_b$), from mating type 1 ($P_{U_1}$), or from mating type 2 ($P_{U_2}$) lead to a range of inheritance patterns that include uniparental inheritance (from both parents) and biparental inheritance (see Table A.27 for some examples).

Lastly, selection against heteroplasmy provides an explanation for the cases in which mitochondria are inherited from one parent while chloroplasts are inherited from the other (e.g. in *Chlamydomonas* and pines [20, 21]). If uniparental inheritance simply evolved to maintain homoplasm in cells, it should not matter which parent donates mitochondria or chloroplasts.
2.5 Discussion

Our model shows that selection against heteroplasmy can lead to the fixation of uniparental inheritance in an ancestrally biparental population. We find that uniparental inheritance replaces biparental inheritance under almost all tested scenarios and parameter values. Our model also explains many of the known exceptions to strict uniparental inheritance. We show that uniparental inheritance can replace biparental inheritance whether mutations lead to neutral or non-neutral haplotypes. Relaxing our initial assumptions of pre-existing mating types and lack of recombination does not prevent uniparental inheritance from evolving. As we make no attempt to resolve the evolution of mating types within the context of mitochondrial inheritance, as others have previously attempted [1, 22], our findings thus leave open the possibility that mating types preceded uniparental inheritance, evolved as a consequence of uniparental inheritance, or evolved after uniparental inheritance.

In contrast to previous models, we show that uniparental inheritance can spread under realistic mutation rates and number of mitochondria per cell. The lowest value of \( \mu \) that we tested \((10^{-10})\) is eight orders of magnitude lower than required by the genomic conflict theory [1] and compares favorably with empirical mutation rates \((10^{-7} \text{ to } 10^{-8} \text{ per site per generation [23–25]})\). Both the genomic conflict and mutation clearance hypotheses require unrealistic mutation rates and number of mitochondria per cell for uniparental inheritance to replace biparental inheritance, while uniparental inheritance cannot replace biparental inheritance under any parameter values in the mitochondrial-nuclear coadaptation model [1]. The genomic conflict model requires a mutation rate of 1% per generation before uniparental inheritance can replace biparental inheritance [1]. The only known example that satisfies this assumption is the petite mutant in *Saccharomyces cerevisiae*, which is a hyper-mutable selfish mitochondrion that can spontaneously arise at a rate of 1% per generation [26]. Under this mutation rate, however, the genomic conflict model requires that cells contain at least 50 mitochondria [1], whereas most extant isogamous species, including *Saccharomyces*, contain fewer than 20 mitochondria at syngamy [16, 18]. As mutant mitochondria lack a transmission advantage over wild type mitochondria in the mutation clearance hypothesis, the mutation clearance model requires even higher mutation rates [1]. To the best of our knowledge, no extant organism satisfies the assumptions of the genomic conflict or mutation clearance hypotheses.

Why do our results differ from the findings of previous models? In the genomic conflict and mutation clearance models, wild type mitochondria mutate to selfish or
deleterious mitochondria. Biparental inheritance results in cells that are heteroplasmic for wild type and mutant mtDNA, while $U_1$ gametes mostly contain wild type mitochondria [1]. Because $U_1$ purges $B_2$ gametes of mutant mitochondria, $B_1 \times B_2$ matings involve increasingly fewer mutant mitochondria as the frequency of $U_1$ increases [1, 5]. $U_1$ is thus subject to negative frequency-dependent selection, and the population reaches equilibrium well before uniparental inheritance replaces biparental inheritance at realistic mutation rates [1]. The mitochondrial-nuclear coadaptation model assumes that mitochondria are well matched or poorly matched to nuclear alleles [1, 8]. Because mutation can lead to matched nuclear-mitochondrial states becoming unmatched, the effective mitochondrial mutation rate is lower in the mitochondrial-nuclear coadaptation model, which prevents uniparental inheritance from displacing biparental inheritance under any parameter values [1].

Evidence for a cost of heteroplasmy comes from a recent study that compared the effect of two mtDNA haplotypes (NZB and 129S6) in a cogenic nuclear background on the functioning of mice [12]. Mice homoplasmic for NZB or 129S6 were phenotypically normal, but NZB-129S6 heteroplasmic mice suffered from reduced activity, lowered food intake, compromised respiration, heightened stress response, and impaired cognition [12]. While the mechanism(s) behind the cost of heteroplasmy is unknown, there are a few possibilities. Heteroplasmy may disrupt cell signaling by altering production of reactive oxygen species (ROS) [27] and there are indications that heteroplasmy can increase mitochondrial ROS levels [28, 29], leading to phenotypes that differ from cells that are homoplasmic for either haplotype [29, 30]. Alternatively heteroplasmy may lead to deleterious interactions between polypeptides from different mitochondria within the same electron transport chain [12, 31]. Because chloroplasts also contain independent genomes, are involved in cellular bioenergetics, and generally show uniparental inheritance [9], our findings likely apply to both mitochondria and chloroplasts.

Although the evidence in mice is compelling [12], it is unknown whether selection against heteroplasmy is a general phenomenon in eukaryotes. While Sharpley and colleagues [12] used different mitochondrial lineages to construct heteroplasmic individuals, our model assumes that mutations accumulated within a single generation can cause mitochondrial types to become sufficiently distinct to lead to negative effects for the cell. At this stage we do not know how different mitochondrial genomes have to be for selection against heteroplasmy to apply. It could also be that there are regions of the genome in which heteroplasmic mutations have a stronger effect on fitness than others. To support or refute our model, we now need solid empirical
data on a range of organisms showing the cost, if any, of heteroplasmy on organism fitness.

While we have referred to $n$ as the number of mitochondria in the cell, $n$ actually refers to the number of segregating units of mtDNA at syngamy. Mitochondria pack DNA into DNA-protein complexes called nucleoids, which themselves may contain multiple copies of mtDNA [32, 33]. It is currently unknown whether the segregating unit is the mtDNA molecule itself, the nucleoid, the mitochondrion or another level of mtDNA organization [33]. But as nucleoids are predominantly homoplasmic, even in heteroplasmic tissues [32, 33], the number of mitochondria may be a reasonable approximation of the number of segregating units in the cell. If the segregating unit is at a lower level of organization (e.g. the mtDNA molecule), then $n$, as used in our model, will apply to the number of segregating units not the number of mitochondria per cell (e.g. $n = 200$ would then apply to a cell with 200 segregating units, which may be a cell with far fewer than 200 mitochondria).

By assuming an infinite population size, a common assumption in studies of this kind [1, 5, 6, 8] we have ignored genetic drift, which can be a powerful force in population genetics. While it is beyond the scope of this study to formally model the effects of genetic drift on the evolution of uniparental inheritance, we can anticipate some of its effects. As the mutation leading to uniparental inheritance has a small advantage when its frequency is low, genetic drift will lead to the frequent loss of those mutations. Thus, the initial invasion of a mutation for uniparental inheritance may be largely determined by genetic drift rather than by positive selection. As the frequency of uniparental inheritance increases, however, so too does its advantage, reducing the probability that the mutation is lost to drift. The potential for rare mutations to be lost to drift is not unique to our model. The genomic conflict hypothesis requires stringent conditions for uniparental inheritance mutations to invade [6, 34]. Under this hypothesis, a mutation for uniparental inheritance must arise within a population that contains selfish mutants but in which the selfish mutant is not fixed. Otherwise, uniparental inheritance cannot become associated with non-selfish mitochondria. Any mutations leading to uniparental inheritance that arise outside of this window will have no selective advantage and will be more likely to be lost by genetic drift [6, 34].

## 2.6 Conclusion

Selection against heteroplasmy has implications for the evolution of the mitochondrial genome. Because of a smaller effective population size, which is more strongly
affected by genetic drift, and higher mutation rates, mtDNA should be less conserved than the nuclear genome [35, 36]. Indeed, mitochondrial transfer RNAs and synonymous sites mutate 5–50 times more frequently than comparable elements in the nuclear genome [35, 37]. Because the mitochondrial genome is effectively asexual, any deleterious mutations in the fittest haplotype cannot be rescued (except by unlikely back mutations). This effect, known as Muller’s ratchet, should eventually lead to irreparable genome meltdown [38, 39]. In stark contrast to theoretical predictions, however, mitochondrial coding genes are more conserved than analogous nuclear oxidative phosphorylation genes [36]. When mtDNA mutates, only one of the many mtDNA molecules in the cell is affected, leading to a heteroplasmic cell. Selection against heteroplasmy should reduce the probability that mtDNA mutations spread throughout the cell, which, in turn, should oppose changes to mtDNA. Thus, selection against heteroplasmy may not only explain the evolution of uniparental inheritance but also why mitochondrial coding genes have thus far managed to resist the effects of Muller’s ratchet.
Bibliography


[1.1.3, 2.2, 2.3, 2.5, 3.2, 3.3, 4.2, 4.3, 4.4.3, 4.5, 5.5, 6.1, 6.1.1, 6.1.3, †, 6.3, ‖, A.5, A.5.1, C.2]


of the Royal Society B-Biological Sciences 279(1734):1865–1872. [1.1.3, 2.2, 2.3, 2.5, 3.2, 4.2, 4.5, 6.1, A.5]


[38] Muller HJ (1964) The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 1(1):2–9. [1.2.2, 2.6, 3.2, 3.4.6, 4.2]

Chapter 3

Uniparental inheritance promotes adaptive evolution in cytoplasmic genomes

3.1 Abstract

Eukaryotes carry numerous asexual cytoplasmic genomes (mitochondria and chloroplasts). Lacking recombination, asexual genomes suffer from impaired adaptive evolution. Yet, empirical evidence indicates that cytoplasmic genomes experience higher levels of adaptive evolution than predicted by theory. In this study, we use a computational model to show that the unique biology of cytoplasmic genomes—specifically their organization into host cells and their uniparental inheritance—enable them to undergo adaptive evolution more effectively than comparable free-living asexual genomes. Uniparental inheritance of cytoplasmic genomes decreases competition between different beneficial substitutions (clonal interference), reduces genetic hitchhiking of deleterious substitutions during selective sweeps, and promotes adaptive evolution by increasing the level of beneficial substitutions relative to deleterious substitutions. When cytoplasmic genome inheritance is biparental, the presumed ancestral state, decreasing the number of genomes transmitted during gametogenesis aids adaptive evolution. Nevertheless, adaptive evolution is always more efficient when inheritance is uniparental. Our findings explain empirical observations that cytoplasmic genomes—despite their asexual mode of reproduction—can readily undergo adaptive evolution.
3.2 Introduction

About 1.5–2 billion years ago, an α-proteobacterium was engulfed by a proto-eukaryote, an event that led to modern mitochondria [1]. Likewise, chloroplasts in plants and algae are derived from a cyanobacterium [2]. These cytoplasmic genomes are essential to extant eukaryotic life, producing much of the energy required by their eukaryotic hosts. Like their ancient ancestors, cytoplasmic genomes reproduce asexually and appear to undergo little recombination with other cytoplasmic genomes [3, 4].

Since they lack recombination, asexual genomes have lower rates of adaptive evolution than sexual genomes unless the size of the population is extremely large [5, 6]. While the theoretical costs of asexual reproduction have long been known [5–9], conclusive empirical evidence is more recent [10–13]. Three factors largely explain why asexual genomes have low rates of adaptive evolution: (1) beneficial substitutions accumulate slowly; (2) deleterious substitutions are poorly selected against; and (3) when beneficial substitutions do spread, any linked deleterious substitutions also increase in frequency through genetic hitchhiking [5, 7, 8, 10, 11].

The lack of recombination in asexual genomes slows the accumulation of beneficial substitutions. Recombination can aid the spread of beneficial substitutions by separating out rare beneficial mutations from deleterious genetic backgrounds (“ruby in the rubbish”) [14]. Furthermore, recombination can reduce competition between different beneficial substitutions (“clonal interference”) [5, 7, 8, 10, 11, 15–17]. Under realistic population sizes and mutation rates, an asexual population will contain multiple genomes—each with different beneficial substitutions—competing with one another for fixation [11, 16]. Ultimately, clonal interference leads to the loss of some beneficial substitutions, reducing the efficiency of adaptive evolution [5, 7, 8, 10, 11, 15–17].

The lack of recombination also makes it more difficult for asexual genomes to purge deleterious substitutions. An asexual genome can only restore a loss of function from a deleterious substitution through a back mutation or a compensatory mutation, both of which are rare [5, 18]. Unless the size of the population is very large, the number of slightly deleterious substitutions should increase over time as the least-mutated class of genome is lost through genetic drift (“Muller’s ratchet”) [5, 18].

If that were not enough, asexual genomes are also especially susceptible to genetic hitchhiking [10, 11], a process by which deleterious substitutions spread through their association with beneficial substitutions [19, 20]. As all loci on an asexual genome are linked, deleterious and beneficial substitutions on the same genome will segregate...
3.2. Introduction

together. When the positive effect of a beneficial substitution outweighs the negative effect of a deleterious substitution, the genome that carries both can spread through positive selection [19, 20]. Even when the additive effect is zero or negative, a beneficial substitution can still aid the spread of a deleterious substitution via genetic drift by reducing the efficiency of selection against the deleterious substitution. Genetic hitchhiking can thus offset the benefits of accumulating beneficial substitutions by interfering with the genome’s ability to purge deleterious substitutions [19, 20].

Free-living asexual organisms generally have very large population sizes [21] and may undergo occasional sexual exchange (e.g. conjugation in bacteria [22]), allowing these organisms to alleviate some of the costs of asexual reproduction [5, 6]. Asexual cytoplasmic genomes, however, have an effective population size much smaller than that of free-living asexual organisms [21, 23]. As a smaller population size increases the effect of genetic drift, cytoplasmic genomes should have less efficient selection than asexual organisms [24, 25] and should struggle to accumulate beneficial substitutions and to purge deleterious substitutions [26–28].

But despite theoretical predictions, cytoplasmic genomes readily undergo adaptive evolution. Mitochondrial protein-coding genes show signatures that are consistent with both low levels of deleterious substitutions [21, 29, 30] and frequent selective sweeps of beneficial substitutions [31, 32]. Indeed, it is estimated that 26% of mitochondrial substitutions that alter proteins in animals have become fixed through adaptive evolution [33]. Beneficial substitutions in the mitochondrial genome have helped animals adapt to specialized metabolic requirements [34–37] and have enabled humans to adapt to cold northern climates [38]. Likewise, it is clear that adaptive evolution has played a role in the evolution of chloroplast genomes [39, 40].

How then do we reconcile empirical evidence for adaptive evolution in cytoplasmic genomes with theoretical predictions that such adaptation should be impaired? Unlike free-living asexual organisms, which are directly exposed to selection, cytoplasmic genomes exist within host cells. The fitness of cytoplasmic genomes is therefore closely aligned with the fitness of their host. Each of these hosts carries multiple cytoplasmic genomes that are generally inherited from a single parent (uniparental inheritance) [41]. During gametogenesis, cytoplasmic genomes can undergo tight population bottlenecks, affecting the transmission of genomes from parent to offspring [42, 43]. Cytoplasmic genomes are thus subject to very different evolutionary pressures than free-living asexual organisms.

Some of the effects of uniparental inheritance and a transmission bottleneck on the evolution of cytoplasmic genomes have already been identified. Both uniparental
inheritance and a transmission bottleneck decrease within-cell variance in cytoplasmic genomes and increase between-cell variance. [41, 44–46]. Uniparental inheritance is known to select against deleterious mutations [45–48] and select for mito-nuclear coadaptation [49]. Similarly, a transmission bottleneck and other forms of within-generation drift are known to slow the accumulation of deleterious substitutions in cytoplasmic genomes [27, 44, 50].

Although the effect of uniparental inheritance and a bottleneck on the accumulation of deleterious substitutions is reasonably well-studied, much less attention has been paid to the other limitations of asexual reproduction: slow accumulation of beneficial substitutions and high levels of genetic hitchhiking. The two studies that have addressed the spread of beneficial substitutions have come to contradictory conclusions. Takahata and Slatkin [50] showed that within-generation drift promoted the accumulation of beneficial substitutions. In contrast, Roze and colleagues [45] found that within-generation drift due to a bottleneck reduced the fixation probability of a beneficial mutation. Takahata and Slatkin found no difference between uniparental and biparental inheritance of cytoplasmic genomes [50] while Roze and colleagues found that uniparental inheritance increased the fixation probability of a beneficial mutation and its frequency at mutation-selection equilibrium [45]. Of the two previous studies, only the model of Takahata and Slatkin was able to examine the accumulation of substitutions [50] (the model of Roze and colleagues only considered a single locus [45]). To our knowledge, no study has looked at how inheritance mode affects genetic hitchhiking in cytoplasmic genomes.

Here we develop theory that explains how cytoplasmic genomes are capable of adaptive evolution despite their lack of recombination. We will show how the biology of cytoplasmic genomes—specifically their organization into host cells and their uniparental inheritance—allows them to accumulate beneficial substitutions and to purge deleterious substitutions more efficiently than comparable free-living asexual genomes.

### 3.3 Model

For simplicity, we base our model on a population of diploid single-celled eukaryotes. We examine the accumulation of beneficial and deleterious substitutions in an individual-based computational model that compares uniparental inheritance of cytoplasmic genomes with biparental inheritance (the presumed ancestral state [41]). We examine each form of inheritance separately. Since genetic drift plays an important
role in the spread of substitutions, we take stochastic effects into account. We vary the size of the transmission bottleneck during meiosis (i.e. the number of cytoplasmic genomes passed from parent to gamete) to alter the level of genetic drift. To examine how the organization of cytoplasmic genomes into host cells affects their evolution, we also include a model of comparable free-living asexual genomes.

We have four specific aims. We will determine how inheritance mode and the size of the transmission bottleneck affect (Aim 1) clonal interference and the accumulation of beneficial substitutions; (Aim 2) the accumulation of deleterious substitutions; (Aim 3) the level of genetic hitchhiking; and (Aim 4) the level of adaptive evolution, which we define as the ratio of beneficial to deleterious substitutions. Although uniparental inheritance and a transmission bottleneck are known to select against deleterious mutations on their own [27, 44–48, 50], the interaction between inheritance mode, transmission bottleneck, and the accumulation of deleterious substitutions has not to our knowledge been examined. Thus we include Aim 2 to specifically examine interactions between inheritance mode and size of the transmission bottleneck. To address our aims, we built four variations of our model. First, we examine clonal interference and the accumulation of beneficial substitutions using a model that considers beneficial but not deleterious mutations (Aim 1). Second, we consider deleterious but not beneficial mutations to determine how inheritance mode and a transmission bottleneck affect the accumulation of deleterious substitutions in cytoplasmic genomes (Aim 2). Third, we combine both beneficial and deleterious substitutions. This allows us to examine the accumulation of deleterious substitutions in the presence of beneficial mutations (genetic hitchhiking; Aim 3) and the ratio of beneficial to deleterious substitutions (Aim 4). For all aims, we compare our models of cytoplasmic genomes to a comparable population of free-living asexual genomes. This serves as a null model, allowing us to examine the strength of selection when asexual genomes are directly exposed to selection.

The population contains \( N \) individuals, each carrying the nuclear genotype \( Aa \), where \( A \) and \( a \) are self-incompatible mating type alleles. Diploid cells contain \( n \) cytoplasmic genomes, and each genome has \( l = 20000 \) linked base pairs. A cytoplasmic genome is identified by the number of beneficial and deleterious substitutions it carries (\( \alpha \) and \( \kappa \) respectively; note, we do not track where on the genome the mutations occur). Cells are identified by the number of each type of cytoplasmic genome they carry. The life cycle has four stages, and a complete passage through the four stages comprises a generation. The first stage is mutation. Initially, all cells carry cytoplasmic genomes with zero substitutions. Mutations can occur at any of the \( l \) base pairs.
The probability that one of these \( l \) sites will mutate to a beneficial or deleterious site is given by \( \mu_b \) and \( \mu_d \) per site per generation respectively (determined via generation of random numbers within each simulation). As the mutation rate in mitochondrial DNA is between \( 7.8 \times 10^{-8} \) and \( 1.7 \times 10^{-7} \) per nucleotide per generation \([51–53]\), we let \( \mu_d = 1 \times 10^{-7} \) per nucleotide per generation. We assume the beneficial mutation rate is lower than the deleterious mutation rate, and as such, examine both \( \mu_b = 1 \times 10^{-8} \) and \( \mu_b = 1 \times 10^{-9} \) per nucleotide per generation \([54]\).

After mutation, cells are subject to selection, assumed for simplicity to act only on diploid cells. We assume that each substitution has the same effect, which is given by the selection coefficient (\( s_b \) for beneficial and \( s_d \) for deleterious) and that fitness is additive. We assume that a cell’s fitness depends on the total number of substitutions carried by its cytoplasmic genomes. As there are few data on the distribution of fitness effects of beneficial substitutions in cytoplasmic genomes, we examine three fitness functions: concave up, linear, and concave down (Figure 3.1A). For deleterious substitutions in cytoplasmic genomes, there is strong evidence that fitness is only strongly affected when the cell carries a high proportion of deleterious genomes \([55, 56]\), and so we use a decreasing concave down function to model deleterious substitutions (Figure 3.1B). When we combine beneficial and deleterious mutations in a single model, we examine all three fitness functions for the accumulation of beneficial substitutions but only a concave down decreasing fitness function for the accumulation of deleterious substitutions (Figure 3.1B).

We focus on selection coefficients that represent mutations with small effects on fitness: \( s_b = 0.01 - 0.1 \) (see the legend of Figure 3.1 for a description of how the selection coefficient relates to fitness). Cells are assigned a relative fitness based on the number of beneficial and deleterious substitutions carried by their cytoplasmic genomes. These fitness values are used to sample \( N \) new individuals for the next generation.

Each of the post-selection diploid cells then undergoes meiosis to produce two gametes, one with nuclear allele \( A \) and the other with nuclear allele \( a \). Each gamete also carries \( b \) cytoplasmic genomes sampled with replacement from the \( n \) cytoplasmic genomes carried by the parent cell (with \( b \leq n/2 \)) \([42]\). We examine both a tight transmission bottleneck (\( b = n/10 \)) and a relaxed transmission bottleneck (\( b = n/2 \)). To maintain population size at \( N \), each diploid cell produces two gametes.

During mating, each gamete produced during meiosis is randomly paired with another gamete of a compatible mating type. These paired cells fuse to produce diploid cells. Under biparental inheritance, both the gametes with the \( A \) and \( a \) alleles
3.3. Model

Figure 3.1: Fitness functions. Additional parameters: \( n = 50, s_b = s_d = 0.1, \gamma = 5 \).

**A.** The three fitness functions used in this study in the case of beneficial mutations only. The selection coefficient is defined such that \( 1 - s_b \) represents the fitness of a cell with zero beneficial substitutions (a cell with \( n\gamma \) beneficial substitutions has a fitness of 1, where \( n \) is the number of cytoplasmic genomes and \( \gamma \) is the number of substitutions each cytoplasmic genome must accumulate before the simulation is terminated). In this case, where \( n = 50 \) and \( \gamma = 5 \), a cell’s fitness is 1 when each cytoplasmic genome in the cell carries an average of 5 substitutions (50 \( \times \) 5 = 250 beneficial substitutions in total). **B.** The deleterious fitness function. Here, a cell with no deleterious substitutions has a fitness of 1, while a cell with \( n\gamma \) substitutions has a fitness of \( 1 - s_d \). We only examine a concave down decreasing function for the accumulation of deleterious substitutions (unless we are comparing cytoplasmic genomes to free-living genomes, in which case we use a linear fitness function). **C.** One of the fitness functions used in the model with both beneficial and deleterious mutations. The beneficial substitution portion of the function can take any of the forms in panel A while the deleterious substitution portion takes the form in panel B. In this example the fitness surface combines a linear function for beneficial substitutions with a concave down fitness function for deleterious substitutions. The color represents the fitness of a cell carrying a given number of deleterious substitutions (x-axis) and beneficial substitutions (y-axis). Equations for the fitness functions can be found in section B.3.2 (A), section B.4 (B), and section B.5.2. (C)

pass on their \( b \) cytoplasmic genomes, while under uniparental inheritance, only the \( b \) genomes from the gamete with the \( A \) allele are transmitted. Finally, \( n \) genomes are restored to each new diploid cell by sampling \( n \) genomes with replacement from the genomes carried by the diploid cell after mating (2\( b \) under biparental inheritance and \( b \) under uniparental inheritance). The model then repeats, following the cycle of mutation, selection, meiosis, and mating described above.

To ensure that our model of free-living asexual genomes can be directly compared to our model of cytoplasmic genomes, we assume a population size of \( N \times n \) free-living genomes. Each free-living genome carries one haploid asexual nuclear genome with \( l \) base pairs. Now there are only two stages to the life cycle: mutation and selection. Mutation proceeds as in the model of cytoplasmic genomes. Selection, however, now depends only on the number of substitutions carried by a genome. We assume that a mutation has the same effect on the fitness of a free-living cell as a mutation on a
cytoplasmic genome has on the fitness of its host cell. (When comparing free-living and cytoplasmic genomes, we always use a linear fitness function for both beneficial and deleterious substitutions because for this function the strength of selection on a new substitution is independent of existing substitution load.) Our intention is not to accurately model extant populations of free-living asexual organisms, as these differ in a number of ways from cytoplasmic genomes (e.g. population size, mutation rate, and genome size [21]), but rather to examine how the organization of multiple cytoplasmic genomes within a host affects their evolution.

When we consider beneficial mutations only (Aim 1), the simulation stops once every cytoplasmic genome in the population has accumulated at least $\gamma$ beneficial substitutions. For the remaining models, we run each simulation for 10,000 generations. For all the models, we average the results of 500 Monte Carlo simulations for each combination of parameter values (we vary $N$, $n$, $b$, $s_b$, $s_d$, and the fitness functions associated with beneficial substitutions). We wrote our model in R version 3.1.2 [57]. For a detailed description of the models, see section B.3—section B.5.

3.4 Results

3.4.1 Cytoplasmic genomes accumulate beneficial mutations faster than free-living genomes

The units of selection differ between cytoplasmic genomes (eukaryotic host cell) and free-living genomes (free-living asexual cell). Cytoplasmic genomes have two levels at which variance in fitness can be generated: variation in the number of substitutions per genome and variation in the relative number of each genome type in a host cell (Figure 3.2A). In contrast, free-living genomes can differ only in the number of substitutions carried per genome. Consequently, cytoplasmic genomes have a greater potential for creating variance between the units of selection than free-living genomes (Figure 3.2B).

For conceptual purposes, we break down the accumulation of beneficial substitutions into two phases. In the first phase (establishment), we determine the time for a genome that carries $\alpha$ substitutions to become established in a population that contains genomes with $\alpha - 1$ or fewer beneficial substitutions. Since we examine small selection coefficients, drift dominates the fate of genomes when they are rare, and the genome with $\alpha$ substitutions is frequently lost to drift when it first arises. The establishment phase starts when we first observe a genome with $\alpha$ substitutions and ends when that genome persists in the population (i.e. it is no longer lost to drift).
3.4. Results

**Figure 3.2: Dynamics in the accumulation of beneficial substitutions.** Parameters: $N = 1000$, $n = 50$, $\mu_b = 10^{-8}$, linear fitness function, and $b = 25$ (relaxed transmission bottleneck) or $b = 5$ (tight transmission bottleneck). As neither fitness function nor selection coefficient qualitatively affect the results, we show a single representative set of parameter values. Error bars represent standard error of the mean. UPI is uniparental inheritance with a relaxed bottleneck, UPI (bot) is uniparental inheritance with a tight bottleneck, BPI is biparental inheritance with a relaxed bottleneck, and BPI (bot) is biparental inheritance with a tight bottleneck. 

**A.** Variance in the number of different cytoplasmic genomes carried by cells (averaged over all cells in the population each generation). As free-living cells carry a single genome, they have no within-cell variance.

**B.** Variance of all cells' fitness values (averaged over each generation). (Note that between-cell variation in the free-living population is depicted but is so low that it appears as zero.)

**C.** The number of generations separating the genome carrying $\alpha$ substitutions from the genome carrying $\alpha + 1$ (averaged over all observed substitutions, but excluding $\alpha = 1$, as the dynamics of $\alpha = 1$ are largely driven by the starting conditions). The establishment phase begins when the genome carrying $\alpha$ substitutions first appears and ends when that genome becomes established in the population (depicted in dark blue). The sweep phase begins with the establishment of the genome with $\alpha$ substitutions and ends upon the first appearance of the genome with $\alpha + 1$ substitutions (depicted in yellow).

**D.** During the establishment period of the genome with $\alpha$ substitutions, shows the probability of losing all genomes with $\alpha$ substitutions ($P(\text{lose } \alpha)$) and the probability of regenerating at least one genome with $\alpha$ substitutions once all genomes with $\alpha$ substitutions have been lost ($P(\text{regain } \alpha)$) (averaged over all observed establishment periods, but excluding $\alpha = 1$).

**E.** During the establishment period of the genome with $\alpha$ substitutions, shows the trajectory of the genome with $\alpha - 1$ substitutions. To calculate the curves, we divided each of the 500 Monte Carlo simulations into 20 equidistant pieces. We rounded to the nearest generation and obtained the frequency of the genome with $\alpha - 1$ substitutions at each of those 20 generation markers. Each curve shows the average of those 20 generation markers (over all establishment periods, excluding $\alpha = 1$, and over all simulations) and is plotted so that the end of the curve aligns with the mean length of the establishment period (shown in panel C).

**F.** The mean number of generations to accumulate a single beneficial substitution. We divide the number of generations to accumulate $\gamma$ substitutions by the mean number of beneficial substitutions accumulated in that time period (averaged over all simulations).
The second phase (sweep) starts at this point and ends when a genome carrying $\alpha + 1$ substitutions first appears in the population. Once a genome with $\alpha + 1$ substitutions appears, the establishment phase of this genome begins and the cycle continues.

In cytoplasmic genomes, fewer generations separate the appearance of the genome with $\alpha$ and the genome with $\alpha + 1$ substitutions than in free-living genomes (Figure 3.2C). Cytoplasmic genomes more easily become established in the population not because they are less likely to be lost by drift—in fact cytoplasmic genomes are more frequently lost to drift than free-living genomes—but because once a genome with $\alpha$ substitutions has been lost, it is more quickly regenerated (Figure 3.2D). The regeneration of the genome with $\alpha$ substitutions is proportional to the rate at which mutations occur on the genome with $\alpha - 1$ substitutions. In cytoplasmic genomes, the genome with $\alpha - 1$ substitutions increases in frequency much more quickly than in free-living genomes (Figure 3.2E). Thus, in cytoplasmic genomes, the genome with $\alpha - 1$ substitutions presents a larger target for de novo mutations, driving regeneration of the genome with $\alpha$ substitutions (Figure 3.2D). As a result, cytoplasmic genomes suffer less from clonal interference (Figure 3.3) and take less time to accumulate beneficial substitutions than free-living genomes (Figure 3.2F).

### 3.4.2 Uniparental inheritance of cytoplasmic genomes promotes the accumulation of beneficial substitutions

Meiosis introduces variation in the cytoplasmic genomes that are passed to gametes. Gametes can thus carry a higher or lower proportion of beneficial substitutions than their parent. Uniparental inheritance maintains this variation in offspring, reducing within-cell variation (Figure 3.2A) while increasing between-cell variation (Figure 3.2B). Biparental inheritance, however, combines the cytoplasmic genomes of different gametes, destroying much of the variation produced during meiosis and reducing between-cell variation (Figure 3.2B). Thus, selection is more efficient when inheritance is uniparental because there is more between-cell variation in fitness on which selection can act (Figure 3.2B). Uniparental inheritance eases the establishment of the genome with $\alpha$ substitutions (Figure 3.2C) by increasing the rate at which the genome with $\alpha$ substitutions is regenerated once lost to genetic drift (Figure 3.2D). Under uniparental inheritance, the genome with $\alpha - 1$ substitutions quickly increases in frequency (Figure 3.2E), driving the formation of the genome with $\alpha$ substitutions. Uniparental inheritance decreases clonal interference (Figure 3.3), reducing
3.4. Results

the time to accumulate beneficial substitutions compared to biparental inheritance (Figure 3.2F; see Figure B.1 for a range of different parameter values).

![Figure 3.3: Uniparental inheritance reduces clonal interference. Parameters: N = 1000, n = 50, s_b = 0.1, and a linear fitness function. The figure depicts a time-series of a single simulation, showing the proportions of genomes carrying different numbers of substitutions (we chose the first completed simulation for each comparison). To quantify the slope of declines in proportion of a genome type (equivalently, the speed at which a genome type is replaced), we report the generations (± se) for the wild type genome to drop from 100% to below 0.5% (averaged over all simulations), which we call g_{0.005}. We also report the mean number of genomes (± se) co-existing in the population (averaged over each generation and over all simulations), which we call c_g. A. In a free-living population, genomes with beneficial substitutions spread slowly through the population (g_{0.005} = 5708±31 generations). As a result, multiple genomes co-exist at any one time (c_g = 7.0 ± 0.02 genomes), increasing the scope for clonal interference. B–C. Biparental inheritance with a relaxed bottleneck (B; b = 25) and tight bottleneck (C; b = 5). Genomes with beneficial substitutions spread more quickly compared to free-living genomes (B: g_{0.005} = 2584±21 generations; C: g_{0.005} = 1377±14 generations), reducing the number of co-existing genomes (B: c_g = 4.8 ± 0.02 genomes; C: c_g = 3.8 ± 0.01 genomes). D–E. Uniparental inheritance with a relaxed bottleneck (D; b = 25) and tight bottleneck (E; b = 5). Under uniparental inheritance, genomes with beneficial substitutions spread much more quickly than free-living and biparentally inherited cytoplasmic genomes (D: g_{0.005} = 463±6 generations; E: g_{0.005} = 453 ± 6 generations). This leads to fewer genomes co-existing in the population (D: c_g = 3.1 ± 0.01 genomes; E: c_g = 2.8 ± 0.01 genomes) and low levels of clonal interference. (continued)
3.4.3 Inheritance mode is more important than the size of the bottleneck

Under biparental inheritance, a tight bottleneck decreases the variation in cytoplasmic genomes within gametes (Figure 3.2A) and increases the variation between gametes (Figure 3.2B). Consequently, under biparental inheritance beneficial substitutions accumulate more quickly than when the transmission bottleneck is relaxed (Figure 3.2F and Figure B.1). Bottleneck size has less of an effect on uniparental inheritance because uniparental inheritance efficiently maintains the variation generated during meiosis even when the bottleneck is relaxed (Figure 3.2B). When \( n \) is larger (\( n = 200 \)), a tight bottleneck reduces the time for beneficial substitutions to accumulate, but even here the effect is minor (Figure B.1C).

Importantly, the accumulation of beneficial substitutions under biparental inheritance and a tight bottleneck is always less effective than under uniparental inheritance, irrespective of the size of the bottleneck during uniparental inheritance (Figure 3.2F and Figure B.1). While a tight transmission bottleneck reduces within-gamete variation, the subsequent mixing of cytoplasmic genomes due to biparental inheritance means that cells have higher levels of within-cell variation and lower levels of between-cell variation than uniparental inheritance (Figure 3.2A–B).

3.4.4 Varying parameter values does not alter patterns

The choice of fitness function has little effect on our findings (Figure B.1). Likewise, varying the selection coefficient does not affect the patterns, although the relative advantage of uniparental inheritance over biparental inheritance is larger for higher selection coefficients (Figure B.1). Increasing the number of cytoplasmic genomes (\( n \)) increases the relative advantage of uniparental inheritance over biparental inheritance, whereas increasing the population size (\( N \)) has little effect (compare Figure B.1C with Figure B.1A).

3.4.5 Uniparental inheritance helps cytoplasmic genomes purge deleterious substitutions

Free-living asexual genomes accumulate deleterious substitutions more quickly than cytoplasmic genomes (Figure 3.4A). Biparental inheritance of cytoplasmic genomes causes deleterious substitutions to accumulate more quickly than when inheritance is uniparental (Figure 3.4). A tight transmission bottleneck slows the accumulation of deleterious substitutions under biparental inheritance, but biparental inheritance
3.4. Results

Figure 3.4: Accumulation of deleterious substitutions in the absence of beneficial mutations. Parameters (unless otherwise stated): $N = 1000$, $n = 50$, $\mu = 10^{-7}$, a concave down fitness function, and $b = 25$ (relaxed transmission bottleneck) or $b = 5$ (tight transmission bottleneck). A. Comparison with free-living genomes (linear fitness function for both free-living and cytoplasmic genomes and $s_d = 0.1$). B. Mean deleterious substitutions per cytoplasmic genome for $s_d = 0.1$. C. Mean deleterious substitutions per cytoplasmic genome for $s_d = 0.01$. Error bars are ± standard error of the mean.

always remains less efficient than uniparental inheritance at purging deleterious substitutions (Figure 3.4).

3.4.6 Uniparental inheritance reduces hitchhiking of deleterious substitutions during selective sweeps

To detect levels of genetic hitchhiking in cytoplasmic genomes, we identified the location of all “beneficial sweeps”, defined as the generation at which the genome that carries the fewest beneficial substitutions is lost from the population. Likewise, we identified the location of all “deleterious sweeps”, which is the generation in which the genome carrying the fewest deleterious substitutions is lost (note that a deleterious sweep is the same as a “click” of Muller’s ratchet [18]) (Figure B.2).

Cycling through each beneficial sweep, we identified the location of the nearest upstream deleterious sweep (i.e. in the same or in a later generation as the beneficial sweep). We measured the number of generations separating the two events and calculated the mean generations of all such instances. To obtain a “genetic hitchhiking index” ($\phi$), we normalized by dividing the mean generations by the expected number of generations for a deleterious sweep to follow a beneficial sweep (see Figure B.2 legend for how we calculate the expected number of generations). If fewer than ex-
3.4. Results

Figure 3.5: Genetic hitchhiking. \( \phi < 1 \) indicates the presence of genetic hitchhiking (the lower the value of \( \phi \), the greater the level of hitchhiking). Parameters: \( N = 1000 \), \( n = 50 \), \( \mu_b = 10^{-8} \), \( \mu_d = 10^{-7} \), and \( b = 25 \) (relaxed transmission bottleneck) or \( b = 5 \) (tight transmission bottleneck). The overall level of genetic hitchhiking in each population, measured by our genetic hitchhiking index (see Figure B.2 for details). Error bars are \( \pm \) standard error of the mean. A. Free-living comparison (linear fitness function for both beneficial and deleterious substitutions with \( s_b = s_d = 0.1 \)). For cytoplasmic genomes, B shows \( s_b = 0.1 \) while C shows \( s_b = 0.01 \). For B–C, the fitness function for beneficial substitutions is shown on the x-axis while the fitness function for deleterious substitutions is concave down.

In all cases, \( \phi < 1 \) (Figure 3.5 and Figure B.3), indicating that genetic hitchhiking plays an important role in aiding the spread of deleterious substitutions. Free-living genomes experience higher levels of hitchhiking than cytoplasmic genomes (Figure 3.5A). Uniparental inheritance reduces levels of genetic hitchhiking compared to biparental inheritance (Figure 3.5B–C and Figure B.3). Uniparental inheritance actually increases the proportion of deleterious substitutions that sweep concurrently with beneficial substitutions (Figure 3.6; leftmost bar). This occurs when the genomes...
3.4. Results

that sweep carry more than the minimum deleterious substitutions in the population. However, uniparental inheritance also increases the proportion of deleterious sweeps in which $\phi$ is large (Figure 3.6), which occur when the genomes that sweep carry the minimum number of deleterious substitutions in the population. Overall, the latter outweigh the former, leading to lower levels of genetic hitchhiking under uniparental inheritance (Figure 3.5 and Figure B.3).

Figure 3.6: Inheritance mode and the distribution of genetic hitchhiking. Parameters: $N = 1000$, $n = 50$, $\mu_b = 10^{-8}$, $\mu_d = 10^{-7}$, $b = 25$, a concave down fitness function for the accumulation of beneficial substitutions, and $s_b = 0.1$ (A) or $s_b = 0.01$ (B). A histogram that shows the distribution of hitchhiking index values for each pair of beneficial and deleterious sweeps. A beneficial sweep occurs when the genome with the fewest beneficial substitutions is lost and a deleterious sweep occurs when the genome with the fewest deleterious substitutions is lost. In both A and B, uniparental inheritance more often leads to cases in which a beneficial sweep is very closely followed by a deleterious sweep (leftmost bar). However, uniparental inheritance also leads to more cases in which the deleterious sweep is greatly separated from the beneficial sweep, indicating that genetic hitchhiking is more often suppressed under uniparental inheritance (right-hand side of the graph). Overall, uniparental inheritance leads to a higher overall hitchhiking index ($\phi$)–and thus lower levels of hitchhiking–than biparental inheritance (A. UPI: 0.79; BPI: 0.59. B. UPI: 0.86; BPI: 0.61). Blue bars pertain to uniparental inheritance, the light pink bars pertain to biparental inheritance, and the dark red bars depict overlapping bars (the dark red bar pertains to whichever color does not show on the top of the bar). (We do not plot cases in which the simulation terminates before a beneficial sweep is followed by a deleterious sweep. However, we do take these into account when generating the hitchhiking index value: see Figure B.2 for details.)
3.4.7 Uniparental inheritance promotes adaptive evolution

Cytoplasmic genomes have higher levels of adaptive evolution than free-living genomes under the same set of conditions (Figure 3.7A). Strikingly, uniparental inheritance of cytoplasmic genomes leads to a ratio of beneficial to deleterious substitutions that is two orders of magnitude higher than in free-living genomes (Figure 3.7A). Among cytoplasmic genomes, uniparental inheritance always leads to higher levels of adaptive evolution than biparental inheritance (Figure 3.7 and Figure B.4). While a tight transmission bottleneck combined with biparental inheritance increases the ratio of beneficial to deleterious substitutions, biparental inheritance always has lower levels of adaptive evolution than uniparental inheritance, regardless of the size of the transmission bottleneck (Figure B.4).

Figure 3.7: Uniparental inheritance promotes adaptive evolution. Parameters: \( N = 1000, n = 50, \mu_b = 10^{-8}, \mu_d = 10^{-7}, s_b = 0.1, \) and \( b = 25 \) (relaxed transmission bottleneck) or \( b = 5 \) (tight transmission bottleneck). A. Comparison with free-living genomes. Here, the fitness function for both beneficial and deleterious substitutions is linear. B–E shows the mean trajectory of the 500 simulations plotted every 500 generations. Here, the fitness function for beneficial substitutions is linear while the fitness function for deleterious substitutions is concave down, decreasing. We calculate the ratio of beneficial to deleterious substitutions as follows. First, we calculate the aggregated mean of the number of beneficial and deleterious substitutions for the population at generation 10,000 (average substitutions per cytoplasmic genome). Second, for each of the 500 simulations we divide the mean number of beneficial substitutions per genome by the corresponding mean number of deleterious substitutions per genome. Finally, we take the average of the ratios of the 500 simulations.
3.5 Discussion

Both theory and experiments indicate that asexual reproduction leads to lower rates of adaptive evolution than sexual reproduction [5, 7, 8, 10, 11, 15–17]. Free-living asexual organisms typically have huge population sizes, allowing them to overcome these limitations of asexual reproduction [21]. Cytoplasmic genomes, however, have much smaller effective population sizes and should be especially susceptible to these limitations of asexual reproduction [26–28]. These predictions, however, are inconsistent with empirical observations that cytoplasmic genomes can readily accumulate beneficial substitutions and purge deleterious substitutions [29, 31, 33, 35].

In this study, we help reconcile theory with empirical observations. We show that the specific biology of cytoplasmic genomes—in particular uniparental inheritance and their organization within hosts—increases the efficacy of selection on cytoplasmic genomes relative to comparable free-living genomes. Furthermore, we show that the mode of inheritance of cytoplasmic genomes has a profound effect on adaptive evolution: uniparental inheritance reduces variation of cytoplasmic genomes within cells and increases variation of fitness between cells, improving the efficacy of selection relative to biparental inheritance.

In particular, uniparental inheritance reduces competition between different beneficial substitutions (clonal interference), causing beneficial substitutions to accumulate on cytoplasmic genomes more quickly than under biparental inheritance. Uniparental inheritance also facilitates selection against deleterious substitutions, slowing the progression of Muller’s ratchet. Finally, uniparental inheritance reduces the level of genetic hitchhiking in cytoplasmic genomes, a phenomenon to which asexual genomes are especially susceptible [10, 11]. Lower levels of hitchhiking under uniparental inheritance means that beneficial (selective) sweeps are less likely to involve excess deleterious substitutions. As these genomes lacking excess deleterious substitutions spread, they remove standing variation in the population, purging genomes that carry excess deleterious substitutions and slowing Muller’s ratchet. Furthermore, both theoretical [58] and empirical [59] evidence suggest that beneficial substitutions can slow Muller’s ratchet by compensating for deleterious substitutions. By increasing the ratio of beneficial to deleterious substitutions, uniparental inheritance effectively increases the ratio of beneficial compensatory substitutions to deleterious substitutions. Thus, the accumulation of beneficial substitutions in cytoplasmic genomes not only aids adaptive evolution [33] but improves the ability of cytoplasmic genomes to resist
3.5. Discussion

Muller’s ratchet [44, 58]. Our findings thus help explain how cytoplasmic genomes are able to undergo adaptive evolution in the absence of sex and recombination.

We explicitly included a transmission bottleneck as previous theoretical work seemed to suggest that this alone could act to slow the accumulation of deleterious substitutions on cytoplasmic genomes [44]. Separate work found that host cell divisions—which act similarly to a transmission bottleneck—promoted the fixation of beneficial mutations and slowed the accumulation of deleterious mutations [50]. In contrast, yet another study found that a tight bottleneck increases genetic drift, reducing the fixation probability of a beneficial mutation and increasing the fixation probability of a deleterious mutation [45]. Here we show that these apparently contradictory findings are entirely consistent. We find that a tight transmission bottleneck indeed increases the rate at which beneficial substitutions are lost when rare (Figure 3.2D). But in a population with recurrent mutation, losing beneficial mutations when rare can be compensated for by a higher rate of regeneration, explaining how a tight bottleneck promotes adaptive evolution despite higher levels of genetic drift. Although a tight transmission bottleneck promoted beneficial substitutions and opposed deleterious substitutions when inheritance was biparental, we show that a bottleneck must be combined with uniparental inheritance to maximize adaptive evolution in cytoplasmic genomes. A transmission bottleneck is less effective in combination with biparental inheritance because the mixing of cytoplasmic genomes after syngamy largely destroys the variation generated between gametes during meiosis. For the parameter values we examined, uniparental inheritance is the key factor driving adaptive evolution, as the size of the bottleneck has little effect on the accumulation of beneficial and deleterious substitutions when inheritance is uniparental.

Our work illustrates that population genetic theory from free-living organisms cannot be blindly applied to cytoplasmic genomes. Consider effective population size ($N_e$). A lower $N_e$ leads to higher levels of genetic drift [24], and it is often assumed that low $N_e$ impairs selection in cytoplasmic genomes [25]. However, this assumes that factors which decrease $N_e$ do not alter selective pressures and aid adaptive evolution in other ways. This assumption is violated in cytoplasmic genomes as halving the $N_e$ of cytoplasmic genomes—the difference between biparental and uniparental inheritance—improves the efficacy of selection and can increase the ratio of beneficial to deleterious substitutions by 2–21 times (Figure B.4).

Although our findings apply most obviously to mitochondria and chloroplasts, they can also be applied to another type of cytoplasmic genomes: obligate endosymbionts such as Rickettsia, Buchnera, and Wolbachia. Endosymbionts share many
traits with cytoplasmic organelles, including uniparental inheritance and multiple copy numbers per host cell. Thus, uniparental inheritance may also be key to explaining known examples of adaptive evolution in endosymbionts [60, 61]
Bibliography


Bibliography


Can selection for beneficial mitochondrial mutations explain the evolution of uniparental inheritance?

4.1 Abstract

During sex, mitochondrial DNA is generally inherited from one parent rather than two. The evolutionary reasons behind this mode of inheritance remain unclear. Recent theoretical work suggests that selection against cells that carry more than one mitochondrial lineage (heteroplasmy) can explain the evolution of uniparental inheritance. Empirical evidence for this hypothesis, however, is currently restricted to a couple of studies in animals. More recently, uniparental inheritance has been shown to promote adaptive evolution in mitochondrial DNA. This finding raises an intriguing possibility: could uniparental inheritance have evolved because of fitness benefits resulting from the accumulation of beneficial mutations? Adaptive evolution of mitochondrial genomes has been found in a wide range of species, supporting the central assumption of this hypothesis. Here, we investigate theoretically whether the accumulation of beneficial mutations can explain why uniparental inheritance evolved. We find that uniparental inheritance indeed spreads at the expense of biparental inheritance, often becoming the dominant mode of inheritance. However, uniparental inheritance experiences negative frequency-dependent selection and only replaces biparental inheritance under a narrow set of assumptions. Facilitating the accumulation
of beneficial mutations is thus likely a fortuitous byproduct—rather than the ultimate cause—of the evolution of uniparental inheritance.

## 4.2 Introduction

An important consequence of sexual reproduction is the biased transmission of cytoplasmic organelles. Some of these organelles—mitochondria and chloroplasts—carry their own genomes. In general, only one partner passes on these cytoplasmic genomes to the zygote [1]. In fact, organisms go to great pains to actively avoid the biparental transmission of cytoplasm and the extranuclear genomes it contains. In male mammals, proteins in the membrane of the mitochondria are tagged with the protein ubiquitin during spermatogenesis [2]. Ubiquitin acts as a marker for protein degradation and recycling. This tagging of mitochondria delivered by sperm to the zygote means that sperm-derived mitochondria can be selectively digested after fertilization [2]. Ubiquitin-tagged mitochondrial proteins are only one of many mechanisms that have evolved to selectively destroy cytoplasmic organelles of one parent [3, 4]. While empirical evidence suggests that having cytoplasmic genomes from two parents is disadvantageous for individuals, it is less clear why this should be so.

Over the last few decades, a number of hypotheses have attempted to explain the evolution of uniparental inheritance of mitochondria. One of the first postulated that uniparental inheritance results from genomic conflict between mitochondrial and nuclear genes [5–9]. The potential for multiple different mtDNA genomes to co-exist within a cell (heteroplasmy) allows for selection at the level of the mitochondrion. Selection for fast replicating mitochondrial genomes could lead to “selfish” mitochondria, which spread rapidly but produce energy inefficiently (the mtDNA contains genes that code for proteins in the electron transport chain, which is used by the cell to produce energy (ATP) via oxidative phosphorylation [10]). According to this hypothesis, uniparental inheritance is beneficial because it allows cells to maintain metabolically efficient mitochondria despite mutations that frequently produce selfish mitochondria [5–9]. An alternative hypothesis suggests that uniparental inheritance is beneficial because it prevents cells from accumulating deleterious mitochondrial mutations (deleterious mutations are essentially selfish mutations that do not increase the replication rate of its mtDNA) [7]. Lastly, because nuclear DNA and mtDNA must work together to produce the protein complexes necessary for the respiratory chain reaction, another hypothesis suggests that uniparental inheritance may have evolved to facilitate coadaptation between nuclear and mitochondrial genes [7, 11].
While the above-described hypotheses seem plausible, mathematical models describing these mechanisms have failed to show how uniparental inheritance could evolve under realistic assumptions [7, 8]. For uniparental inheritance to replace biparental inheritance under the genome conflict and mutation clearance hypotheses, mutation rates of $>1\%$ per genome per generation must be coupled with large numbers of mitochondria per cell [7], assumptions that do not apply to any extant organisms [12]. In a model that examined whether uniparental inheritance could have evolved to facilitate mito-nuclear coadaptation, uniparental inheritance could not replace biparental inheritance under any combination of parameter values [7].

In chapter 2, we proposed a new hypothesis to explain the evolution of uniparental inheritance: the need to prevent the mixing of multiple mitochondrial types (heteroplasmy) in a single cell [12]. Our model assumes that maintaining multiple types of mitochondria in cells is costly to the organism, an assumption that has been experimentally validated in mice [13] and nematodes [14]. Unlike earlier hypotheses, selection against heteroplasmy can lead to a population with strict uniparental inheritance under a wide range of realistic modelling conditions, making it the only hypothesis with both theoretical and empirical support [12–14]. However, empirical support for selection against heteroplasmy is restricted to two studies in animals [13, 14], and it is unknown whether this phenomenon can be generalized to all eukaryotes [12].

Despite the extensive work on the evolution of uniparental inheritance, we know relatively little about how natural selection acts on mtDNA or the role that uniparental inheritance plays in this process. Historically, mtDNA evolution was thought to be dominated by slightly deleterious mutations [15, 16] that accumulate due to the asexual reproduction of mtDNA [17, 18]. Mitochondrial genomes also have a much smaller effective population size than free-living asexual organisms, making them more vulnerable to the effects of genetic drift [19]. The combination of asexual reproduction and small effective population size should theoretically curtail adaptive evolution [17, 18, 20]. Yet evidence of adaptive evolution of mitochondrial DNA is extremely strong [21–26]. Mitochondrial genome evolution in animals is consistent with frequent selective sweeps [21], and it is estimated that a quarter of non-synonymous substitutions in animals have been fixed through adaptive evolution [26]. In chapter 3, we showed that this apparent contradiction between theory and empirical evidence can be explained by the way in which mtDNA is passed across generations: uniparental inheritance promotes adaptive evolution in cytoplasmic genomes, enabling mtDNA to efficiently accumulate beneficial mutations. Indeed, cells with uniparental inheritance accumulate beneficial mitochondrial mutations about three times faster.
than cells with biparental inheritance (chapter 3). It follows that individuals with uniparental inheritance should have higher fitness than individuals with biparental inheritance because they more efficiently accumulate beneficial mutations (chapter 3).

The finding that uniparental inheritance increases the fitness of individuals by promoting the accumulation of beneficial mutations in mtDNA raises an intriguing question: is the ability of uniparental inheritance to promote the accumulation of beneficial mutations sufficient to explain its evolution? Here we investigate theoretically if selection for the accumulation of beneficial mutations could have driven the evolution of uniparental inheritance in an ancestral population in which both parents transmit their mitochondria to offspring (biparental inheritance).

4.3 Model

We base our model on the life cycle of an infinite population of single-celled eukaryotes with haploid and diploid stages that carry $n$ mitochondria, an approach used by other theoretical studies that have examined the evolution of uniparental inheritance [7, 8, 12]. Haploid cells (gametes) carry a nuclear allele that determines mitochondrial inheritance ($U$ for uniparental inheritance and $B$ for biparental inheritance). We also assume that gametes carry a nuclear allele that encodes for one of two self-incompatible mating types, 1 or 2 (whereby mating type 1 can only mate with mating type 2 and vice versa). We assume that these loci are tightly linked, preventing recombination. The population is equally split between self-incompatible mating types 1 and 2. (Later in the paper we examine a scenario in which the ancestral population has no mating types.) When we initialize the model, most gametes carry the $B$ allele (roughly evenly split between mating types 1 and 2) while a small proportion carry the $U$ allele. The $U$ allele is linked to mating type 1, simulating a population with biparental inheritance in which there has been a recent mutation from biparental to uniparental inheritance. (The exact starting conditions are 49% type $B_1$, 1% type $U_1$, and 50% type $B_2$.) Gametes contain $n/2$ mitochondria and initially all mitochondria are wild type. For simplicity, we assume that each mitochondrion carries a single mtDNA molecule. Hence, in what follows we use mitochondria and mtDNA interchangeably.

The life cycle includes four stages. First, gametes undergo random mating, in which haploid gametes unite to form diploid cells. Matings between two gametes with $B$ alleles lead to biparental inheritance, in which the mitochondria carried by the zygote are a summation of the mitochondria found in both gametes. Matings
between a gamete carrying the $U_1$ allele and a gamete carrying the $B_2$ allele lead to uniparental inheritance. Here, the zygote’s mitochondria are derived solely from the gamete carrying the $U_1$ allele. Matings between two gametes carrying the $U_1$ allele are not possible because both are of mating type 1.

The second stage of the model is mutation. Each wild type mitochondrion has a probability, $\mu$, of mutating into a beneficial mutant (we ignore back mutation). We also extend this model to two loci, in which a wild type mitochondrion can mutate into a mitochondrion with one beneficial mutation, and a mitochondrion with one beneficial mutation can mutate into a mitochondrion with two beneficial mutations.

The third stage is selection, in which we assign cells a fitness value based on the number of mutants they contain. The relative benefit of carrying a mutant is given by the selection coefficient, $s_b$, where $1 - s_b$ is the fitness of a cell with only wild type mitochondria and 1 is the fitness of a cell with only mutants. Cells with more mutants have higher fitness and will be relatively more frequent in the population after selection. Since there are few data that describe how fitness increases as the number of mutants increase in a cell, we examine three common functions to describe fitness in mitochondria: concave up, linear and concave down (Figure 4.1).

The final stage is meiosis, in which diploid cells with $n$ mitochondria produce haploid gametes with $n/2$ mitochondria. As apportioning of mitochondria into gametes is probabilistic, diploid cells with both wild type and mutant mitochondria can produce gametes with more or less of one mitochondrial type than in the parent. The simulation runs until the frequencies of $B_1$ and $U_1$ change less than $\epsilon = 10^{-12}$ across a single generation, which we consider the equilibrium frequency ($B_2$ remains at 0.5 throughout the simulation because it is the only possible mating partner for both $B_1$ and $U_1$). We track the allele ($U_1, B_1$ and $B_2$) and genotype ($B_1B_2, U_1B_2$) frequencies across time.

We run four analyses to investigate the possibility that uniparental inheritance evolved because it facilitates the accumulation of beneficial mutations. First, we examine a single locus model. While a single locus model is both easier to construct and analyze, in reality mtDNA molecules have multiple loci. Hence in our second analysis we compare the single-locus model with a two-locus model to gain insight into how multiple loci alter the dynamics of our model. Third, we relax the assumption of pre-existing mating types with tight linkage between mating type and inheritance loci. Here we examine how uniparental inheritance spreads in a model in which the ancestral population has biparental inheritance but no mating types. Fourth and finally, we examine the spread of uniparental inheritance in a fluctuating environment.
Here, we assume that a change in environment alters the selection pressure on mtDNA, causing a previously beneficial mutation to no longer benefit the cell (e.g. change of temperature as a species expands its distribution into a new habitat [27]). We then ask if the ability to adapt to such environmental fluctuations could have driven the evolution of uniparental inheritance. For mathematical descriptions of the various models, see section C.2–section C.5.

To ensure our findings are robust, we compare the frequency of uniparental inheritance at equilibrium under a wide range of parameter values (we vary $s_b$, $n$, $\mu$, and the fitness functions for the accumulation of beneficial mutations).

### 4.4 Results

#### 4.4.1 Single-locus model: uniparental inheritance invades but does not replace biparental inheritance

The frequency of uniparental inheritance in the population is determined by the frequency of the $U_1B_2$ genotype. Uniparental inheritance reaches equilibrium at a frequency lower than 1 (Figure 4.2, and Table C.1–Table C.4). At equilibrium both modes of inheritance are present in the population and all cells contain mutant mitochondria only (Table C.1–Table C.4; column 5). The frequency of uniparental inheritance at equilibrium is highest when the selection coefficient is high (Figure 4.2B–D).
Similarly uniparental inheritance reaches a higher equilibrium frequency when the fitness function relating the proportion of mutant mitochondria to cell fitness is concave up (Figure 4.2A). Mutation rate has little effect on equilibrium (Table C.1–Table C.4). For details of the dynamics underpinning equilibrium, see Figure 4.3.

![Figure 4.2: Frequency of uniparental inheritance at equilibrium.](image)

**Figure 4.2:** *Frequency of uniparental inheritance at equilibrium.* We examined different parameter values for the number of mitochondria ($n$) and selection coefficient ($s_b$), ranging from $n = 2$ to $n = 100$ (in $n = 2$ increments) and from $s_b = 0.02$ to $s_b = 0.98$ (in $s_b = 0.02$ increments). A shows the frequency of uniparental inheritance (i.e. the frequency of $U_1B_2$ cells) at equilibrium for all three fitness functions. B–D show more detailed heat maps for the concave down, linear and concave up fitness functions respectively. For B–D, the frequency of uniparental inheritance at equilibrium for a particular value of $n$ and $s_b$ is given by the colour of the graph at that point. Additional parameters: $\mu = 10^{-6}$.

### 4.4.2 Two-locus model: uniparental inheritance again invades but still does not replace biparental inheritance

Although a single locus model is easier to conceptualize, mtDNA molecules in fact contain multiple linked loci. As modelling the mtDNA as a single locus might miss important dynamics that occur for multiple loci, we also analyzed a two-locus model. Now each mitochondrion has two loci. Mitochondria can be wild type, have a single
beneficial mutation, or have two beneficial mutations. Cells that carry only mitochondrial DNA with two beneficial mutations have the highest fitness. As in the single-locus model, uniparental inheritance does not spread to fixation in any simulation. Generally, we find little difference in the frequency of uniparental inheritance at equilibrium when comparing the single-locus and two-locus models (Figure 4.4A–B and Table C.5). When the selection coefficient is high ($s_b \geq 0.5$) and the fitness function is linear or concave up, uniparental inheritance spreads to a higher equilibrium frequency in the two-locus model than in the single-locus model, but never approaches fixation (Figure 4.4C–D and Table C.5).
Figure 4.3 (previous page): Relative fitness of alleles and their change in frequency per generation. Here we examine the relative fitness of gametes to illustrate how they spread through the population. As gametes technically do not have fitness in our model, we adapt the fitness function so it applies to gametes. To calculate the relative or mean fitness of a particular allele, we multiply the frequency of each gamete carrying that allele by its fitness value, sum these products and divide the resulting number by the frequency of the allele in the population (see section C.2.6 for a mathematical description). 

A. Cells that contain more mutant mitochondria have a higher fitness than those with no or fewer mutants. Uniparental inheritance increases between-cell variance, with respect to the type of mitochondria a cell carries (wild type or mutant), on which selection can act. Selection then favours those $U_1B_2$ cells with high levels of mutant mitochondria, causing them to increase in frequency. As a result, the $U_1$ allele quickly becomes associated with mutant mitochondria, and the relative fitness of $U_1$ increases more quickly than the relative fitness of $B_1$ and $B_2$ (A-B). 

C. Because $U_1$ gametes carry a high proportion of mutant mitochondria, matings between gametes carrying $U_1$ and $B_2$ alleles result in $U_1B_2$ cells that also contain mutant mitochondria. Now the $B_2$ allele becomes associated with mutant mitochondria because $U_1B_2$ cells that are homoplasmy for mutant mitochondria produce $B_2$ gametes with mutant mitochondria. Consequently, matings between $B_1$ and $B_2$ result in $B_1B_2$ cells that carry mutant mitochondria, which in turn leads to $B_1$ gametes that carry mutant mitochondria. In effect, mutant mitochondria leak from gametes carrying the $U_1$ allele to gametes carrying the $B_2$ allele, and from gametes carrying the $B_2$ allele to gametes carrying the $B_1$ allele. This leakage is not instant, and so the relative fitnesses of $B_1$ and $B_2$ follow the relative fitness of $U_1$ after a delay (A). As the frequency of the $U_1$ allele increases, a greater proportion of gametes carrying the $B_2$ allele mate with gametes carrying the $U_1$ allele, which accelerates the leakage of mutant mitochondria from gametes with the $U_1$ allele to gametes with the $B_1$ allele (B) and slows the rate of decrease of gametes with the $B_1$ allele (D). Thus, the $U_1$ allele is under negative frequency-dependent selection and does not replace the $B_1$ allele. 

Relative fitness reaches 1 once all gametes carry beneficial mitochondria only (A–B) and the alleles stop spreading (C–D). Parameters: $n = 50$, $s_b = 0.5$, concave up fitness function, and $\mu = 10^{-9}$. 


4.4. Results

- **A** Single-locus model - low $s_b$, fitness function: concave up
  - Generations 0 to 350
  - Proportion
  - UPI trajectory
  - $B_B$ cells
  - $U_B$ cells

- **B** Two-locus model - low $s_b$, fitness function: concave up
  - Generations 0 to 600
  - Proportion
  - UPI trajectory
  - $B_B$ cells
  - $U_B$ cells

- **C** Single-locus model - high $s_b$, fitness function: concave down
  - Generations 0 to 80
  - Proportion
  - UPI trajectory
  - $B_B$ cells
  - $U_B$ cells

- **D** Two-locus model - high $s_b$, fitness function: concave down
  - Generations 0 to 160
  - Proportion
  - UPI trajectory
  - $B_B$ cells
  - $U_B$ cells
Comparing single-locus and two-locus models. A–B show simulations in which the frequency of uniparental inheritance at equilibrium does not change between the single locus and two-locus models (weak selection, $s_b = 0.1$, and a concave up fitness function). C–D show simulations in which the frequency of uniparental inheritance at equilibrium is higher in the two-locus model (strong selection, $s_b = 0.9$, and a concave down fitness function). In the single-locus models A and C, we designate cells with wild type mitochondria as $M_0$, heteroplasmy cells as $M_{0&1}$, and mutant as $M_1$. In the two locus models (B and D), we designate cells with wild type mitochondria as $M_0$, those with one mutant as $M_1$, heteroplasmy cells as $M_{0&1}$, $M_{0&2}$, $M_{1&2}$, or $M_{0,1&2}$, and those carrying two mutants as $M_2$. The solid line indicates the increase in the frequency of uniparental inheritance; below this line we find the proportion of $U_1B_2$ cells and above this line the proportion of $B_1B_2$ cells. A–B. When the selection coefficient is low, there is little qualitative difference in the dynamics between the single- and two-locus models. In both A and B, the population is initially dominated by $B_1B_2$ cells that only carry wild type mitochondria. Around a third of the way through the simulation, uniparental inheritance generates cells homoplasmic for $M_1$ in the single-locus model or $M_2$ in the two-locus model. Because of their fitness advantage, these cells increase in frequency. Concurrently, mutant mitochondria leak from $U_1B_2$ cells to $B_1B_2$ cells, converting wild type $B_1B_2$ cells into heteroplasmic $B_1B_2$ cells. As in our earlier models, eventually leakage prevents uniparental inheritance from spreading further. The single- and two-locus models differ little because $U_1B_2$ cells in the two-locus model (B) do not become homoplasmic for the intermediate ($M_1$) mitochondrial type, instead moving directly from $M_0$ mitochondria to $M_2$ mitochondria because of the small selective advantage of carrying $M_1$ mutants. In contrast in C and D, when the selection coefficient is high and the fitness function concave down, there is strong selection for mutants when they first arise. Under this scenario we do find differences in the dynamics between the single- and two-locus models. Cells carrying $M_1$ mitochondria now have a larger advantage over $M_0$ mitochondria, resulting in the $M_1$ mutant spreading through the population before the $M_2$ mutant starts spreading. Around generation 40 in D, $U_1B_2$ cells start to become homoplasmic for $M_2$ mutants, replacing $U_1B_2$ cells carrying $M_1$ mitochondria. During the spread of $M_2$ mutants, $U_1B_2$ cells also increase in frequency at the expense of $B_1B_2$ cells, until leakage occurs from $U_1B_2$ to $B_1B_2$ cells, arresting the spread of uniparental inheritance. Since the proportion of $U_1B_2$ cells increases as $M_2$ mutants replace $M_1$ mutants, uniparental inheritance reaches a higher frequency at equilibrium in the two-locus model compared with the single-locus model. Parameters for A and B: $n = 20$, $s_b = 0.1$, $\mu = 10^{-6}$, using a concave up fitness function. Parameters for C and D: $n = 20$, $s_b = 0.9$, $\mu = 10^{-6}$, using a concave down fitness function.
4.4. Results

Figure 4.5: Frequency of uniparental inheritance at equilibrium assuming no mating types. The frequency of uniparental inheritance at equilibrium is always lower when we assume no mating types, regardless of the assumptions about the type of inheritance resulting from $U \times U$ matings. A. Selection coefficient of 0.1. B. Selection coefficient of 0.5. C. Selection coefficient of 0.9. Additional parameter values: $n = 50$, $\mu = 10^{-6}$, and a concave up fitness function.

4.4.3 No mating types: uniparental inheritance reaches a lower frequency at equilibrium than when mating types are present

Here we relax the assumption of pre-existing mating types with tight linkage between the inheritance and mating type loci. We examine the spread of uniparental inheritance in a population with a single locus that does not contain mating types. Unlike our previous models, gametes carrying a $U$ allele can now mate with another gamete carrying a $U$ allele. In line with previous work [7, 12], we examine two possible outcomes of such matings—uniparental inheritance or biparental inheritance. We find little difference in the frequency of the $U$ allele at equilibrium when we compare our single-locus model with mating types with our model without mating types (Table C.6). This result is independent of the type of inheritance, uniparental or biparental, that results from $U \times U$ matings (Table C.6). Although the absence of mating types has little effect on the frequency of the $U$ allele at equilibrium, it does change which alleles can be combined during mating. This, in turn, alters the frequency of the diploid genotypes—and thus the frequency of uniparental inheritance—at equi-
librium (Figure 4.5). For any non-zero frequency of the $U$ allele, the frequency of uniparental inheritance at equilibrium is always higher when we assume pre-existing mating types (Figure 4.5). The reason is that when we assume mating types, matings between gametes carrying $B_1$ and $B_1$, $U_1$ and $U_1$, and $B_2$ and $B_2$ are not allowed. This reduces the frequency of matings between gametes carrying $B$ alleles (two combinations) more than the frequency of matings between gametes carrying $U$ alleles (one combination). As a result, fewer mating combinations lead to biparental inheritance when we assume mating types, decreasing the frequency of biparental inheritance at equilibrium (Figure 4.5).

### 4.4.4 Fluctuating selection environment: uniparental inheritance can replace biparental inheritance

Lastly, we examine the spread of uniparental inheritance in an environment in which selection pressures fluctuate, changing which allele is selectively advantageous. For example, as mitochondrial mutations aid adaptation to thermal environments [27, 28], regular fluctuations in temperature [29] could switch selection pressure on mtDNA haplotypes. While in our previous analyses we assumed that beneficial mutations could occur at any time, we now assume that the fitness consequences of mutations are context-dependent so that an allele can be selected for in a particular environment but not in another.

![Fitness functions and uniparental inheritance frequency](image)
Every 1000 generations, a change in the environment alters the selection pressure so that the currently beneficial haplotype no longer provides a benefit to the cell that carries it. The alternative haplotype now provides a fitness advantage. Initially, all cells in the population carry the beneficial (adapted) haplotype. Although mutation can change the beneficial haplotype into a maladapted haplotype, cells carrying maladapted haplotypes are selected against, meaning that only a tiny proportion of cells carry the maladapted haplotype in the first 1000 generations. As a result, uniparental inheritance does not appreciably increase in frequency. After 1000 generations, an environmental change causes selection against the currently beneficial haplotype, causing it to become maladapted to its environment. The majority of cells thus now carry a maladapted haplotype, which decreases the fitness of the cell. Uniparental inheritance soon generates cells that are homoplasmic for the new beneficial haplotype, resulting in an increase in the frequency of $U_1B_2$ cells at the expense of $B_1B_2$ cells. As in the previous variations of the model, however, the benefits of uniparental inheritance leak from $U_1B_2$ cells to $B_1B_2$ cells, causing $U_1B_2$ cells to stop spreading appreciably once the entire population carries the beneficial haplotype (note the maladapted haplotype is still generated by mutation and so exists at a very low mutation-selection frequency in $B_1B_2$ cells). The population remains in a stationary phase until the environment changes again. The frequency of uniparental inheritance thus increases in a step-wise manner.

A. Using a concave up fitness function results in uniparental inheritance replacing biparental inheritance in relatively few environmental changes under all selection coefficients. B. A linear fitness function requires more environmental changes before uniparental inheritance replaces biparental inheritance, especially when the selection coefficient is small. C. When using a concave down fitness function, uniparental inheritance only replaces biparental inheritance under very strong selection, but even then it takes many environmental changes before biparental inheritance is replaced. Assuming weaker selection, uniparental inheritance reaches an intermediate frequency. In this case, regular environmental changes cause perturbations in the frequency of uniparental inheritance, initially resulting in a decrease in frequency before the frequency returns to a steady-state (insets). Some $B_1B_2$ cells carry a small number of maladapted haplotypes before the environment changes. The combination of a concave down selection function and a small selection coefficient reduces selection against cells that carry a small number of maladapted haplotypes. This, in turn, increases the level of maladapted haplotypes in $B_1B_2$ cells at mutation-selection equilibrium. Thus when the environment changes, and the maladapted haplotype is now selected for, $B_1B_2$ cells have a very short head start on $U_1B_2$ cells, causing the frequency of uniparental inheritance to decline briefly. Additional parameter values: $n = 50, \mu = 10^{-6}$. 
As in our earlier models, we assume that the population contains two haplotypes, but now these haplotypes are either adapted or maladapted. An “adapted” haplotype will increase the fitness of the cell that carries it, whereas a “maladapted” haplotype will not. We start by assuming the population carries the beneficial (adapted) haplotype. Every 1000 generations, there is a change in the environment that causes the beneficial haplotype to become maladapted, reducing the fitness of a cell carrying only this haplotype to drop from 1 to $1 - s_b$. At the same time, the formerly maladapted haplotype becomes the new beneficial haplotype, raising the fitness of a cell carrying only this haplotype from $1 - s_b$ to 1. Mutation can now occur bidirectionally (i.e. we include mutation from adapted to maladapted and mutation from maladapted to adapted). All other aspects of the model remain the same. We run the simulations until equilibrium is reached or there have been 1000 fluctuations in the selection environment.

Under conditions that change which haplotype is selected for, uniparental inheritance goes to fixation in the majority of simulations (Figure 4.6 and Table C.7). Uniparental inheritance does not go to fixation when the fitness function is concave down and the selection coefficient is 0.1 or 0.5 (Table C.7). When uniparental inheritance becomes fixed, it takes between 5 and 122 changes in selection pressure for uniparental inheritance to replace biparental inheritance (Table C.7).

4.5 Discussion

In chapter 3, we showed that uniparental inheritance facilitates the accumulation of beneficial mutations despite the mitochondrial genomes asexual mode of reproduction. Here we ask if the advantages of accumulating beneficial mitochondrial mutations are sufficient to have driven the evolution of uniparental inheritance. Although uniparental inheritance indeed spreads in an ancestral population with biparental inheritance, uniparental and biparental inheritance always coexist at equilibrium unless selection constantly changes which haplotype is selectively advantageous. Uniparental inheritance fails to go to fixation because it experiences negative frequency-dependent selection. As the frequency of uniparental inheritance increases, beneficial mutants leak from cells with uniparental inheritance to cells with biparental inheritance (a similar effect occurs in models of uniparental inheritance that examine deleterious mutations [7, 8]). Uniparental inheritance more efficiently invades a population with biparental inheritance when beneficial mutations have strong effects on fitness ($s_b \geq 0.9$). Under these assumptions, uniparental inheritance reaches a
higher equilibrium when selection is constant, and requires fewer changes of environment when selection is variable. But mutations of this magnitude (corresponding to a $10 \times$ increase in fitness) are, at best, vanishingly rare and perhaps impossible [30]. When assuming more realistic selection coefficients ($s_b \leq 0.1$), uniparental inheritance is much less efficient at invading a population in which mitochondria inherit biparentally.

When changes to the environment alter which haplotype is selectively advantageous, uniparental inheritance was able to replace biparental inheritance. However, there are several reasons why a benefit conveyed in a fluctuating environment falls short of a general explanation for the evolution of uniparental inheritance. First, it requires a large number of environmental changes, especially when assuming realistic selection coefficients. Second, uniparental inheritance is only beneficial when at least some members of the population carry a maladapted haplotype. Whenever the population is fixed for the adapted haplotype (the majority of the time; see Figure 4.6), uniparental inheritance is selectively neutral and would be susceptible to genetic drift in a real population.

By ignoring genetic drift, our model does not take into account a more general problem with this hypothesis: a uniparental modifier must become associated with a beneficial mitochondrial haplotype for uniparental inheritance to have an advantage. In an essentially infinite population (as in our model), cells carrying the uniparental modifier mutate a proportion of wild type mitochondria to beneficial mitochondria no matter how rare the modifier is when introduced. This allows uniparental inheritance to quickly become associated with beneficial mitochondria. In a finite population, however, the uniparental modifier would arise in a single cell. In any given generation, the cell carrying the modifier for uniparental inheritance would only have a very small probability of accumulating a beneficial mitochondrial mutation. Thus, for the uniparental modifier to become associated with beneficial mitochondria, it must either occur in a cell heteroplasmic for wild type and beneficial mitochondria, or occur in a cell that by chance produces a beneficial mutation early on. Otherwise, the uniparental modifier will have no advantage and will be susceptible to being lost by genetic drift.

In our model in chapter 2, which examined selection against heteroplasmy, uniparental inheritance replaced biparental inheritance for conditions analogous to those examined here [12]. Hence, selection against heteroplasmy remains the only hypothesis with both theoretical [12] and empirical support [13, 14]. Despite the strong empirical evidence for the accumulation of beneficial mutations in mtDNA across a
4.5. Discussion

A wide range of species [21–26], we must conclude that the ability of uniparental inheritance to facilitate adaptive evolution is more likely a fortuitous side effect than the driving force behind its evolution.

Selection for the accumulation of beneficial mutations is one of several beneficial consequences of uniparental inheritance unable to explain the evolution of uniparental inheritance (using deterministic models). Similarly, uniparental inheritance selects against selfish mtDNA mutations [7–9], purges deleterious mtDNA mutations [7, 16], and facilitates mito-nuclear coadaptation [7, 11]. In contrast to previous hypotheses on the evolution of uniparental inheritance, selection for the accumulation of beneficial mutations is not sensitive to mutation rate (Table C.1–Table C.4). In effect, beneficial mutations drive themselves because they have a beneficial effect on their host. Selfish and deleterious mtDNA mutations, however, are selected against. Thus, these harmful mutations only provide a sufficient driving force for the evolution of uniparental inheritance when mutation rates are unrealistically high (note that under mito-nuclear coadaptation no mutation rate is sufficient to drive the evolution of uniparental inheritance) [7, 12].

While it is unlikely that uniparental inheritance evolved primarily to aid the spread of beneficial mutations, we have shown that selection for the accumulation of beneficial mutations does promote the spread of uniparental inheritance, even when such mutations are rare and have a small effect on fitness. Therefore, the ability of uniparental inheritance to facilitate adaptive evolution may have played a role in its evolution. Studies examining the evolution of uniparental inheritance typically limit their models to one type of mitochondrial mutation (e.g. beneficial, deleterious or selfish [7]), ignoring possible interactions between different types of mitochondrial mutations. It would be premature to discount selection for beneficial mutations simply because it alone cannot explain the evolution of uniparental inheritance. In fact, in chapter 2 we showed that when selection against heteroplasy is combined with selection for mtDNA that carry beneficial mutations, uniparental inheritance replaces biparental inheritance more quickly than selection against heteroplasy alone. Thus, while selection for beneficial mutations in mtDNA is unlikely to have driven the evolution of uniparental inheritance on its own, it may have acted in concert with selection against heteroplasy to aid the spread of uniparental inheritance.
Bibliography


Oogamy, multicellularity, and germline protect arthropods against harmful endosymbionts

5.1 Abstract

Obligate symbiotic partnerships between eukaryote hosts and bacteria have played fundamental roles in eukaryote evolution. To explain how these important symbioses came to be, we must first understand the conditions that allow these bacteria to invade and spread in eukaryotes. Some endosymbiotic bacteria, such as *Wolbachia*, can harm hosts and dramatically manipulate their host’s reproduction in order to promote their own spread. Although a range of different eukaryotes contain endosymbionts, bacteria that manipulate reproduction are generally found in arthropods. Here we investigate whether this is because endosymbionts can only invade arthropods when they manipulate host reproduction. We use a computational model to examine how three characteristics of arthropods—oogamy, multicellularity, and the presence of a germline—affect the uptake and spread of endosymbionts, using protists as a null model. We find that endosymbionts invade protists more easily than they invade arthropods, particularly when endosymbionts have a harmful effect on hosts. Endosymbionts grow quickly within protists, but selection on protist hosts that carry harmful endosymbionts is weak. As a result, harmful endosymbionts quickly spread through a population of protists. By contrast, endosymbionts grow slowly within arthropods and selection on arthropod hosts is strong. Beneficial endosymbionts spread quickly because of host-level selection, but harmful endosymbionts only
spread when they manipulate their host’s reproduction. Oogamy, multicellularity, and the presence of a germline all promote selection on hosts while impeding endosymbiont growth. Interestingly, the apportioning of endosymbionts between the soma and germline introduces a discordance in endosymbiont load between the two tissues. Because selection acts on endosymbiont load within the soma, which is not well matched to the load carried within the germline, selection becomes less effective. However, the germline also increases variation in endosymbiont load between hosts, which increases the strength of selection on hosts, compensating for the soma-germline discordance. Oogamy, multicellularity, and the presence of a germline help arthropods withstand invasion by harmful endosymbionts, without precluding the spread of beneficial endosymbionts. This provides an explanation for the prevalence of reproductive manipulation by the harmful endosymbionts of arthropod hosts.

5.2 Introduction

Complex eukaryote life can trace its origin to several key ancient endosymbiotic events. One of these, the fusion of a proto-eukaryote with an α-proteobacterium, led to modern mitochondria. Another symbiotic event, involving a cyanobacterium, led to modern day chloroplasts in algae and plants [1, 2]. Mitochondria and chloroplasts play crucial metabolic roles within cells, and their presence has sparked key innovations in eukaryotes [3, 4]. Although these organelles are the product of ancient symbioses, associations between eukaryotes and bacteria continue to play crucial roles in eukaryote evolution. Bacterial endosymbionts can have an array of positive effects in eukaryote hosts, such as compensating for limitations in their host’s metabolic system and protecting hosts against invading pathogens [5]. From the perspective of the host, symbioses can be obligate or facultative [6]. Obligate endosymbionts generally provide nutrients or otherwise aid in the metabolism of their host. Nutritional endosymbionts are widespread, being present in roughly 10% of insect species [5]. Aphids, for example, rely on an obligate bacterial endosymbiont, *Buchnera aphidicola*, for proper metabolic function, including the synthesis of essential amino acids [5].

Facultative endosymbionts are not necessary from the perspective of the host, but most facultative endosymbionts require a host in which to replicate [5, 6]. Many facultative endosymbionts have a beneficial effect on their hosts, providing benefits such as improving host nutrition and protecting against pathogens and temperature fluctuations [5]. The pea aphid, *Acyrthosiphon pisum*, is known to carry at least three facultative endosymbionts: *Hamiltonella defensa*, *Regiella insecticola*, and *Serratia*
symbiotica. H. defensa protects aphids against parasitoids, R. insecticola protects against fungal infections and improves fecundity of aphids feeding on clover, while S. symbiotic protects aphids against heat stress [5].

Despite the dependence of most facultative endosymbionts on a host for reproduction, not all endosymbiotic bacteria have a beneficial effect (here we use symbiosis in its most general sense to refer to any association—beneficial or harmful—between host and endosymbiont [7]). Some host-associated bacteria are pathogenic, causing harm as they replicate inside hosts [8, 9]. Such negative effects have not prevented some parasitic bacteria from maintaining long associations with their hosts. Certain strains of bacteria from the genus Wolbachia, for example, reduce the fecundity and survival of their arthropod hosts [9, 10]. As selection will act against hosts that carry harmful endosymbionts, parasitic endosymbionts will only persist if they find a method to invade and spread within their hosts [11].

Wolbachia has taken this to an extreme, dramatically manipulating the reproduction of its host to bias its transmission [12]. Wolbachia can cause cytoplasmic incompatibility, feminization, male killing, and parthenogenesis in its hosts [12]. Like cytoplasmic organelles, endosymbionts are mainly transmitted via the cytoplasm of females in animals. Males are therefore an evolutionary dead end for endosymbionts [7]. Feminization, male-killing, and parthenogenesis all cause infected females to bias the sex-ratio of their offspring towards females, aiding the spread of the maternally-inherited endosymbiont [12]. Wolbachia-induced cytoplasmic incompatibility causes infected males to modify their sperm in such a way that when they mate with uninfected females the union produces inviable offspring. Infected females thus gain a selective advantage over uninfected females, aiding the spread of the endosymbiont through the population [12]. Feminization can cause genotypic males infected with Wolbachia to develop as phenotypic females. Male-killing can cause infected females to abort their male offspring early in development, which frees up resources to be spent on her female offspring capable of transmitting Wolbachia [12]. Wolbachia-induced parthenogenesis affects some haplo-diploid species (mites, hymenopterans, and thrips), in which females are produced from diploid fertilized eggs and males from haploid unfertilized eggs (arrhenotoky). Parthenogenesis causes infected females to produce unfertilized eggs that develop into infected diploid females (thelytoky) instead of haploid males, aiding the spread of Wolbachia [12].

The extent of reproductive-manipulating endosymbionts in arthropods is staggering. Wolbachia alone infects two-thirds of arthropods [13], and it is but one of many bacterial endosymbionts that can manipulate host reproduction. Bacteria in
the genera *Rickettsia*, *Arsenophonus*, *Spiroplasma*, and *Cardinium* can all affect sex-determination in arthropods [14]. While we know a lot about endosymbionts that infect arthropods, bacterial endosymbionts are also widespread among protists, aiding in photosynthesis, nitrogen-fixing, and methanogenesis [15]. Yet bacterial endosymbionts in protists do not appear to manipulate host reproduction in the ways observed in arthropods. Why have the bacterial endosymbionts of arthropods, but not those of protists, evolved strategies to manipulate reproduction?

One clue might lie in the inheritance of endosymbionts via females in arthropods. Not only does this select for the endosymbiont to influence the sex ratio of their host, but it also affects the evolutionary pressures the endosymbiont is exposed to. Uniparental inheritance of cytoplasmic organelles is known to dramatically affect the accumulation of beneficial and deleterious substitutions in their genomes (chapter 3 and [16]). Uniparental inheritance aids adaptive evolution in organelle genomes by increasing the efficacy of selection, favoring those individuals that carry beneficial cytoplasmic genomes and selecting against those with harmful cytoplasmic genomes (chapter 3). Thus, parasitic endosymbionts in arthropods may struggle to spread due to strong selection against hosts that carry the harmful endosymbionts. In single-celled protists, however, both gametes contribute cytoplasm to the zygote, presumably resulting in the biparental inheritance of endosymbionts (note that little is known about endosymbiont inheritance in protists [15]). Protists that carry harmful endosymbionts may thus be more weakly selected against, allowing parasitic endosymbionts to spread more easily. Arthropods also differ from protists in that they contain multicellular tissues, as well as a distinction between soma and germline. If these arthropod traits together increase the efficiency of selection against harmful endosymbionts, it might explain why many bacterial endosymbionts have evolved an array of sex ratio distorters to aid their spread in arthropods but not in protists. Here we test this idea and investigate the effect of uniparental inheritance of endosymbionts (oogamy), multicellularity, and soma-germline separation on the spread of endosymbionts in a model of arthropods, using a model of protists as a null model.

### 5.3 Model

We developed two individual-based models to examine the spread of facultative endosymbionts in protists and arthropods. The first models a single-celled protist with...
5.3. Model

biparental inheritance of cytoplasm, while the second models a multicellular arthropod with oogamy and a distinct soma and germline. In both models, the population carries $N$ hosts and each of a host’s cells can carry anywhere from 0 to $K$ endosymbionts (where $K$ is the carrying capacity of the cell). We generally fix carrying capacity at $K = 20$, as *Drosophila* and *Aedes* infected with *Wolbachia* typically show a *Wolbachia* density per host cell of 20 or lower [17–20]. Endosymbionts can have a positive, neutral, or negative effect on their host’s fitness [6]. In the arthropod model, endosymbionts may also interfere with their host’s reproduction to bias their own transmission (we examine feminization and cytoplasmic incompatibility). Since some vertically-inherited endosymbionts can also be transmitted horizontally [7], we build two variants of each model: (1) vertical inheritance but no horizontal transmission (hereafter “vertical transmission”); and (2) vertical and horizontal transmission (hereafter “mixed transmission”). For the vertical transmission model, we examine the probability that an endosymbiont becomes fixed when a single host is inoculated with the endosymbiont (over 10,000 Monte Carlo simulations). For the mixed transmission model, we examine how long it takes for an endosymbiont to invade (over 500 Monte Carlo simulations). For both models, we consider “fixation” of the endosymbiont to occur when every individual in the population carries at least $K/2$ endosymbionts. In the model of mixed transmission, simulations stop if fixation does not occur within 50,000 generations (at about 15 generations per year for an arthropod [21], 50,000 generations represents more than 3,000 years).

5.3.1 Protist with biparental inheritance of cytoplasm

5.3.1.1 Mixed transmission

In this model, mating is controlled by nuclear-encoded self-incompatible mating types ($A$ and $a$). Mating is determined by the gametic allele, which means that only $A \times a$ matings are valid, leading to a genotype of $Aa$. The first stage of the life cycle is horizontal transmission, in which each protist has a probability ($\mu$) of receiving $b$ endosymbionts through contact with an external source (e.g. endosymbionts from another species of arthropod or from a free-living population of bacteria [7]; for simplicity, we ignore horizontal transmission of endosymbionts between conspecific hosts). The next stage is endosymbiont growth, in which endosymbionts replicate within the host. We assume that growth is logistic and that the number of endosymbionts after growth is Poisson-distributed, where the expected value ($\gamma$) depends on the growth rate ($r$) and the carrying capacity ($K$). In the next stage, selection,
5.3. Model

5.3.1.1 Horizontal transmission

In this version of the model, we initialize the population with one host that carries \( b \) endosymbionts (simulating a very rare case of horizontal transmission), whereas the remainder of the \( N - 1 \) hosts carry no endosymbionts. No further horizontal transmission occurs and each simulation stops either when all endosymbionts are lost or spread to fixation. The life cycle is the same as section 5.3.1.1 but now there is no horizontal transmission phase. For a detailed description of the protist model, see section D.5.

5.3.2 Multicellular arthropod with oogamy and soma-germline distinction

5.3.2.1 Mixed transmission

In the arthropod model, hosts are either male or female (determined by a XX/XY- or XX/XO-like genetic sex determination system). Each sex produces one type of specialized gamete (sperm for males and eggs for females). Initially, the host is an undifferentiated single-celled zygote. In the first stage of the life cycle, horizontal transmission, each zygote can receive \( b \) endosymbionts through contact with an external source (with probability \( \mu \)). The next stage is the division of the soma and germline. To simplify this process, we track two types of tissues (soma and germline). (Note that we only track one representative cell for each tissue type, and thus we...
5.3. Model

5.3. Model 98

do not explicitly model within-tissue variation; however, see meiosis for how we introduce within-tissue variation later on.) A proportion of the zygote’s cytoplasm \( p_s \) becomes the precursor of the germline and the remainder \( 1 - p_s \) becomes the precursor of the soma. We sample \( K \) pieces of cytoplasm with replacement from the soma and germline precursors to form the mature soma and germline, respectively. Our soma-germline division captures two important features of multicellular development: (1) the endosymbiont load in the soma can differ from that of the germline, depending on the density of endosymbionts in each of the precursors; and (2) the endosymbiont load in both the soma and germline can differ from the endosymbiont load in the zygote because of stochastic effects. The third stage is endosymbiont growth, in which endosymbionts replicate within both the soma and germline. In the next stage, selection, hosts are assigned a fitness value based on the endosymbiont load in their soma (the germline does not affect a host’s fitness). As in the protist model, an arthropod’s probability of surviving and reproducing (or not) is proportional to its fitness. In the next stage, meiosis, females produce eggs and males produce sperm. Since we do not explicitly track within-tissue variation (i.e. we underestimate the variance between cells within a tissue), we sample \( K \) cytoplasmic pieces for eggs with replacement from their mother’s germline (which introduces variation in the endosymbiont load of gametes produced by a germline with a given endosymbiont load). Finally, eggs and sperm randomly pair up during mating to form the zygote. Unlike in the protist model, in which both gametes can transmit endosymbionts, only the egg passes on endosymbionts in arthropods. Finally, we randomly assign a sex to the zygote.

5.3.2.2 Feminization

Here, we assume that endosymbionts can cause feminization of their hosts. (Since male-killing, feminization, and parthenogenesis all involve the endosymbiont biasing the sex-ratio towards females, for simplicity we only examine feminization.) We assume that penetrance of feminization is proportional to endosymbiont load, as indicated in empirical studies [22]. We assume that the penetrance of feminization is under negative frequency-dependent selection and is given by 

\[
1 - \frac{e_i}{K} \left(1 - \frac{\sum_{x=1}^{N} e_x}{NK}\right),
\]

where \( e_i \) is the endosymbiont load in the \( i \)th male zygote and \( \frac{\sum_{x=1}^{N} e_x}{NK} \) accounts for the endosymbiont load in the population (\( e_x \) represents the number of endosymbionts in the \( x \)th zygote). We assume negative frequency-dependent selection on the trait for feminization because feminization traits typically do not reach fixation in natural populations, as this leads to the extinction of males (negative frequency-dependent
5.4. Results

5.3.2.3 Cytoplasmic incompatibility

Cytoplasmic incompatibility means that infected males prevent uninfected females from producing viable offspring. Males never (or at least very rarely) transmit endosymbionts [7]. Cytoplasmic incompatibility instead occurs due to sperm modifications in the male germline [12], with the penetrance of cytoplasmic incompatibility being proportional to endosymbiont density in the germline [25, 26]. We assume that the penetrance (equivalently, the probability of survival) of cytoplasmic incompatibility is zero when the number of endosymbionts in the female germline is greater than or equal to the number of endosymbionts in the male germline. When the number of endosymbionts in the female germline is less than that of the male germline, we assume that the penetrance of cytoplasmic incompatibility is directly proportional to the difference in the number of endosymbionts between the male and female germline (i.e. \( p_{ij}^c = (e_i - e_j)/K \), where \( e_i \) is the number of endosymbionts in the \( i \)th male and \( e_j \) is the number of endosymbionts in the \( j \)th female).

5.3.2.4 Vertical transmission

As in section 5.3.1.2, we initialize the population with 1 zygote with \( b \) endosymbionts and \( N - 1 \) zygotes with no endosymbionts. The endosymbionts either spread to fixation or are lost. The life cycle is the same as in section 5.3.2.1, but now there is no horizontal transmission stage. For a detailed description of the arthropod model, see section D.6.

5.4 Results

5.4.1 Endosymbionts more easily invade in the protist model

When transmission is vertical, endosymbionts more easily invade hosts in the protist model, irrespective of the endosymbiont’s growth rate or the effect that the endosymbiont has on host fitness (Figure 5.1A). When transmission is mixed, endosymbionts always take less time to invade protists unless the endosymbiont has a large benefit on its host, grows slowly within its host, and horizontal transmission occurs relatively frequently (Figure 5.1B and Figure D.1).
Endosymbionts invade less readily in the arthropod model. Note that the y-axes differ between plots. The numbers above bars indicate the fixation probability (A) and number of generations until fixation (B). A. Vertical transmission. The fixation probability of endosymbionts when horizontal transmission is extremely rare. B. Mixed transmission. The number of generations before endosymbionts become fixed. Simulations were terminated once 50,000 generations had passed without fixation of the endosymbiont (when endosymbionts were costly, they never invaded in the arthropod model). Parameters: \(N = 1000\) and \(K = 20\). For “High growth rate”, \(r = 2\), and for “Low growth rate”, \(r = 0.1\). \(s_d = 0.5\) for “Large cost”, \(s_d = 0.05\) for “Small cost”, \(s = 0\) for “Neutral”, \(s_b = 0.05\) for “Small benefit”, and \(s_b = 0.5\) for “Large benefit”. The inoculum size is \(b = 5\) and the rate of horizontal transmission is \(\mu = 10^{-3}\).

In the arthropod model, endosymbionts with harmful effects do not spread, regardless of whether transmission is vertical or mixed, while in the protist model, the same endosymbionts easily reach fixation (Figure 5.1 and Figure D.1).

When transmission is vertical, beneficial endosymbionts are much more likely to reach fixation in protists than in arthropods (Figure 5.1A and Figure 5.2). But, when transmission is mixed, beneficial endosymbionts quickly invade both arthropods and protists (Figure 5.1B and Figure D.1).
5.4. Results

Figure 5.2: Beneficial endosymbionts become fixed less frequently in the arthropod model. The heat map shows the relative advantage of fixation probability under the protist model compared to the arthropod model over a range of selection coefficients ($s_b$) and growth rates ($r$). The advantage of fixation probability under the protist model is given by $P(F_p)/P(F_a)$, where $P(F_p)$ is the probability of fixation under the protist model and $P(F_a)$ is the probability of fixation under the arthropod model (fixation probability is the number of simulations (out of 10,000) in which the endosymbiont becomes fixed). Additional parameters: $N = 1000$ and $K = 20$. Endosymbionts are beneficial and transmission is vertical. Each simulation is initialized by randomly choosing one host to be inoculated with $b$ endosymbionts. We vary the inoculum size across $A$ ($b = 2$), $B$ ($b = 5$), $C$ ($b = 10$), and $D$ ($b = 20$). Higher variance in the arthropod model (Figure 5.3) leads to higher levels of genetic drift, increasing the probability that beneficial endosymbionts are lost. In addition, males cannot transmit endosymbionts in the arthropod model; if the randomly chosen host that carries the endosymbionts at initialization is male, fixation of the endosymbiont is impossible. Contrast this to the protist model, in which all randomly chosen hosts can pass on endosymbionts.
Figure 5.3: The arthropod model increases variation in endosymbiont load between hosts, promoting selection on hosts but reducing endosymbiont growth within hosts. We compare the distribution of endosymbionts among hosts in the model of protists with the model of arthropods (for the arthropod model, we depict the distribution of endosymbionts in the germline). Vertical bars show the mean endosymbiont load for each model. Parameters: $N = 1000$, $K = 20$, and $r = 1$. A. The distribution of endosymbionts produced by zygotes that carry $K/2$ endosymbionts after a single generation (averaged over 10 million zygotes). We initially omit endosymbiont growth and selection. Arthropods generate a much broader range of endosymbiont loads among their offspring than protists. B. The distribution of endosymbionts among hosts after endosymbiont growth. We start from the distribution in A and allow endosymbionts to grow. We calculate the relative level of growth as $\bar{r}_m = (e_g - e_n)/e_n$, where $e_n$ is the mean number of endosymbionts after no growth or selection, $e_g$ is the mean number of endosymbionts after growth, and $m$ is the model being examined (a for arthropod and p for protist). Because protists generate less variance in endosymbiont load among offspring than arthropods, the protist model has higher levels of growth ($\bar{r}_p = 0.48$) than the arthropod model ($\bar{r}_a = 0.26$) (see section D.3), increasing the mean endosymbiont load (vertical lines). We use the distribution in B and apply positive selection ($s_b = 0.5$; note that selection is not depicted). We calculate the relative level of selection as $\bar{\omega}_{m,s} = (e_a - e_g)/e_g$, where $e_s$ is the mean number of endosymbionts after growth and selection, and $s$ is the endosymbiont’s effect on the host (b for beneficial and d for deleterious). Because variation between hosts is high, selection is more efficient in arthropods ($\bar{\omega}_{a,b} = 0.061$) than in protists ($\bar{\omega}_{p,b} = 0.014$). When we apply negative selection ($s_d = 0.5$), again selection is more efficient in arthropods ($\bar{\omega}_{a,d} = -0.077$) than in protists ($\bar{\omega}_{a,d} = -0.019$).
5.4.2 Within-host growth drives the spread of endosymbionts in the protist model

The protist model generates low variation in endosymbiont load between hosts (Figure 5.3A). Low between-host variation increases the growth rate of endosymbionts within hosts (Figure 5.3B, Figure 5.4A,D, and Table D.1). (When growth is logistic, the number of endosymbionts generated after growth decreases with an increase in variation in endosymbiont load between hosts; see section D.3 for a general proof.) Low between-host variation also relaxes selection on hosts (Figure 5.3, Figure 5.4B,E, and Table D.1). As a result, the spread of endosymbionts in the protist model is primarily driven by endosymbiont growth within hosts rather than selection on endosymbiont load in hosts (Figure 5.3 and Figure 5.4). In addition, when within-host growth of endosymbionts is high, biparental transmission causes endosymbionts to quickly spread throughout offspring—endosymbionts grow quickly within infected hosts and all infected hosts can pass on endosymbionts (Figure 5.4C). These factors combine to drive the rapid spread of endosymbionts through the population in the protist model.

5.4.3 Harmful endosymbionts cannot spread in the arthropod model

In the arthropod model, variation in endosymbiont load between hosts is high (Figure 5.3A). This leads to poor growth of endosymbionts (Figure 5.3B and Figure 5.4A,D; section D.3) and strong selection against hosts that carry harmful symbionts (Figure 5.3). Combined, these factors prevent harmful endosymbionts from invading arthropods, even when the growth rate of an endosymbiont is high and its harmful effect is low (Figure 5.1 and Figure D.1).
5.4. Results

Figure 5.4: The arthropod model increases selection on hosts but decreases endosymbiont growth. This figure compares the arthropod model with the protist model, showing the dynamics of the vertical transmission model. Endosymbionts have a beneficial effect on their host. We only present data for the simulations in which the endosymbiont became fixed in the population. Parameters: \( N = 1000, K = 20, \) and \( s_b = 0.5 \) (large benefit). Error bars are standard error of the mean. A–C show a high growth rate \( (r = 2) \), while D–F show a low growth rate \( (r = 0.1) \). A and D. The relative growth of endosymbionts. For each generation, we compared the mean endosymbiont load per host before and after endosymbiont growth (the y-axis shows the mean increase in endosymbionts per host after growth). Protists have high levels of endosymbiont growth, whereas arthropods experience almost no endosymbiont growth. In the arthropod model, once the number of endosymbionts in a host tissue reaches the carrying capacity \( (K) \), growth is no longer possible. When a host’s germline reaches \( K \) endosymbionts, each egg will inherit \( K \) endosymbionts and pass these \( K \) endosymbionts to the zygote, preventing endosymbiont growth. This differs from the protist model, in which biparental transmission spreads endosymbionts among hosts rather than concentrating them within hosts. B and E. In the arthropod model, selection between individuals is stronger because arthropods have higher variation in endosymbiont load across generations than protists (see Figure 5.3). The y-axis shows the mean increase in endosymbiont load per host before and after selection. The strength of selection levels off in the final generations as the endosymbiont becomes fixed in the population (which removes the variation on which selection can act). C and F. The proportion of hosts carrying > 0 endosymbionts through time. C. When endosymbionts grow quickly within hosts, hosts become “infected” with endosymbionts more quickly in the protist model. In arthropods, only infected females can transmit endosymbionts to offspring, which concentrates endosymbionts within fewer hosts and minimizes endosymbiont growth. F. When endosymbionts grow slowly within hosts, hosts become infected with endosymbionts more quickly in the arthropod model. Selection on hosts now outweighs within-host growth of endosymbionts, allowing arthropods carrying beneficial endosymbionts to quickly spread.
5.4.4 Strong selection on hosts drives the spread of endosymbionts in the arthropod model

Again, endosymbionts grow slowly within hosts (Figure 5.3B and Figure 5.4A,D). But, unlike in the previous section, strong selection on hosts now favors those that carry endosymbionts (Figure 5.3 and Figure 5.4B,E). As a result, beneficial endosymbionts can spread relatively quickly in the arthropod model, although generally they spread more slowly than in the protist model (Figure 5.1 and Figure D.1).

When endosymbionts have a large benefit on their host, grow slowly within hosts, and the rate of horizontal transmission is relatively high (Figure 5.1B and Figure D.1A) then endosymbionts spread more quickly in the arthropod model. Under these conditions, strong between-host selection in the arthropod model causes the rapid spread of beneficial endosymbionts, while weak between-host selection and slow growth of endosymbionts in the protist model slow the spread of endosymbionts. Although the fixation probability of endosymbionts remains lower in the arthropod model when inheritance is vertical (Figure 5.2), “successful” fixations occur in fewer generations in the arthropod model (for the simulation depicted in Figure 5.4D–F; 21.3 ± 2.8 generations in arthropods versus 63.0 ± 8.3 generations in protists (mean ± sd)). When horizontal transmission occurs frequently, the rate at which endosymbionts spread becomes more important than the fixation probability, allowing endosymbionts to spread more quickly in the arthropod model.

5.4.5 The relative effect of oogamy, multicellularity, and soma-germline separation

It is clear that selection acts differently in the arthropod model compared with the protist model, affecting the spread of endosymbionts. The arthropod model differs from the protist model in three key aspects: oogamy (uniparental inheritance of cytoplasm), multicellularity, and a separation of the soma and germline. We will now identify how each of these differences affects the model’s dynamics by altering the arthropod model one characteristic at a time while keeping all else constant.
Figure 5.5: Oogamy, multicellularity, and germline promote selection on hosts at the expense of endosymbiont growth. We alter the arthropod model by changing one characteristic (either oogamy, multicellularity, or the soma-germline separation), while holding everything else constant. The panels show how endosymbionts are distributed among offspring across a generation (averaged over 10 million hosts). (Note that we depict the distribution of endosymbionts in the germline for all models except the no germline model.) We also present the relative level of endosymbiont growth within hosts ($\bar{r}_m = (e_g - e_n)/e_n$) and selection between hosts ($\bar{\omega}_{m,s} = (e_s - e_g)/e_g$), where $m$ takes values in $a$ for arthropod, $o$ for oogamy, $mc$ for multicellular, and $g$ for germline; see Figure 5.3. For a description of the three models, see section D.4. The vertical bars show the mean endosymbiont load for each model. Parameters: $N = 1000$, $K = 20$, $r = 1$, and $p_s = 0.2$. Removing oogamy (A), multicellularity (B), and the germline (C) reduces variance in endosymbiont load in offspring. D–F: Higher variation in endosymbiont load means that oogamy, multicellularity, and the germline all decrease endosymbiont growth (section D.3). Relative endosymbiont growth is lowest in the arthropod model ($\bar{r}_a = 0.26$), followed by removing oogamy ($\bar{r}_o = 0.31$), while it is highest when removing multicellularity ($\bar{r}_{mc} = 0.42$) and the germline ($\bar{r}_g = 0.42$). These characteristics also increase the strength of selection on hosts. When we apply selection on hosts, the arthropod model experiences stronger positive ($\bar{\omega}_{a,b} = 0.061$) and negative ($\bar{\omega}_{a,d} = -0.076$) selection than removing oogamy ($\bar{\omega}_{o,b} = 0.027$ and $\bar{\omega}_{o,d} = -0.035$, respectively), removing multicellularity ($\bar{\omega}_{mc,b} = 0.006$ and $\bar{\omega}_{mc,d} = -0.007$, respectively), and removing the germline ($\bar{\omega}_{g,b} = 0.037$ and $\bar{\omega}_{g,d} = -0.049$, respectively).
5.4. Results

5.4.5.1 Oogamy, multicellularity, and the presence of a germline all increase selection between hosts but reduce endosymbiont growth within hosts

Oogamy, multicellularity, and a germline each act to increase variation in endosymbiont load between hosts. Each characteristic therefore opposes endosymbiont growth while strengthening selection on hosts (Figure 5.5). Oogamy has the smallest effect on the distribution of endosymbiont load among offspring, while multicellularity has the largest. Counterintuitively, removing the germline—which produces a distribution midway between removing oogamy and multicellularity—results in the most efficient selection and the highest level of endosymbiont growth (Figure 5.5 and Table D.2). The higher-than-expected endosymbiont growth in the absence of the germline is easily explained: sampling endosymbionts from the entire zygote, rather than from a portion (the germline precursor), leads to a narrower distribution and thus higher endosymbiont growth††. Why removing the germline leads to a higher efficacy of selection than removing oogamy (even though the latter has a 3-fold greater variance in endosymbiont load before selection; Table D.2), however, is less obvious.

5.4.5.2 Endosymbiont loads negatively covary in soma and germline

The stronger than expected between-host selection in the absence of a germline can be explained by the relationship between soma and germline. The separation of the zygote into soma and germline tissues in arthropods has two important implications for the spread of endosymbionts. First, while a host’s fitness depends on the endosymbiont load in its soma, it is the germline, not the soma, that is actually transmitted to the next generation. Second, since the germline is derived from a smaller portion of the zygote than the soma, it experiences more genetic drift than the soma, leading to a broader distribution in endosymbiont loads in the germline than in the soma (Figure 5.6A).

Since sampling of the soma and germline from the zygote is without replacement, when the precursor of one tissue (e.g. soma) carries more endosymbionts than expected by chance, the precursor of the other tissue (germline) must carry fewer endosymbionts than expected. This leads to negative covariance between endosymbiont load in the soma and germline (relative to the load in the zygote) (Figure 5.6B). Consequently, the soma-germline separation means that the soma, which dictates host fitness, and the germline, which produces the next generation, regularly lie on the

††Before we apply growth to the distribution in Figure 5.5A–C, the soma and germline form from the zygote.
Figure 5.6: Negative covariance in endosymbionts carried by soma and germline. Parameters: \( N = 1000, K = 20, p_s = 0.2 \). Since the soma and germline are derived from different portions of the zygote, they can differ in endosymbiont load. **A.** The variation in endosymbiont load in the soma and germline tissues (averaged over 10 million hosts). Since the germline is formed from a smaller proportion of the zygote \((p_s)\) than the soma \((1 - p_s)\), variation in the number of endosymbionts is higher in the germline than in the soma.) **B.** Not only can the soma and germline differ in endosymbiont load, but they, in fact, negatively covary (Pearson’s product moment correlation coefficient: \( r = -0.421, 95\% CI = [-0.422, -0.421]; P < 0.001 \)). **B** shows covariance between the endosymbionts in soma and the endosymbionts in germline as the area of squares, plotted from the respective means (given by the dotted lines) to data points (we plot the first 300 data points here, but use 10 million to calculate the correlation coefficient). In the top right and bottom left quadrants, the endosymbionts in the soma and germline show positive covariance (both are less or both are greater than their means). In the top left and bottom right, the soma and germline covary negatively (note that the sum of the area of the squares is higher in the negative quadrants). The soma and germline negatively covary because whenever the portion of the zygote that forms the germline (germline precursor) by chance carries more endosymbionts than expected, the soma precursor must therefore carry fewer endosymbionts than expected (and vice versa) since sampling is without replacement. **C.** The size of the carrying capacity does not affect the negative covariance (Pearson’s correlation coefficient is shown in brackets). However, the relative difference in endosymbionts carried by the soma and the germline monotonically decreases as carrying capacity increases. The relative discordance in endosymbionts in soma and germline measures the magnitude of the difference in endosymbiont load between the soma and germline relative to the carrying capacity. It is given by \( d_{sg} = \frac{\sum_{i=1}^{a} |C_{t,S}^{(1,i)} - C_{t,G}^{(2,i)}|}{aK} \), where \( C_{t,S}^{(1,i)} \) represents the number of endosymbionts in the soma of the ith host, \( C_{t,G}^{(2,i)} \) represents the number of endosymbionts in the germline of the ith host, and \( a \) represents the number of hosts (here \( a = 10,000,000 \)).
opposite side of the endosymbiont load found in the zygote (Figure 5.6B). This discordance reduces the efficacy of selection—since selection acts on a soma that does not accurately reflect the load present in the germline—explaining why selection on hosts is stronger than expected in the model that removes the germline. Overall, however, the presence of a germline increases the efficacy of selection on hosts because the increased variation in endosymbiont load in offspring more than outweighs the negative covariance introduced by the soma-germline separation (compare the strength of selection between the arthropod model and no-germline model in Table D.2).

The size of the carrying capacity has little effect on the level of negative covariance (Figure 5.6C). However, the discordance in load between the soma and germline (relative to carrying capacity) approaches zero as the carrying capacity becomes very large due to a reduction in genetic drift (Figure 5.6C). (When genetic drift is low, the load in the soma and germline precursors will more closely match that of the zygote.) The smaller the germline precursor as a proportion of the zygote (e.g. late germline formation during development), the higher the variance in endosymbiont load between hosts (Figure D.2). This increases selection on hosts and decreases endosymbiont growth (Figure D.2). The smaller the germline precursor, the less negative the covariance—but the larger the discordance—between soma and germline (Figure D.2). When the germline precursor is small, its contents have less of an impact on the contents of the soma precursor, explaining the weaker negative covariance. But, a small germline precursor also increases variance in germline load relative to the soma, which increases the magnitude of the discordance between the two tissues (Figure D.2).

5.4.6 Endosymbionts can invade arthropods when they manipulate host reproduction

When endosymbionts manipulate reproduction by causing feminization or cytoplasmic incompatibility in their hosts, harmful endosymbionts can now invade in the

‡‡To explain this in a more intuitive way, imagine an extreme example in which a tiny germline precursor forms from one “unit” of cytoplasm (where each cell is composed of 100 units). If 50/100 units of a zygote’s cytoplasm contain an endosymbiont, then sampling the single unit germline precursor (which will either contain an endosymbiont or not) will have little effect on the endosymbiont load in the soma precursor (either 50/99 or 49/99 after sampling). Regardless, after sampling the soma from the soma precursor (with replacement), it will carry $\approx 50$ endosymbionts. As a result, the covariance between soma and germline will be less negative because sampling the germline has little impact on the load in the soma. After the germline develops from the germline precursor, however, it will either carry 0 or 100 endosymbionts, depending on whether the single unit germline precursor contained an endosymbiont or not. Thus, even when covariance is positive, the discordance between the two tissues will always be very large.
5.4. Results

arthropod model (Figure 5.7 and Figure D.3), but only when their cost to the host is small. Under our assumptions, feminization is a more effective strategy for harmful endosymbionts to spread in arthropods than cytoplasmic incompatibility (Figure 5.7 and Figure D.3).

Figure 5.7: By manipulating reproduction, harmful endosymbionts can invade arthropods. We compare cytoplasmic incompatibility (CI) and feminization (F) with no reproductive manipulation (none). A. Mixed transmission. The number of generations before endosymbionts become fixed. Note that the y-axes differ between plots (numbers above bars indicate the generations). Simulations were terminated once 50,000 generations had passed without fixation of the endosymbiont. Note that mildly harmful endosymbionts can now reach fixation relatively quickly under cytoplasmic incompatibility and feminization (when endosymbionts are very harmful, however, they cannot invade within 50,000 generations, even if they manipulate reproduction). B. Vertical transmission. The fixation probability of endosymbionts when horizontal transmission is extremely rare. Both cytoplasmic incompatibility and feminization have a non-zero fixation probability when endosymbionts are mildly harmful (although the fixation probability is low under cytoplasmic incompatibility). Parameters: $N = 1000$ and $K = 20$. For “High growth rate”, $r = 2$, and for “Low growth rate”, $r = 0.1$. $s_d = 0.5$ for “Large cost”, $s_d = 0.05$ for “Small cost”, $s = 0$ for “Neutral”, $s_b = 0.05$ for “Small benefit”, and $s_b = 0.5$ for “Large benefit”. The inoculum size is $b = 5$ and the rate of horizontal transmission is $\mu = 10^{-3}$. 
5.5 Discussion

Our results show that endosymbionts easily invade both protists and arthropods, unless the endosymbionts of the latter harm their hosts. Harmful endosymbionts can only invade arthropods if the endosymbiont manipulates host reproduction to bias its own transmission. The arthropod life cycle strengthens selection on hosts but minimizes endosymbiont growth rate. Strong selection on arthropods facilitates the spread of endosymbionts when they benefit hosts but suppresses their spread when they harm hosts. Our work thus provides an explanation for the widespread occurrence of reproductive manipulation by many endosymbionts of arthropods.

Three characteristics of the arthropod life cycle in particular—oogamy, multicellularity, and the presence of a germline—increase variation in endosymbiont load between hosts, promoting selection on hosts while minimizing endosymbiont growth within hosts. Under the assumptions made in our models, multicellularity was the most important factor determining the strength of selection on hosts, followed by oogamy and the presence of a germline respectively. Surprisingly, removing the germline led to stronger selection on hosts than expected based on between-host variation. This occurs because the soma and germline negatively covary in endosymbiont load. As the soma is the tissue exposed to selection, while the germline produces the next generation, this discordance reduces the efficacy of selection.

Other work has shown that factors which increase variance in cytoplasmic genomes (of which endosymbionts are an example) slow the spread of harmful cytoplasmic genomes. Uniparental inheritance, transmission bottlenecks, and mitotic cell divisions—all of which increase between-host variation in cytoplasmic genomes—slow the spread of harmful cytoplasmic genomes (chapter 3; [16, 27–31]). Some of the factors that protect against the invasion of harmful endosymbionts (poor growth and no transmission through males) also reduce the spread of endosymbionts that benefit their hosts; in our arthropod model, such endosymbionts spread more slowly than in the protist model unless endosymbiont growth rate is low, endosymbionts have a large beneficial effect, and horizontal transmission is relatively high. These results appear to contradict our findings in chapter 3, where we showed that that factors which increase variance between hosts (uniparental inheritance and a transmission bottleneck) promote the spread of cytoplasmic genomes carrying beneficial mutations. The discrepancy between the two studies is due to the presence of within-host selection in the model of endosymbionts. While mutations in cytoplasmic genomes primarily
5.5. Discussion

affect the host’s fitness, our model of endosymbionts includes selection on the endosymbionts themselves (within-host growth), as well as selection on the host. While the arthropod model strengthens selection on hosts, it also reduces growth of endosymbionts within hosts. However, when between-host selection is stronger than within-host selection (Figure 5.1B and Figure D.1), beneficial endosymbionts spread quickly. A similar interaction of between-host selection and within-host growth has been shown in endosymbionts that are subject to transmission bottlenecks [30].

The notion that the presence of a germline will impact endosymbionts is not new [32, 33], although previous arguments have differed substantially from each other and from ours. In our study, the presence of a germline increases variation in endosymbiont load between hosts, strengthening selection on hosts and decreasing endosymbiont growth. Hurst [32] and Frank [33] proposed that a segregated germline would reduce the “virulence” of harmful endosymbionts. Hurst envisioned that competition between cells to form the germline would be won by cells with low parasite diversity, leading to the evolution of low parasite virulence [32]. Frank envisaged the germline as a randomizing force that prevented an endosymbiont from biasing its transmission to the germline. By removing the incentive for the endosymbiont to act selfishly, the endosymbiont’s virulence would decrease as its fitness aligns with that of the host [33]. Although the arguments differ, all predict that the presence of a germline will reduce the impact of within-host selection, thus more closely aligning the fitness of an endosymbiont with that of its host.

Our work also identified a potential evolutionary disadvantage to the soma-germline separation. The apportioning of endosymbionts between soma and germline leads to negative covariance in endosymbiont load between the two tissues. As the soma-germline discordance applies equally to heteroplasmy of mitochondria and chloroplasts, this may have important implications for selection on cytoplasmic genomes. By uncoupling the tissue exposed to selection (soma) from the one that produces the next generation (germline), the soma-germline separation reduces the efficacy of selection.

The observation that oogamy, multicellularity, and a germline reduce the spread of both harmful and beneficial endosymbionts may provide an evolutionary answer to the widespread occurrence of manipulation of host reproduction by endosymbionts of arthropods. Not all endosymbionts known to manipulate their host’s reproduction have negative effects on their host. For example, Wolbachia can increase host fitness by improving fecundity or survival in beetles, wasps, and flies [34–38]. Likewise, Rickettsia sp. nr. bellii, a facultative symbiont of the invasive whitefly Bemisia
**5.5. Discussion**

Tabaci, causes feminization in its host but also improves survival of its host, accelerates its host’s development [21], and confers resistance to the pathogenic bacterium *Pseudomonas syringae* [39]. Combined, *Rickettsia’s* fitness benefits and reproductive manipulation allowed it to sweep through whitefly populations of the south-western United States in just six years, dramatically affecting the ecology and invasiveness of the whitefly [21].

The transition from a facultative to an obligate symbiosis involves the alignment of an endosymbiont’s fitness with that of its host. One way for the host to accomplish this is to take control over the replication of the endosymbiont; for example, by exporting the genes for the endosymbiont’s replication to the host’s nucleus [40, 41]. We have shown that several features of the arthropod life cycle dramatically reduce the scope for growth in facultative endosymbionts, acting to align the fitness of the endosymbiont with that of its host. Since the endosymbionts of arthropods can do little to promote their spread by replicating within hosts, it is not surprising that they have instead evolved mechanisms to promote their spread by manipulating their host’s reproductive system.


[1.1.3, 2.2, 2.3, 2.5, 3.2, 3.3, 4.2, 4.3, 4.4.3, 4.5, 5.5, 6.1, 6.1.1, 6.1.3, †, 6.3, ‖, A.5, A.5.1, C.2]


Cytoplasmic genomes are exposed to a unique set of evolutionary pressures. They exist within hosts in multiple copies. Consequently, cytoplasmic genomes can be selected on at different levels: within-host, by modulating their replication rate, and between-host, through their effects on hosts. Arguably the most striking evolutionary feature of cytoplasmic genomes is their transmission through a single parent. In chapter 2 and chapter 4, I investigated two novel hypotheses to explain the evolution of uniparental inheritance. I conclude that one of these hypotheses — selection against heteroplasmy — has stronger theoretical support than any existing hypothesis. The origin of uniparental inheritance is not simply an interesting evolutionary question, but something that has important implications for the spread and evolution of cytoplasmic genomes in general. In chapter 3, I show that uniparental inheritance can explain the unexpectedly high levels of adaptive evolution found in animal mitochondrial genomes. Finally, in chapter 5, I show that uniparental inheritance (oogamy) of facultative endosymbionts, multicellularity, and the presence of a germline protect arthropods against invasion by harmful or pathogenic cytoplasmic endosymbionts. These traits so strongly select against harmful endosymbionts that such endosymbionts must evolve traits to manipulate their host’s reproduction or they cannot spread.

6.1 Evolution of uniparental inheritance

Conflict between selfish cytoplasmic genomes and the nuclear genome of their host has been the dominant evolutionary explanation for uniparental inheritance for the past four decades [1–4]. Other hypotheses, which have received less attention, include
improved purging of deleterious mutations [5] and co-adaptation between the nuclear and mitochondrial genomes [5, 6]. While uniparental inheritance indeed selects against selfish cytoplasmic genomes and deleterious mutations, and for mito-nuclear co-adaptation, strict uniparental inheritance does not evolve in mathematical models of these hypotheses under realistic assumptions [5]. In chapter 4, I add selection for the accumulation of beneficial mutations to that list. While uniparental inheritance improves the accumulation of beneficial mutations, strict uniparental inheritance only evolves under a narrow set of assumptions. Of the hypotheses that had been modeled before this thesis, the selfish conflict hypothesis requires the least restrictive biological conditions in order for uniparental inheritance to replace biparental inheritance (notwithstanding the requirement for mutations to have selfish effects) [5]. Below I make a case for why I believe this hypothesis falls short of a general explanation for the evolution of uniparental inheritance.

6.1.1 Selfish conflict hypothesis

The first concern about the selfish conflict hypothesis was pointed out long ago by its first modelers [4, 7]. Unless the mutation that causes uniparental inheritance occurs in a population that is polymorphic for the selfish mutation, it has no adaptive value. Since selfish cytoplasmic genomes are expected to be either quickly lost or to sweep through a population with biparental inheritance, the window during which the trait for uniparental inheritance can arise is very narrow [4, 7]. (A similar concern applies to the hypothesis for the accumulation of beneficial mutations (chapter 4). The window is even narrower in the hypothesis I tested in chapter 4, as the mutation that causes uniparental inheritance must occur within an individual that is polymorphic for a beneficial cytoplasmic genome mutation (chapter 4).) The second (and more worrying) concern relates to the biological conditions postulated by the selfish conflict hypothesis.

Modeling has shown that uniparental inheritance is subject to negative frequency-dependent selection under the selfish conflict hypothesis [3, 5]. However, this dynamic can be overridden if four conditions are met: (1) selfish genomes replicate faster than wild type genomes; (2) selfish genomes arise at a frequency of at least 1% per cytoplasmic genome per generation; (3) selfish genomes have a lethal effect on hosts when they are the only genome present; and (4) hosts carry at least \( n = 50 \) cytoplasmic genomes with a bottleneck size of \( n/2 \) ( [5]; Fig. S6 in the supplementary material). Conditions (1) and (2) provide a strong driving force behind the selfish mutant, while condition (3) provides a large selective advantage for hosts that can
maintain an association with wild type cytoplasmic genomes (or, at least, with few selfish genomes). Condition (4) lowers variance in cytoplasmic genomes between hosts carrying the trait for uniparental inheritance. In combination with conditions (1) and (2), condition (4) prevents uniparental inheritance from efficiently clearing selfish genomes from the population, which is what causes negative frequency-dependent selection against the trait [5].

Conditions (1) and (3) seem reasonable as there are biological examples of selfish mitochondrial genomes, some of which have the ability to kill their host [8–11]. Condition (4) is not met by many single-celled eukaryotes, which often have tighter mitochondrial bottlenecks than required (e.g. slime mold gametes carry \( \approx 15 \) mitochondria [12] while yeast only have a couple of mitochondria [13]). But condition (4) is satisfied in animals, which have more relaxed bottlenecks [14]. It is hard to believe that condition (2), however, is met in any extant organism.

Although Ephrussi reported that the petite mutation in yeast (section 1.1.3) occurs at a frequency of around 1% [15], this estimate is almost certainly much higher than the mutation rate per mitochondrial genome per generation. A mutation accumulation study on *Saccharomyces cerevisiae* estimated the mtDNA mutation rate at \( 12.2 \times 10^{-9} \) per nucleotide per host cell division [16]. At a genome size of 85 kb [17], that gives a mutation rate of \( 0.001^* \) per mtDNA genome. Since it is implausible that petite mutants are generated at a rate 10-fold that of the total mutation rate per mtDNA genome, it seems likely that Ephrussi’s estimate of 1% encompassed multiple mitochondria and multiple cell divisions (Ephrussi’s petite colonies developed over a few days [15] whereas yeast divide approximately every two hours [18]). In animals, generating selfish mitochondria at a rate of 1% per generation seems equally implausible. Using the “animal-like” parameter values for mutation and mtDNA genome length in chapter 3 \( (u_d = 10^{-7} \) and \( l = 20000 \) bp), the mutation rate per mtDNA genome is 0.002.

Using extant mitochondrial genomes to make predictions about the ancient past is not without its problems, of course, as mitochondrial genomes were likely structured quite differently back when uniparental inheritance evolved. If, during the evolution of uniparental inheritance, the mitochondrial genome was more like the genome of an extant facultative endosymbiont such as *Wolbachia*, then it would have been substantially larger than it is today (e.g. the genome of *Wolbachia* is \( \approx 1 \) Mb [19]). I am not aware of any direct estimates of the mutation rate in *Wolbachia*, but a study using a genomics approach estimated it to be 100-fold lower than the mutation rate

\[
1 - (85000) \left( 12.2 \times 10^{-9} \right)^0 (1 - 12.2 \times 10^{-9})^{85000}
\]
at synonymous sites in the mitochondrial genome [20]. A genome size of 1 Mb and a mutation rate of $10^{-9}$ per nucleotide per generation gives the same mutation rate per genome as yeast (0.001).

All this means that the mutation rate required by models of the selfish conflict hypothesis to override the negative frequency-dependent selection against uniparental inheritance is simply too high. And this is before taking into account two additional problems. The first additional problem is that only a subset of the genome is likely to be a candidate for mutations that result in a selfish phenotype as envisioned in models (i.e. enhanced replication and lethal for the host when homoplasmic) [5]. For example, in animal mitochondrial genomes, the non-coding control region appears to be the candidate region for selfish mutations [10], a part of the genome that is typically less than 10% the length of the mtDNA [21]. The second problem is that presumably the vast majority of mutations that do occur in the candidate region will not lead to a selfish phenotype. When these factors are taken into account, the effective mutation rate per genome becomes even more unrealistic. The stringent biological conditions required by the selfish conflict hypothesis—and the fact that no extant organism appears to satisfy them—poses a serious challenge to the ability of the prevailing hypothesis to explain the evolution of uniparental inheritance.

### 6.1.2 Selection against heteroplasmic hypothesis

In chapter 2, I propose that selection against heteroplasmicity is the best explanation for the evolution of uniparental inheritance. Unlike selfish conflict, selection against heteroplasmicity is not sensitive to mutation rate, number of cytoplasmic genomes per host, or assumed cost of heteroplasmicity (chapter 2). Its fundamental tenet—that heteroplasmicity imposes a cost on the host—is backed up by an empirical study, which showed that heteroplasmicity disrupted the behavior and physiology of mice [22]. Since the publication of chapter 2 [23], more empirical evidence for costs related to heteroplasmicity has emerged in *C. elegans* [24]. This study found that heteroplasmicity involving a mtDNA mutant with a 3.1 kb deletion increased embryo lethality by 23-fold. Even transient heteroplasmacy, due to delayed removal of paternal mtDNA, led to a 5–7 fold increase in embryo lethality. Delayed removal of paternal mtDNA in crosses involving two slightly different wild type mtDNA haplotypes also led to significantly reduced survival in embryos [24].

In chapter 2, I assumed that incompatible haplotypes were generated through mutations accumulated within a single generation. Although this hypothesis is not
sensitive to mutation rate, if mtDNA haplotypes must be highly divergent for heteroplasmy to be costly, it is doubtful whether mutation could generate sufficient differences over a single generation (or even over a few generations, if mutations were able to accumulate). The fact that even slightly different wild type mtDNA haplotypes can cause a cost of heteroplasmy in *C. elegans* [24] supports the generality of selection against heteroplasmy as an explanation for the evolution of uniparental inheritance (the mtDNA haplotypes in the study that found a cost of heteroplasmy in mice were quite divergent [22]). Another scenario that could lead to heteroplasmy of incompatible mitochondria is the hybridization of populations with divergent haplotypes; similar incompatibilities can occur with mito-nuclear interactions in hybrid offspring [25–27].

While the empirical evidence for selection against heteroplasmy is accumulating, we still do not know how or why heteroplasmy imposes a cost on the host. As discussed in section 2.5, two possible explanations are Lane’s hypothesis of disrupted signaling between mitochondria and the nucleus [28, 29], and Wallace’s hypothesis of incompatibility between mitochondrial polypeptides encoded by different mtDNA genomes [22, 30]. I do not want to elaborate on these hypotheses, nor support one over the other, as the dearth of empirical evidence means that such comments would be little more than speculation. However, I will point out that if Wallace’s hypothesis is correct, it may have salient implications for the evolutionary forces on mutations that cause selection against heteroplasmy. If mismatched mito-mito protein interactions cause heteroplasmy to be costly [22, 30], then it is likely that mismatched mito-nuclear protein interactions should also be costly. Since the number of mitochondrial proteins in the electron transport chain of animals (approximately 13 [31]) is fewer than the number of nuclear proteins in the transport chain (approximately 73 [31]), it is a reasonable expectation that mutations in mitochondrial protein genes will be more likely to disrupt mito-nuclear interactions than to disrupt mito-mito interactions. Therefore, models that examine a cost of heteroplasmy may need to combine these effects with mito-nuclear interactions, as the same underlying mechanism will likely lead to both phenomena. Since selection for mito-nuclear coadaptation exerts strong negative frequency-dependent selection on the spread of the trait for uniparental inheritance, it is possible that uniparental inheritance would not be able to evolve in a system in which both dynamics were operating, especially if mito-nuclear effects were more dominant than mito-mito effects.

A group of authors in the 1980s and 1990s—Cosmides and Tooby, Hoekstra, Hurst, and Frank—were the first to propose that a mixture of cytoplasmic genomes might
impose a cost on the host [2, 32–34], although to my knowledge these ideas were never modeled in the context of uniparental inheritance. One of the earliest ideas, proposed by Cosmides and Tooby, was that biparental inheritance might harm the host if different cytoplasmic genomes acted to destroy one another [2]. Using a verbal argument, they proposed that the evolution of anisogamy might have been a nuclear adaptation to minimize conflict between different cytoplasmic genomes [2]. If competition between two cytoplasmic genomes was highest when both were present at 50% (which seems reasonable), then Cosmides and Tooby’s verbal argument would lead to very similar dynamics as those seen under selection against heteroplasmy (chapter 2) [2].

Another prominent idea was that a host should attempt to reduce the diversity of different cytoplasmic genomes it carries. By reducing diversity among its cytoplasmic genomes, a host could minimize the virulence of its cytoplasmic genomes (i.e. the harmful effect of the genome on the host) by aligning the fitness of the cytoplasmic genome with its own fitness [32–34]. However, this latter idea leads to quite different dynamics than selection against heteroplasmy, as low virulence will take some time to evolve [32, 34]. Thus, a modifier that causes a host to homogenize its cytoplasmic genomes will initially have no adaptive benefit—even if the host carries a single type of cytoplasmic genome, that genome will still be a virulent form [32, 34].

To lead to similar dynamics as selection against heteroplasmy, carrying a mixture of cytoplasmic genomes must be costly, and that cost must immediately be lessened when carrying a single type of cytoplasmic genome (chapter 2). Other ideas discussed by Hurst and Frank had closer parallels to selection against heteroplasmy [33, 34]. Hurst proposed that different types of endosymbionts might occupy different niches within a host cell. A higher diversity of endosymbionts might therefore deplete the host of different resource types, a scenario that Hurst suggested might be more costly than depletion of a single resource type [33]. Frank discussed the possibility that endosymbionts could directly compete with other species within a host if diversity were high, such as the secretion of a toxin (to which the endosymbiont is itself protected against) to kill competitors [34]. If such a toxin harmed the host, and were only secreted when unrelated endosymbionts were detected, then this scenario could lead to dynamics similar to those found when assuming selection against heteroplasmy (chapter 2). To my knowledge, there is no direct empirical evidence supporting the scenarios proposed by Hurst and Frank. They are nonetheless interesting propositions and highlight that the dynamics seen under selection against heteroplasmy need not be restricted to incompatibilities between mitochondrial haplotypes.
6.1.3 Relaxed inheritance hypothesis

Since the publication of chapter 2 [23], a new hypothesis has been proposed to jointly explain the evolution of uniparental inheritance and protection against the accumulation of deleterious mutations through Muller’s ratchet [35]. Greiner and colleagues propose that, as in the selfish conflict hypothesis, uniparental inheritance evolved to protect against selfish cytoplasmic genomes. In addition, they suggest that the threat of Muller’s ratchet drives relaxation of uniparental inheritance. This leads to bouts of paternal leakage or biparental transmission, allowing recombination between cytoplasmic genomes, which avoids mutational meltdown through Muller’s ratchet [35]. The hypothesis of Greiner and colleagues is, unfortunately, couched solely in “for the good of the population” terms [35]. Below I will show that, when framed in evolutionary terms, this hypothesis does not hold.

First, as discussed in section 6.1.1, for selfish conflict to explain uniparental inheritance one must envision very specific conditions: a population in which harmful selfish mutants arise at an extraordinary frequency [5]. Under these conditions, a trait for uniparental inheritance can spread because it allows hosts to maintain association with wild type cytoplasmic genomes despite the constant generation of selfish genomes [5].

Second, selection cannot act on populations to avoid mutational meltdown via Muller’s ratchet. Muller’s ratchet is an outcome, not an evolutionary force. Rather, selection acts on individuals that carry the fewest slightly deleterious mutations in their cytoplasmic genomes. If the efficacy of selection on individuals is weak, then those individuals with the fewest slightly deleterious mutations can be lost and Muller’s ratchet is a natural outcome. If the efficacy of selection on individuals is strong, then Muller’s ratchet will progress slowly or halt. Thus, the evolutionary force underlying Muller’s ratchet is simply the efficacy of selection on individuals.

It follows that the same general evolutionary force that drives uniparental inheritance also slows Muller’s ratchet: increased efficacy of selection for hosts that carry few or no deleterious mutations in their cytoplasmic genomes. But the strength of the selection pressure generated by selfish conflict and by Muller’s ratchet are markedly different. Clearly, no evolutionary mechanism can favor selection against the very gradual accumulation of slightly deleterious mutations in cytoplasmic genomes (chapter 3 and [36]) under Muller’s ratchet over selection against harmful selfish cytoplasmic genomes that are produced in huge numbers each generation†, which is what drives

†For a mutation rate of 1% per genome per generation [5], a population of 1,000 hosts, each with 50 cytoplasmic genomes, would be bombarded with 500 selfish genomes every generation.
and stabilizes the evolution of uniparental inheritance under the selfish conflict hypothesis [5]. Thus, the claim of Greiner and colleagues that uniparental inheritance evolved due to selfish conflict is incompatible with the claim that Muller’s ratchet can relax organelle inheritance and lead to periods of paternal leakage or biparental inheritance [35]. The evolutionary forces driving and stabilizing uniparental inheritance are much stronger, and occur over shorter time-scales, than the forces underlying Muller’s ratchet. Indeed, modeling has previously shown that even low levels of paternal leakage have devastating effects on fitness when selfish genomes arise within a population [37]. Finally, my work in chapter 3 suggests that recombination between cytoplasmic genomes in animals is likely to exacerbate Muller’s ratchet, not slow it, as I will discuss in the next section.

### 6.2 Evolutionary consequences of uniparental inheritance

Theory has long predicted that asexual genomes, such as those of cytoplasmic genomes, should suffer low rates of adaptive evolution because they lack recombination [38–43]. These predictions have been confirmed in a wide range of empirical studies, supporting the evolutionary advantage of recombination [44–47]. Yet, widespread evidence indicates that adaptive evolution is pervasive in the mitochondrial genomes of animals [48–58]. In chapter 3, I show that existing theory does not account for the specific evolutionary forces acting on cytoplasmic genomes, in particular their uniparental mode of inheritance. When these forces are taken into account, cytoplasmic genomes can undergo high levels of adaptive evolution despite their lack of sex and recombination.

#### 6.2.1 Adaptive evolution in cytoplasmic genomes does not require recombination

There is an oft-repeated notion in the literature that low levels of recombination, made possible by paternal leakage or occasional biparental inheritance, prevents mitochondrial genomes from accumulating deleterious mutations and succumbing to Muller’s ratchet [35, 59–62]. Certainly paternal leakage can occur in animals, and it may even be relatively widespread [63–65]. Recombination between animal mitochondrial DNA has also been observed [66, 67], but whether this occurs often enough to alter evolutionary dynamics is unknown [68, 69]. For example, studies observing paternal leakage in natural populations have failed to detect recombinant mtDNA [64].
Whether recombination between cytoplasmic genomes is possible is a separate issue to whether recombination between cytoplasmic genomes meaningfully increases genetic variation in natural populations [68]. While recombinant mtDNA has been found in humans [70], showing that recombination is possible, the fact that there is only one documented case strongly suggests that recombination plays no meaningful role in the evolution of the human mitochondrial genome.

For recombination between cytoplasmic genomes to generate genetic variation, heteroplasmy is required [68]. Heteroplasmy can be generated through paternal leakage (or biparental inheritance) or a de-novo mutation in a homoplasmic cell. Although I am not aware of evidence either way, I suggest that de-novo mutations are highly unlikely to result in genetic variation through recombination. Homologous recombination will only produce genetic variation when the recombining haplotypes differ at more than one site; such haplotypes will rarely be generated through de-novo mutations. Paternal leakage, however, could create conditions conducive to producing genetic variation through recombination by producing a heteroplasmic zygote that contains divergent haplotypes [63, 68]. But, I showed in chapter 3 that heteroplasmy resulting from biparental inheritance reduces the efficacy of selection on hosts and can reduce the level of adaptive evolution up to 30-fold (Figure 3.7A). Thus, the conditions that promote recombination between cytoplasmic genomes also impair adaptive evolution (chapter 3). This means that any benefits of recombination with respect to alleviating Muller’s ratchet must overcome the acceleration of Muller’s ratchet due to inefficient selection against deleterious mutations. Given that uniparental inheritance and homoplasmy dramatically improve the ability of cytoplasmic genomes to undergo adaptive evolution (chapter 3), it seems unlikely that occasional recombination will adequately compensate.

Obviously, the proposition that recombination between cytoplasmic genomes will impair adaptive evolution should be formally modeled before we can draw any firm conclusions about the interplay between paternal leakage, recombination, and adaptive evolution in cytoplasmic genomes. To my knowledge, the role of recombination on cytoplasmic genome evolution has never been modeled. Instead the proposition that recombination between cytoplasmic genomes might alleviate Muller’s ratchet [35, 60, 62] relies on the findings of models based on free-living asexual genomes, which suggest that low levels of recombination are sufficient to halt Muller’s ratchet (e.g. [71, 72]). In chapter 3 I clearly showed that one cannot simply extrapolate findings from free-living asexual genomes to cytoplasmic genomes. I strongly suspect the same is the case with respect to the alleged role of recombination in relaxing Muller’s
ratchet. Importantly, my work in chapter 3 shows that cytoplasmic genomes can readily undergo adaptive evolution with no recombination whatsoever. This removes the need to invoke the unsupported proposition that paternal leakage and recombination act to alleviate Muller’s ratchet in cytoplasmic genomes.

### 6.2.2 Two-level selection and adaptive evolution

My results in chapter 5 differ from some of those in chapter 3 with respect to the dynamics of beneficial cytoplasmic genomes. In chapter 3, uniparental inheritance and a transmission bottleneck promoted the accumulation of beneficial substitutions in cytoplasmic genomes. In chapter 5, however, uniparental inheritance (oogamy), multicellularity, and the presence of a germline generally slowed the spread of beneficial endosymbionts. This difference is somewhat surprising because both chapters examined how uniparental inheritance (and other factors that increase between-host variation) affect selection on beneficial cytoplasmic genomes, be they beneficial substitutions in cytoplasmic genomes (chapter 3) or endosymbionts that increase the fitness of their host (chapter 5). The most important difference between the two models lies in the presence or absence of within-host selection, a factor that has previously been shown to affect selection on endosymbiont genomes [73]. In chapter 3, I focused on cytoplasmic organelles. As the within-host replication of these organelles appears to be primarily under host control [74, 75], there is little scope for within-host selection on cytoplasmic genomes. Facultative endosymbionts, which are the focus of chapter 5, can grow within their hosts, and this growth is not regulated by the host [76, 77]. As such, endosymbionts are subjected to both within-host and between-host selection, in contrast to the cytoplasmic genomes in chapter 3.

In addition, in chapter 3 I allowed cytoplasmic genomes to accumulate mutations, while in chapter 5 I tracked the spread of a single type of endosymbiont without mutation. However, horizontal transmission of endosymbionts in the mixed transmission model in chapter 5 functions in a similar manner to mutation\(^1\). For both mutation and horizontal transmission, each host can gain beneficial or deleterious cytoplasmic genomes with some probability. Consequently, some of the scenarios examined in chapter 5 have close parallels with those examined in chapter 3.

\(^1\)Obviously, the details are substantially different. For horizontal transmission, there is a constant probability of obtaining \(b\) endosymbionts. For mutation, the probability of a particular mutation varies depending on the types of genomes within a host. It is also highly improbable that more than one genome within a host mutates in a given generation. Nevertheless, both processes generate beneficial cytoplasmic genomes.
Specifically, the dynamics in chapter 3 and chapter 5 should be consistent when two conditions are met: (1) horizontal transmission in chapter 5 occurs at roughly the same rate as mutation in chapter 3; and (2) between-host selection dominates within-host selection in chapter 5 (i.e. endosymbionts have a large benefit on their hosts and grow slowly within hosts). In, chapter 3, I examined two mutation rates for beneficial mutations: $10^{-8}$ and $10^{-9}$ per nucleotide per generation. If a host carries 50 cytoplasmic genomes, each with a genome size of 20 kb, then the mutation rate per host per generation is 0.01 ($10^{-8}$) or 0.001 ($10^{-9}$). These mutation rates are equivalent to the rate of horizontal transmission in Figure D.1A and Figure 5.1B. Indeed, for these rates of horizontal transmission, when endosymbionts provide a large benefit to their hosts and grow slowly within hosts, endosymbionts spread more quickly in the arthropod model than in the protist model (bottom right bar graph in Figure D.1A and Figure 5.1B). In both chapter 3 and chapter 5, the mechanism is similar: uniparental inheritance and similar traits\(^5\) cause beneficial cytoplasmic genomes to sweep to fixation more quickly. Rapid selective sweeps can compensate for the higher levels of genetic drift associated with these traits, which can increase the probability that a single beneficial cytoplasmic genome is lost from the population (Figure 3.2D and Figure 5.2; also see [78]). So long as the rate of horizontal transmission or mutation is not extremely low\(^6\), then losing any one beneficial mutation carries little cost because shortly thereafter another beneficial mutation can take hold and rapidly sweep to fixation.

In chapter 5, when endosymbionts are beneficial, within-host selection (i.e. endosymbiont growth) acts in the same direction as between-host selection. This explains why beneficial endosymbionts generally spread more quickly in the protist model—which strongly promotes endosymbiont growth—than in the arthropod model. When endosymbionts are harmful, however, within-host selection acts in the opposite direction to between-host selection. While harmful endosymbionts grow quickly within hosts and spread in the protist model, they are unable to invade in the arthropod model unless they manipulate reproduction. This explains why so many endosymbionts that impose a cost on arthropods interfere with their host’s reproduction [80, 81]: if they do not evolve such traits, they simply will not persist.

\(^5\)i.e. transmission bottleneck, multicellularity, the presence of a germline—all of which increase between-host variation in cytoplasmic genomes

\(^6\)Beneficial mutations would have to be extraordinarily rare for fixation probability to matter more than the speed at which selective sweeps occur in chapter 3. Parameters of $N = 1000$, $l = 20000$, and $\mu_b = 10^{-9}$ generate beneficial mutations at a more than sufficient rate, and increasing $N$ to a more realistic value [79] would further increase the number of beneficial mutations that are generated in the population.
The ease with which harmful endosymbionts invade in the protist model implies that protists should be overrun with costly endosymbionts. If so, there will be strong selection pressure on protists to evolve counter-adaptations or risk extinction [82]. It may well be that costly endosymbionts exert a significant extinction pressure on protists; it would be very difficult to detect such events, given that protists are understudied relative to animals [83], and their small size and often inaccessible habitats make them difficult to study in the wild. Alternatively, protists may have evolved counter-defences against harmful endosymbionts that were not accounted for in chapter 5. Unlike arthropods, protists actively hunt and consume bacteria [84]. In fact, they so efficiently consume bacteria that predation by protists is responsible for a range of evolutionary adaptations in bacteria [84–86]. Harmful endosymbionts may therefore face a number of obstacles before they even make it to the cytoplasm of protists.

6.3 Future directions

While I have provided explanations for several important evolutionary questions, my work has also raised a number of additional questions that could be examined in future work. Models examining the evolution of uniparental inheritance have thus far taken a deterministic approach [3–5, 23, 87]. While these models are easier to analyze, this approach obscures the role of genetic drift. For example, in the hypotheses for uniparental inheritance that are subject to negative frequency-dependent selection (selfish conflict, deleterious mutations, mito-nuclear coadaptation, and beneficial mutations), it is possible that the trait for uniparental inheritance could drift to fixation once it reaches equilibrium. Stochastic models might shed light on important evolutionary dynamics that are missed in deterministic models. Existing models have also focused on a single type of mitochondrial mutation (e.g. deleterious, selection against heteroplasmy, or mito-nuclear interactions [5, 23]). In reality, these different types of mutations coexist and may interact with each other in counterintuitive or unforeseen ways, which may have implications for the evolution of uniparental inheritance.

Throughout this thesis, I assumed that mitochondrial genomes could not undergo recombination. As there is evidence of recombination among cytoplasmic genomes

---

∥ The polymorphic equilibrium between uniparental and biparental inheritance due to negative frequency-dependent selection is not a stable equilibrium ([5]; Figure 1B). Therefore, once uniparental inheritance reaches equilibrium, selection will not apply downward pressure to prevent it from drifting to fixation.

∗∗ For example, models examining deleterious or beneficial mutations in isolation will miss the effect of genetic hitchhiking.
(limited in animals [68] but more common in plants and fungi [61]), it would be worth examining the effect of recombination between cytoplasmic genomes on adaptive evolution. This would support or refute the notion that infrequent recombination between cytoplasmic genomes is sufficient to prevent the progression of Muller’s ratchet [35, 59–62].

Mitochondrial genomes differ widely between plants, animals, fungi, and protists in both structure and function. Why are the huge mitochondrial genomes of plants [88] structured so differently to the streamlined mitochondrial genomes of animals [21]? Patterns of genome evolution can even differ widely within the same molecule. Why are the protein-coding mtDNA genes of animals highly conserved relative to nuclear genes [89] while their mitochondrial tRNA genes are less conserved relative to nuclear genes [90, 91]? While I showed how cytoplasmic genomes can undergo adaptive evolution despite lacking recombination, there remain many unanswered questions about the specifics of cytoplasmic genome evolution.

How does adaptive evolution differ between single-celled eukaryotes with uniparental inheritance and animals with oogamy, multicellularity, and a soma-germline separation? There is increasing evidence that purifying selection can act between oocytes within the germline [92–95], favoring those oocytes that do not carry harmful mitochondrial mutations. Exactly how this within-tissue quality control works is unknown, but it implies that there may be a third level of selection in animals with a distinct soma-germline: within-cells, between-cells within the germline, and between-hosts. This may help explain the unexpectedly high levels of purifying selection in animal mitochondrial genomes [92] and could be incorporated into future models.

The presence of a germline in animals could be further explored in other ways. In chapter 5, I showed that the soma-germline separation introduces negative covariance in endosymbiont loads between these tissues. Mitochondrial heteroplasmy during embryo development will also cause the soma and germline to negative covary with respect to the proportion of each haplotype in the different tissues. Since this should impair selection against deleterious mitochondrial mutations, it suggests a potential evolutionary disadvantage to the soma-germline distinction. This raises interesting questions. How does the soma-germline discordance affect the evolution of cytoplasmic genomes and has it played a role in the evolution of the timing of the soma-germline separation during development? It is clear that, while this thesis has furthered our understanding of the evolution of uniparental inheritance and its implications for the spread and evolution of cytoplasmic genomes, many questions are still waiting to be answered.
Bibliography


Bibliography


[58] Rollins LA et al. (2016) Selection on mitochondrial variants occurs between and within individuals in an expanding invasion. *Molecular Biology and Evolution*. [1.2.4, 6.2]


A.1 Supplementary figures
Figure A.1: **Uniparental inheritance replaces biparental inheritance when we consider three mitochondrial types.** Parameters: $n = 20$, $\mu = 10^{-6}$, $c_h = 0.1$, and concave fitness (unless indicated otherwise). **A.** $U_1$ replaces $B_1$ leading to complete uniparental inheritance. **B.** Number of generations to reach equilibrium for varying costs of heteroplasmy under concave and convex fitness. $U_1$ is more advantageous when it takes fewer generations to reach equilibrium. **C.** Number of generations to reach equilibrium for varying mutation rates. $U_1$ replaces $B_1$ under all tested values of $\mu$. **D.** Number of generations to reach equilibrium for different number of mitochondria per cell (as the model with three mitochondrial types is very computationally-intensive, we were unable to examine values of $n$ above 40).
Figure A.2: *Fitness and distribution of cell types at a lower mutation rate.* Parameters: $n = 20$, $\mu = 10^{-7}$, and concave fitness. (A) Relative advantage of the two genotypes throughout time. The distribution of $U_1B_2$ is shown in (B) and $B_1B_2$ is shown in (C).
Figure A.3: **Fitness and distribution of gamete types at a lower mutation rate.**
Parameters: $n = 20$, $\mu = 10^{-7}$, $c_h = 0.2$ and concave fitness. **A.** Relative advantage of the three alleles throughout time. The distribution of $U_1$ is shown in **B**, $B_1$ is shown in **C** and $B_2$ is shown in **D**.
A.1. Supplementary figures

Figure A.4: **Uniparental inheritance spreads more quickly when U<sub>1</sub> mutates in a heteroplasmic B<sub>1</sub> gamete compared to a homoplasmic gamete.** The case in which U<sub>1</sub> mutates into a homoplasmic cell is shown in A–D, while the heteroplasmic case is shown in E–G. We let U<sub>1</sub> mutate in the most heteroplasmic B<sub>1</sub> gamete that had a frequency of > 0.01 at the equilibrium between B<sub>1</sub> and B<sub>2</sub> (which was a gamete with two mutant mitochondria). U<sub>1</sub> gametes appear at generation 0. The heteroplasmic U<sub>1</sub> gametes are quickly lost (first few generations in E and F), leading to much higher levels of U<sub>1</sub> gametes with mutant mitochondria (compare F with B). In turn, this leads to much higher levels of heteroplasmy in B<sub>1</sub> and B<sub>2</sub> (generations 0–450 in G and H), which results in a steeper drop in $\omega_{B_1}$ and $\omega_{B_2}$ (compare E with A) and a faster production of B<sub>2</sub> gametes that carry mutant mitochondria (about generation 400 in H compared to 1400 in D). Consequently, U<sub>1</sub> replaces B<sub>1</sub> in about half the number of generations when it mutates in a heteroplasmic B<sub>1</sub> gamete compared to a homoplasmic gamete.
Figure A.5: $U_1$ replaces $B_1$ \textit{when $U_1$ is introduced at lower frequencies}. $U_{in}$ is the frequency of $U_1$ when it mutates from the $B_1$ gamete. It takes longer for $U_1$ to replace $B_1$ when it starts at a lower frequency. Parameters: $n = 20$, $\mu = 10^{-7}$, $c_h = 0.2$ and concave fitness.

Figure A.6: \textit{Relative fitness and levels of heteroplasmy for different costs of heteroplasmy}. Parameters: $n = 20$, $\mu = 10^{-7}$, and concave fitness. (Note that the y-axis differs by two orders of magnitude between D–F.) Selection against heteroplasmy is strongest in A and D, which leads to very low levels of heteroplasmy in $B_1B_2$ cells because few $B_2$ gametes with mutant mitochondria are produced. Consequently it takes many generations before $\omega_{B_1B_2}$ starts to drop substantially and $U_1$ takes longer to replace $B_1$ as a result. In B and E, selection against heteroplasmy is lower, which leads to more heteroplasmic $B_1B_2$ cells and a faster spread of $U_1$. While the levels of heteroplasmy rise dramatically as selection against heteroplasmy weakens further (C and F), this cannot compensate for the fact that heteroplasmic $B_1B_2$ cells are weakly selected against. Thus, $U_1$ takes longer to replace $B_1$ compared to B and E.
Figure A.7: Generations for $U_1$ to replace $B_1$ for different numbers of mitochondria per cell and costs of heteroplasmy. $U_1$ takes increasingly longer to replace $B_1$ as the number of mitochondria per cell and cost of heteroplasmy increases. Parameters: $\mu = 10^{-7}$ and concave fitness.
Figure A.8: Relative fitness and levels of heteroplasmy under the three fitness functions. Parameters: \( n = 20 \), \( \mu = 10^{-4} \), and \( c_h = 0.2 \). Selection against heteroplasmy is weakest under the concave fitness function, followed by linear and convex fitness respectively (see Figure 2.1A). Under concave fitness (A–D), this leads to higher levels of \( U_1 \) gametes that carry the mutant haplotype (B). In turn, this leads to more \( B_2 \) gametes that carry the mutant haplotype (D) and higher levels of heteroplasmy in \( B_1B_2 \) cells (which can be seen through the high levels of heteroplasmy in the \( B_1 \) gametes (C)). Levels of heteroplasmy in the \( B_1 \) gamete are lower under linear (E–H) and convex (I–L) fitness functions because these functions select more strongly against heteroplastic cells. \( U_1 \) replaces \( B_1 \) in a similar number of generations for each fitness function under these set of parameters because lower levels of heteroplasmy under linear and convex fitness is offset by stronger selection against heteroplastic \( B_1B_2 \) cells (see Figure 2.1F). \( U_1 \) spreads at a similar rate for all three fitness functions when \( c_h = 0.2 \).
Figure A.9 (previous page): **Non-neutral haplotypes with strong effects.** Parameters: \( s_d = s_a = 0.1, \ n = 20, \ \mu = 10^{-7}, \ \text{and} \ c_h = 0.2. \) In all these cases, the accumulation of mutations is modeled using a concave fitness function. Concave/convex, as noted on the Figure, refers to the fitness function governing selection against heteroplasmy. \( U_1 \) replaces \( B_1 \) unless both the accumulation of mutations and selection against heteroplasmy are modeled using a concave function (black-solid and red-dashed lines). In these cases, the advantageous and deleterious scenarios converge to the same polymorphic equilibrium with a low level of uniparental inheritance. In the advantageous concave case (black-solid), mutant mitochondria quickly replace wild type mitochondria as the dominant haplotype (this corresponds to the rapid rise in \( U_1 \) frequency to about 0.16). \( B_1 \times B_2 \) matings are now less costly because almost all matings involve mutant mitochondria (this stops the rapid spread of \( U_1 \)). At this point, the advantageous and deleterious scenarios are actually equivalent to each other (mutating from the advantageous mutant to the “normal” wild type is the same as mutating from the “normal” wild type to the deleterious mutant since the selection coefficients are the same in both cases). Thus, both cases converge to the same equilibrium. \( U_1 \) does not replace \( B_1 \) because it is more advantageous for \( B_1B_2 \) cells to have low levels of heteroplasmy (but large numbers of mutant mitochondria) than it is for \( U_1B_2 \) to have a low frequency of cells that are homoplasmic for the wild type haplotype (recall that \( U_1B_2 \) cells quickly segregate into homoplasmic cells; thus, mutations from the advantageous mutant to wild type become segregated in homoplasmic wild type cells). This is because the mutant haplotype confers such a large advantage when \( s_a = 0.1. \) Contrast this with the advantageous case in which selection against heteroplasmy is convex (blue-dotted). Here, too, \( U_1 \) stops its rapid spread once the mutant haplotype has replaced the wild type haplotype (\( U_1 \) frequency of about 0.35), but now the \( U_1 \) slowly spreads until it replaces \( B_1. \) Because selection against heteroplasmy is convex in this case, which translates into stronger selection against low levels of heteroplasmy compared to concave selection, it is now less advantageous for \( B_1B_2 \) cells to have low levels of heteroplasmy than it is for \( U_1B_2 \) to have a low frequency of cells that are homoplasmic for the wild type haplotype. As a result, \( U_1 \) slowly replaces \( B_1. \)
A.1. Supplementary figures

Figure A.10: Relative advantage and distribution of gamete types when mutations are advantageous, neutral and deleterious. In A–D, $U_1$ spreads more quickly when under $s_a = 0.001$. $U_1$ produces gametes that carry the mutant haplotype, which then rapidly spread in $U_1B_2$ cells due to their fitness advantage (compare B to F). Because the mutant haplotype is linked to $U_1$ (and to $B_2$ through $U_1 \times B_2$ matings), $U_1$ spreads more rapidly in this scenario. In I–L, $U_1$ produces much fewer gametes that carry the mutant haplotype (compare J to F) because $U_1B_2$ cells that only carry the mutant haplotype are more strongly selected against than $U_1B_2$ cells that are homoplasmic for wild type mitochondria. This reduces the number of $B_2$ gametes with mutant haplotypes (L), which reduces heteroplasmy in $B_1B_2$ cells (seen in the lower level of heteroplasmy in $B_1$ gametes (K)) and slows the spread of $U_1$. 

$s_a = 0.001$ 

$s_d = 0.001$ 

Neutral
Figure A.11: Probability of recombination does not affect equilibrium when it is above a threshold. A. $P_r$ is below the threshold, which leads to the fixation of the $U_1B_2$ genotype. When $P_r$ is above the threshold (B–D), the trajectories of the $U_1B_2$ and $U_2B_1$ genotypes converge. When $P_r$ is above the threshold but is much lower than 0.5 (B), the frequency of $U_1B_2$ is initially higher than that of $U_2B_1$ (because the $U_2$ gamete initially arises due to recombination between $U_1$ and $B_2$ gametes during $U_1 \times B_2$ matings). But, because there are initially more $U_1B_2$ cells than $U_2B_1$ cells, there are more recombination events in $U_1B_2$ cells than in $U_2B_1$ cells, which drives the $U_1 : U_2$ ratio towards $U_2$. The frequency of $U_2$ continues to increase relative to $U_1$ until $P(U_1) = P(U_2)$, at which point the frequencies of $U_1B_2$ and $U_2B_1$ converge (B).
Figure A.12: Uniparental inheritance is not maximized when \( U \times U \) matings have biparental inheritance and fitness is concave. Additional parameters: \( n = 20, \mu = 10^{-4}, \) and \( c_h = 0.2. \) Under these conditions, the frequency of uniparental inheritance at equilibrium is 0.118. A. The relative advantage of the three genotypes. B–D show the relative proportion of the UB (B), BB (C) and UU (D) cell types, where the heteroplasmy category includes all cells with any level of heteroplasmy. E–F show a more detailed distribution of the UB (E), BB (F) and UU (G) cell types at generation 80,000. H–I show the distribution of gamete types for the U (H) and B (I) alleles. The fitness of UU (\( \tilde{\omega}_{UU} \)) drops sharply in the very early stages of the simulation (A) because of an increase in U gametes homoplasmic for mutant mitochondria (H). \( \tilde{\omega}_{UU} \) decreases because U gametes homoplasmic for mutant mitochondria mate with U gametes homoplasmic for wild type mitochondria, which leads to highly heteroplasmic UU cells. Shortly afterwards (up until about \( 1 \times 10^4 \) generations), U gametes homoplasmic for mutant mitochondria drop in frequency (H). \( \tilde{\omega}_{UU} \) increases because there are now fewer \( U \times U \) matings between mutant and wild type gametes. But it never reaches the level of \( \tilde{\omega}_{BB} \) (A) because U gametes homoplasmic for mutant haplotypes remain (compare H to I). Thus, although UU cells have a lower proportion of heteroplasmic cells, these cells have higher levels of heteroplasmy than BB cells (compare F with G; recall that cells with low levels of heteroplasmy are weakly selected against when fitness is concave). Because uniparental inheritance is under negative frequency-dependent selection, it does not spread to its maximum level.
Uniparental inheritance is maximized when $U \times U$ matings have biparental inheritance and fitness is linear or convex. Additional parameters: $n = 20$, $\mu = 10^{-4}$, $c_h = 0.2$, convex fitness and assuming no mating types. A. The relative advantage of the three genotypes. B–D show the relative proportion of the UB (B), BB (C) and UU (D) cell types, where the heteroplasmy category includes all cells with any level of heteroplasmy. E–F show a more detailed distribution of the UB (E), BB (F) and UU (G) cell types at generation 60,000. H–I show the distribution of gamete types for the $U$ (H) and $B$ (I) alleles. Compared to the situation under concave fitness (Figure A.12), when fitness is linear or convex a negligible amount of $U$ gametes are homoplasmic for mutant mitochondria (H). Consequently, there is no noticeable difference between $U \times U$ and $B \times B$ biparental inheritance matings (compare F to G) and $\bar{\omega}_{UU}$ converges with $\bar{\omega}_{BB}$ (A). Because $U \times B$ matings are more advantageous than the biparental inheritance matings (A), uniparental inheritance spreads to its maximum level under a linear or convex fitness function.
A. The concave and convex fitness functions used in the model that considers three mitochondrial types. 

A. A three-dimensional fitness function that is similar to the two-dimensional concave function. Low levels of heteroplasmy incur a relatively small fitness cost.

B. A three-dimensional fitness function that is similar to the two-dimensional convex function. Low levels of heteroplasmy incur a relatively large fitness cost.

Figure A.14: The concave and convex fitness functions used in the model that considers three mitochondrial types. A. A three-dimensional fitness function that is similar to the two-dimensional concave function. Low levels of heteroplasmy incur a relatively small fitness cost. B. A three-dimensional fitness function that is similar to the two-dimensional convex function. Low levels of heteroplasmy incur a relatively large fitness cost.
**A.2 Supplementary tables**

*Table A.1: General model: $n = 20$ and $\mu = 10^{-4}$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium.*

<table>
<thead>
<tr>
<th>Parameters</th>
<th>n</th>
<th>$\mu$</th>
<th>Fitness</th>
<th>$c_h$</th>
<th>Generations</th>
<th>UPI frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>concave</td>
<td>0.01</td>
<td>9,559</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>concave</td>
<td>0.1</td>
<td>3,152</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>concave</td>
<td>0.2</td>
<td>2,822</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>concave</td>
<td>0.5</td>
<td>3,440</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>concave</td>
<td>1</td>
<td>6,921</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>linear</td>
<td>0.01</td>
<td>7,266</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>linear</td>
<td>0.1</td>
<td>2,712</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>linear</td>
<td>0.2</td>
<td>2,708</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>linear</td>
<td>0.5</td>
<td>4,411</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>linear</td>
<td>1</td>
<td>14,837</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>convex</td>
<td>0.01</td>
<td>5,973</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>convex</td>
<td>0.1</td>
<td>2,531</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>convex</td>
<td>0.2</td>
<td>2,841</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>convex</td>
<td>0.5</td>
<td>6,229</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>$10^{-4}$</td>
<td>convex</td>
<td>1</td>
<td>35,175</td>
<td>1</td>
</tr>
</tbody>
</table>
Table A.2: General model: \( n = 20 \) and \( \mu = 10^{-7} \). Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the \( U_1B_2 \) genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table A.3: General model: \( n = 20 \) and \( \mu = 10^{-10} \). Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the \( U_1B_2 \) genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( \mu )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-10} )</td>
</tr>
</tbody>
</table>
Table A.4: General model: \( n = 50 \) and \( \mu = 10^{-4} \). Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the \( U_1B_2 \) genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n )</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
<tr>
<td>50</td>
<td>10^{-4}</td>
</tr>
</tbody>
</table>
Table A.5: General model: $n = 50$ and $\mu = 10^{-7}$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ convex</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ convex</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ convex</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ convex</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-7}$ convex</td>
</tr>
</tbody>
</table>

Table A.6: General model: $n = 50$ and $\mu = 10^{-10}$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ concave</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ linear</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ convex</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ convex</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ convex</td>
</tr>
<tr>
<td>50</td>
<td>$10^{-10}$ convex</td>
</tr>
</tbody>
</table>
Table A.7: General model: $n = 100$ and $\mu = 10^{-4}$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>
Table A.8: General model: $n = 100$ and $\mu = 10^{-7}$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$ $\mu$ Fitness $c_h$</td>
<td>Generations UPI frequency</td>
</tr>
<tr>
<td>100 $10^{-7}$ concave 0.01</td>
<td>276,437 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ concave 0.1</td>
<td>213,769 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ concave 0.2</td>
<td>296,028 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ concave 0.5</td>
<td>903,755 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ concave 1</td>
<td>8,724,257 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ linear 0.01</td>
<td>185,960 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ linear 0.1</td>
<td>202,757 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ linear 0.2</td>
<td>373,481 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ linear 0.5</td>
<td>2,170,969 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ linear 1</td>
<td>42,742,860 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ convex 0.01</td>
<td>149,338 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ convex 0.1</td>
<td>236,540 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ convex 0.2</td>
<td>584,651 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ convex 0.5</td>
<td>5,688,431 1</td>
</tr>
<tr>
<td>100 $10^{-7}$ convex 1</td>
<td>10,766,453 1</td>
</tr>
</tbody>
</table>

Table A.9: General model: $n = 100$ and $\mu = 10^{-10}$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$ $\mu$ Fitness $c_h$</td>
<td>Generations UPI frequency</td>
</tr>
<tr>
<td>100 $10^{-10}$ concave 0.01</td>
<td>8,841,849 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ concave 0.1</td>
<td>6,671,529 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ concave 0.2</td>
<td>9,198,479 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ concave 0.5</td>
<td>27,795,323 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ concave 1</td>
<td>256,319,556 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ linear 0.01</td>
<td>5,368,580 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ linear 0.1</td>
<td>5,912,779 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ linear 0.2</td>
<td>10,981,277 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ linear 0.5</td>
<td>71,317,346 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ convex 0.01</td>
<td>4,247,519 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ convex 0.1</td>
<td>6,821,227 1</td>
</tr>
<tr>
<td>100 $10^{-10}$ convex 0.2</td>
<td>17,243,560 1</td>
</tr>
</tbody>
</table>
### Table A.10: General model: $n = 200$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>200</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>200</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>200</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>200</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>200</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>200</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>200</td>
<td>$10^{-10}$</td>
</tr>
</tbody>
</table>

### Table A.11: Deleterious model: $n = 20$ and $\mu = 10^{-7}$. Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the $U_1B_2$ genotype at equilibrium. Fitness is the fitness function governing the cost of heteroplasmy. The accumulation of deleterious mutations is modeled using a concave fitness function.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
</tbody>
</table>
Table A.12: Deleterious model: \( n = 100 \) and \( \mu = 10^{-7} \). Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the \( U_1B_2 \) genotype at equilibrium. Fitness is the fitness function governing the cost of heteroplasmy. The accumulation of deleterious mutations is modeled using a concave fitness function.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( \mu )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-7} )</td>
</tr>
</tbody>
</table>

Table A.13: Advantageous model: \( n = 20 \) and \( \mu = 10^{-9} \). Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the \( U_1B_2 \) genotype at equilibrium. Fit (het) is the fitness function governing the cost of heteroplasmy. Fit (accum) is the fitness function that governs the accumulation of advantageous mutants.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( \mu )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-9} )</td>
</tr>
</tbody>
</table>
Table A.14: Advantageous model: \( n = 100 \) and \( \mu = 10^{-9} \). Generations means the number of generations to reach equilibrium. UPI frequency is the frequency of the \( U_1B_2 \) genotype at equilibrium. Fit (het) is the fitness function governing the cost of heteroplasmy. Fit (accum) is the fitness function that governs the accumulation of advantageous mutants.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n )</td>
<td>( \mu )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-9} )</td>
</tr>
</tbody>
</table>

Table A.15: Non neutral scenario when we consider three mitochondrial types. Values represent the number of generations (\( \times 10^3 \)) to reach equilibrium for varying values of \( s_a \) (advantageous selection coefficient) and \( s_d \) (deleterious selection coefficient). When both haplotypes have equal fitness, the population reaches equilibrium in \( 26(\times 10^3) \) generations under the same set of parameters. Uniparental inheritance becomes fixed in all cases. Parameters: \( n = 20 \), \( \mu = 10^{-7} \), \( c_h = 0.1 \) and concave fitness.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_d = s_a = 10^{-2} )</td>
<td>( s_d = s_a = 10^{-3} )</td>
</tr>
<tr>
<td>Deleterious</td>
<td>2428</td>
</tr>
<tr>
<td>Advantageous</td>
<td>725</td>
</tr>
</tbody>
</table>
Table A.16: Recombination and no mating types for $U \times U$ with uniparental inheritance: $n = 20$ and $\mu = 10^{-4}$. Gen means the number of generations to reach equilibrium ("r" represents recombination, while "nmt" represents no mating types). UPI freq is the frequency of uniparental inheritance at equilibrium ($U_1U_2$ for recombination and $UU$ for no mating types). Additional parameters: $P_r = 0.5$ (for recombination).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>

Table A.17: Recombination and no mating types for $U \times U$ with uniparental inheritance: $n = 20$ and $\mu = 10^{-7}$. Gen means the number of generations to reach equilibrium ("r" represents recombination, while "nmt" represents no mating types). UPI freq is the frequency of uniparental inheritance at equilibrium ($U_1U_2$ for recombination and $UU$ for no mating types). Additional parameters: $P_r = 0.5$ (for recombination).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-7}$</td>
</tr>
</tbody>
</table>
Table A.18: Recombination and no mating types for $U \times U$ with uniparental inheritance: $n = 100$ and $\mu \equiv 10^{-4}$. Gen means the number of generations to reach equilibrium ("r" represents recombination, while "nmt" represents no mating types). UPI freq is the frequency of uniparental inheritance at equilibrium ($U_1U_2$ for recombination and $UU$ for no mating types). Additional parameters: $P_r = 0.5$ (for recombination).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
<th>Gen (r)</th>
<th>Gen (nmt)</th>
<th>UPI freq (r)</th>
<th>UPI freq (nmt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$ $\mu$ Fitness $c_h$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 $10^{-4}$ concave 0.01</td>
<td>7,045,472</td>
<td>9,960,598</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>100 $10^{-4}$ linear 0.01</td>
<td>7,023,898</td>
<td>9,930,778</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>100 $10^{-4}$ concave 0.5</td>
<td>2,143,254</td>
<td>2,715,278</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>100 $10^{-4}$ linear 0.5</td>
<td>5,335,851</td>
<td>6,729,353</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
A.2. Supplementary tables

Table A.19: Recombination and no mating types for $U \times U$ with biparental inheritance: $n = 20$ and $\mu = 10^{-4}$. UPI is maximized at 0.5 when $U \times U$ have biparental inheritance (see main text for explanation). UPI (r) (frequency of uniparental inheritance at equilibrium, assuming recombination) is evenly split between the $U_1B_2$ and $U_2B_1$ genotypes at equilibrium, while the UPI (nmt) (frequency of uniparental inheritance at equilibrium, assuming no mating types) refers to the frequency of the $UB$ genotype at equilibrium. The column for “UPI max?” denotes whether the frequency of uniparental inheritance was maximized at equilibrium (section A.10). Additional parameters: $P_r = 0.5$ (for recombination).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>
Table A.20: Recombination and no mating types for U × U with biparental inheritance: \( n = 20 \) and \( \mu = 10^{-7} \). UPI is maximized at 0.5 when U × U have biparental inheritance (see main text for explanation). UPI (r) (frequency of uniparental inheritance at equilibrium, assuming recombination) is evenly split between the \( U_1B_2 \) and \( U_2B_1 \) genotypes at equilibrium, while the UPI (nmt) (frequency of uniparental inheritance at equilibrium, assuming no mating types) refers to the frequency of the UB genotype at equilibrium. The column for “UPI max?” denotes whether the frequency of uniparental inheritance was maximized at equilibrium (section A.10). Additional parameters: \( P_r = 0.5 \) (for recombination).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>Fitness ( c_h )</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-7} ) concave 0.01</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-7} ) linear 0.01</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-7} ) convex 0.01</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-7} ) concave 0.5</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-7} ) linear 0.5</td>
</tr>
<tr>
<td>20</td>
<td>( 10^{-7} ) convex 0.5</td>
</tr>
</tbody>
</table>

Table A.21: Recombination and no mating types for U × U with biparental inheritance: \( n = 100 \) and \( \mu = 10^{-4} \). UPI is maximized at 0.5 when U × U have biparental inheritance (see main text for explanation). UPI (r) (frequency of uniparental inheritance at equilibrium, assuming recombination) is evenly split between the \( U_1B_2 \) and \( U_2B_1 \) genotypes at equilibrium, while the UPI (nmt) (frequency of uniparental inheritance at equilibrium, assuming no mating types) refers to the frequency of the UB genotype at equilibrium. The column for “UPI max?” denotes whether the frequency of uniparental inheritance was maximized at equilibrium (section A.10). Additional parameters: \( P_r = 0.5 \) (for recombination).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>Fitness ( c_h )</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-4} ) concave 0.01</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-4} ) linear 0.01</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-4} ) convex 0.01</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-4} ) concave 0.5</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-4} ) linear 0.5</td>
</tr>
<tr>
<td>100</td>
<td>( 10^{-4} ) convex 0.5</td>
</tr>
</tbody>
</table>
Table A.22: Recombination and no mating types for $U \times U$ with mixed uniparental/biparental inheritance: $n = 20$ and $\mu = 10^{-4}$. The frequency of uniparental inheritance assuming recombination (UPI freq (r)) is given by $P(U_1B_2) + P(U_2B_1) + P(U_1U_2)(1 - P_b)$ (at equilibrium), while the frequency of uniparental inheritance assuming no mating types (UPI freq (nmt)) is given by $P(UB) + P(UU)(1 - P_b)$ (at equilibrium). Additional parameters: $P_r = 0.5$ (for recombination). The column for “UPI max?” denotes whether the frequency of uniparental inheritance was maximized at equilibrium (section A.10).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>
### Table A.23: Recombination and no mating types for $U \times U$ with mixed uniparental/biparental inheritance

The frequency of uniparental inheritance assuming recombination ($\text{UPI freq (r)}$) is given by $P(U_1B_2) + P(U_2B_1) + P(U_1U_2)(1 - P_b)$ (at equilibrium), while the frequency of uniparental inheritance assuming no mating types ($\text{UPI freq (nmt)}$) is given by $P(UB) + P(UU)(1 - P_b)$ (at equilibrium). Additional parameters: $P_r = 0.5$ (for recombination). The column for $\text{UPI (max)}$ denotes whether the frequency of uniparental inheritance was maximized at equilibrium (section A.10).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>100</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>
Table A.24: Modeling Saccharomyces: vegetative segregation (mitosis) occurs before selection can act. Generations means the number of generations to reach equilibrium. UPI freq is the frequency of uniparental inheritance at equilibrium. Divisions is the number of mitotic divisions. In rows 7, 8 and 10, in which there are few mitochondria, multiple mitotic divisions, and selection against heteroplasmy after mitosis, U\textsubscript{1} has no selective advantage and does not spread beyond its introductory frequency (when U\textsubscript{1} is introduced at a frequency of 0.01, the frequency of UPI is 0.02). Under these conditions, a mutation for uniparental inheritance could only spread via genetic drift; thus, biparental inheritance would be expected to remain stable if it were the ancestral condition.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>µ</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-4}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>8</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>8</td>
<td>10\textsuperscript{-7}</td>
</tr>
</tbody>
</table>

* The simulation in row 5 was stopped after 2 billion generations (before reaching equilibrium); while the spread of UPI was slowed in this simulation, it was not stopped.

Table A.25: Modeling Saccharomyces: selection acts midway through vegetative segregation (mitosis). In this case, we apply selection after cells have gone through half of their mitotic divisions. After selection, we apply the second half of the mitotic divisions (e.g. in row one: 10 divisions, selection, 10 divisions).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>µ</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-4}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
<tr>
<td>4</td>
<td>10\textsuperscript{-7}</td>
</tr>
</tbody>
</table>
Table A.26: Modeling Physarum: $U_1 \times B_2$ matings have mixed uniparental/biparental inheritance. Generations means the number of generations to reach equilibrium. UPI freq is the frequency of uniparental inheritance at equilibrium, and it is given by $P(U_1B_2)(1 - P_b)$ (at equilibrium).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
</tbody>
</table>
Table A.27: Modeling Didymium and Chlamydomonas: $U_1 \times U_2$ matings have a mixture of uniparental (from either parent) and biparental inheritance. We generated random parameter values for $P_b$, $P_{U_1}$, and $P_{U_2}$ using the “twister” MATLAB® rng. The rng values were normalized so that they sum to 1 because $P_b + P_{U_1} + P_{U_2} = 1$. $UPI(U_1)$ is given by $P_{U_1}(U_1 U_2) + U_1 B_2$, $UPI(U_2)$ is given by $P_{U_2}(U_1 U_2) + U_2 B_1$ and $BPI$ is given by $P_b(U_1 U_2) + B_1 B_2$.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$\mu$</td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A.2. Supplementary tables

Table A.28: Mutation variables describing the transition between a pre- and post-mutation cell in the model with 3 mitochondrial types.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>the number (out of $i$) of type $I$ mitochondria that mutate</td>
</tr>
<tr>
<td>$b$</td>
<td>the number (out of $j$) of type $J$ mitochondria that mutate</td>
</tr>
<tr>
<td>$c$</td>
<td>the number (out of $k$) of type $K$ mitochondria that mutate</td>
</tr>
<tr>
<td>$x_i$</td>
<td>the number (out of $a$) of mutations from type $I$ mitochondria to type $J$ mitochondria</td>
</tr>
<tr>
<td>$x_j$</td>
<td>the number (out of $b$) of mutations from type $J$ mitochondria to type $I$ mitochondria</td>
</tr>
<tr>
<td>$x_k$</td>
<td>the number (out of $c$) of mutations from type $K$ mitochondria to type $I$ mitochondria</td>
</tr>
<tr>
<td>$a - x_i$</td>
<td>the number (out of $a$) of mutations from type $I$ mitochondria to type $K$ mitochondria</td>
</tr>
<tr>
<td>$b - x_j$</td>
<td>the number (out of $b$) of mutations from type $J$ mitochondria to type $K$ mitochondria</td>
</tr>
<tr>
<td>$c - x_k$</td>
<td>the number (out of $c$) of mutations from type $K$ mitochondria to type $J$ mitochondria</td>
</tr>
</tbody>
</table>

Table A.29: Fitness function parameters in the neutral scenario. Parameters: $n = 20$.

<table>
<thead>
<tr>
<th></th>
<th>( t )</th>
<th>( v )</th>
<th>( y )</th>
<th>( z )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concave ($c_h = 0.1$)</td>
<td>0.25</td>
<td>0.68</td>
<td>2.5</td>
<td>0.9</td>
</tr>
<tr>
<td>Convex ($c_h = 0.1$)</td>
<td>0.05</td>
<td>1.5</td>
<td>1</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Table A.30: Values of $\phi$ used to generate the selection coefficients for the deleterious scenario (column 1) and values of $\zeta$ used produce the selection coefficients for the advantageous scenario (column 2) in the model with 3 mitochondrial types.

<table>
<thead>
<tr>
<th>$s_d = s_a$</th>
<th>$\phi$</th>
<th>$\zeta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-2}$</td>
<td>0.2</td>
<td>0.2089</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>0.0628</td>
<td>0.063</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>0.0198</td>
<td>0.0199</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>0.00628</td>
<td>0.00628</td>
</tr>
</tbody>
</table>
A.3 Supplementary text: detailed model dynamics

Once $B_1$ and $B_2$ gametes have reached mutation-selection equilibrium, part of the population is heteroplasmic (mutation-selection equilibrium is generation 0 in Figure 2.2 and Figure 2.3). When a mutation from $B_1$ to $U_1$ occurs in a gamete homoplasmic for the wild type haplotype, the proportion of $B_1$ and $B_2$ gametes with any level of heteroplasmy initially decreases (generations 0–100 in Figure 2.2C and Figure 2.3C–D). The influx of $U_1$ gametes homoplasmic for the wild type haplotype converts some heteroplasmic $B_2$ gametes into homoplasmic $B_2$ gametes. In turn, this drives down the proportion of heteroplasmy in $B_1$ gametes via $B_1 \times B_2$ matings. (When the mutation rate is smaller, this initial drop in heteroplasmy is less noticeable (Figure A.2 and Figure A.3).)

After about 100 generations, $U_1$ gametes homoplasmic for mutant mitochondria begin to increase in frequency (Figure 2.3B). As described earlier, this leads to matings between $B_2$ gametes homoplasmic for mutant mitochondria and $B_1$ gametes carrying the wild type haplotype, which result in heteroplasmic $B_1B_2$ cells (Figure 2.2C and Figure 2.3C–D; note that most $B_1$ gametes are homoplasmic for the wild type haplotype or only carry a few mutant mitochondria at this stage). This results in an increase in the proportion of heteroplasmic $B_1$ and $B_2$ gametes and $B_1B_2$ cells (generations 100–1350 in Figure 2.2C and Figure 2.3C–D). Selection against heteroplasmy thus decreases the relative fitness of $B_1 (\bar{\omega}_{B_1})$ and $B_2 (\bar{\omega}_{B_2})$ gametes and $B_1B_2 (\bar{\omega}_{B_1B_2})$ cells (Figure 2.2A and Figure 2.3A). From generations 1350–1820, the proportion of heteroplasmic $B_1$ and $B_2$ gametes and $B_1B_2$ cells decreases (Figure 2.2C and Figure 2.3C–D). Despite this, $\bar{\omega}_{B_1}$ and $\bar{\omega}_{B_1B_2}$ continue to decrease ($\bar{\omega}_{B_2}$, however, starts to converge with $\bar{\omega}_{U_1}$). While the proportion of heteroplasmic $B_1$ and $B_2$ gametes and $B_1B_2$ cells decreases during this period, the level of heteroplasmy within heteroplasmic gametes and cells increases (Figure 2.2C–E and Figure 2.3C–F). The increased levels of heteroplasmy outweigh the reduced proportion of heteroplasmic cells, and the net effect is increased selection against heteroplasmic $B_1B_2$ cells (Figure 2.2A and Figure 2.3A).

From generations 1350–1820, $U_1$ rapidly spreads through the population, increasing from 0.077 to 0.474. During this period, $U_1 \times B_2$ matings become more frequent, increasing the proportion of homoplasmic $B_2$ gametes. In turn, this increases the proportion of homoplasmic $B_1B_2$ cells and $B_1$ gametes through $B_1 \times B_2$ matings.
(Figure 2.2C–E and Figure 2.3C–F). More $B_2$ gametes are now homoplasmic for mutant mitochondria (Figure 2.3D; note that these $B_2$ gametes begin to appear around generation 1400 in Figure 2.3D). $B_1 \times B_2$ matings involving $B_2$ gametes homoplasmic for mutant mitochondria become more common, leading to $B_1B_2$ cells with high levels of heteroplasmy (compare Figure 2.2D with Figure 2.2E). Increased levels of heteroplasmy within $B_1B_2$ cells drives down $\bar{\omega}_{B_1B_2}$ and $\bar{\omega}_{B_1}$ (Figure 2.2A and Figure 2.3A).

As the frequency of $B_1$ decreases, $\bar{\omega}_{B_2}$ becomes increasingly determined by $U_1 \times B_2$ matings and $\bar{\omega}_{B_2}$ converges to $\bar{\omega}_{U_1}$ around generation 1900 (Figure 2.3A). During the remainder of the simulation, $\bar{\omega}_{B_1B_2}$ and $\bar{\omega}_{B_1}$ decrease further as $U_1$ replaces $B_1$.

Since there are few cells homoplasmic for mutant mitochondria at the beginning of the simulation, the relative advantage of $U_1$ over $B_1$ is low when the frequency of $U_1$ is low (e.g. $\bar{\omega}_{U_1} = 0.99984$ and $\bar{\omega}_{B_1} = 0.99881$ at generation 50 in Figure 2.3A, giving a relative advantage for $U_1$ of 0.001). As the frequency of $U_1$ gametes with mutant mitochondria increases, so too does the relative advantage of $U_1$ (e.g. $\bar{\omega}_{U_1} = 0.99984$ and $\bar{\omega}_{B_1} = 0.98476$ at generation 2500 in Figure 2.3A, giving a relative advantage for $U_1$ of 0.015). For more details about the change in gamete and cell type distributions as $U_1$ spreads, see S1–S2 Videos.
A.4 Supplementary text: no mating types

In the absence of mating types, there are two gametes (B and U) and three genotypes (UU, UB, and UU). As in the recombination case, when U × U matings lead to uniparental inheritance, the UU genotype always spreads until it is fixed in the population, leading to complete uniparental inheritance (Table A.16–Table A.18). When U × U matings lead to biparental inheritance or a mixture of uniparental inheritance and biparental inheritance, uniparental inheritance does not become fixed (again, as in the recombination case) (Table A.19–Table A.23). The only difference between the no mating type and recombination scenarios is that the UB genotype (no mating types) has the same frequency as the sum of the U_1B_2 and U_2B_1 genotypes (recombination) at equilibrium (provided that P_r is sufficiently large) (Figure 2.4A,F and Table A.16–Table A.23). Thus, the no mating type case can be inferred from the recombination case in the main text by setting P(UB) = P(U_1B_2) + P(U_2B_1).
A.5 Model description

Our model tracks the distribution of cell types through each stage of the life cycle across multiple generations. The redistribution of cell types is based on probability theory, but the model itself is deterministic. We assume that the population is effectively infinite and unaffected by genetic drift, as is regularly assumed in models such as ours [1–4]. Consequently, the probability that a cell takes a particular state equates to the proportion of that cell type in the population. We take a similar approach to previous models [1, 2], but our model differs slightly in our treatment of mutation. Hastings does not include mutation [2], while Hadjivasiliou and colleagues treat mutation as a one-way process from wild-type to mutant mitochondria in the conflict and mutation clearance models [1]. When examining the mitochondrial-nuclear coadaptation model, however, Hadjivasiliou and colleagues allow mutation to proceed both ways as we have done here [1]. In our model, mutation is designed to capture the ability of a mitochondrial type to mutate from its current state to other haplotypes (one type in our main model and two types in our supplementary model, but an extremely large number of haplotypes in reality).

Diploid cell types are described by the vector \( M^{t,\tau_\alpha} = (i, G) \), where \( i \) corresponds to the number of mutant mitochondria and takes values in \( \{0, 1...n\} \), \( t \) indicates the generation, and \( \tau_\alpha \) indicates the stage of the life cycle. If we know the number of mutant mitochondria \( i \), the number of wild type mitochondria (which we denote \( j \)) is fixed as \( j = n - i \). \( G \) indicates the nuclear genotype and takes values in \( \{U_1B_2, B_1B_2\} \). Gametes are described by the vector \( M^{t,\tau_\alpha} = (p, g) \), where \( p \) is the number of mutant mitochondria and takes values in \( \{0, 1...n/2\} \) and \( g \) represents the nuclear allele and takes values in \( \{U_1, B_1, B_2\} \). The probability of obtaining a particular diploid cell type is written as \( P(M^{t,\tau_\alpha} = (i, G)) \), and the probability of obtaining a particular gamete is written as \( P(M^{t,\tau_\alpha} = (p, g)) \). These probabilities can also be thought of as the proportion of the population with that particular cell or gamete type.

There are \( n + 1 \) total mitochondrial states for diploid cells and \( n/2 + 1 \) possible mitochondrial states for haploid cells. For the case in which mating type and inheritance loci are linked, the total number of diploid cell types is \( 2(n + 1) \) while the total number of haploid cell types is \( 3(n/2 + 1) \). We obtained numerical solutions to our model via scripts that we developed in MATLAB® (version 2013b).
A.5. Model description

A.5.1 Initialization

The starting population is evenly split between $B_1$ and $B_2$ gametes, and all gametes contain type wild type mitochondria (i.e. $P(M^{0,\tau_1} = (0, B_1)) = 0.5$, $P(M^{0,\tau_1} = (0, B_2)) = 0.5$, and $P(M^{0,\tau_1} = (p, g)) = 0$, $\forall p > 0$ and $g = U_1$). We first allow this population to reach equilibrium, which we define as the point at which the proportion of cell types change by less than $10^{-12}$ (except when the probability that a mitochondrion mutates into another mitochondrion is $\mu = 10^{-10}$, in which case we define equilibrium to be a change of less than $10^{-13}$). We then introduce $U_1$ gametes that are homoplasmic for wild type mitochondria by setting $P(M^{g_1,\tau_1} = (0, U_1)) = 0.01$, where $g_1$ is the number of generations taken to reach the first equilibrium. To maintain the total proportion of the population at 1, we remove the corresponding proportion of cells from the $B_1$ population. In two instances, we alter the way in which $U_1$ is introduced. In Figure A.4, we introduce $U_1$ into the most heteroplasmic gamete with a frequency greater than 0.01, and in Figure A.5 we vary the introductory frequency of $U_1$. Our life cycle is very similar to the life cycle used by Hadjivasiliou and colleagues [1], which examined the genomic conflict, mutational clearance, and mitochondrial-nuclear coadaptation hypotheses.

A.5.2 Random mating

Gametes with $n/2$ mitochondria randomly mate with the opposite mating type to produce diploid cells containing $n$ mitochondria. In effect, this is random mating in which all matings between the same mating type (i.e. $U_1U_1$, $B_1B_1$, $B_2B_2$, and $U_1B_1$) are lethal, and the only viable genotypes are $U_1B_2$ and $B_1B_2$.

A.5.2.1 Biparental mating

Consider a biparental mating of a gamete in state $M^{k,\tau_1} = (p, B_1)$, where $\tau_1$ is the gamete stage of the life cycle. For this gamete to produce a diploid cell with type $M^{k,\tau_2} = (i, B_1B_2)$, where $\tau_2$ is the diploid stage of the life cycle that precedes mutation, it must mate with a gamete of type $M^{i,\tau_1} = (i - p, B_2)$. The probability of this mating is $2 [P(M^{k,\tau_1} = (p, B_1)) P(M^{i,\tau_1} = (i - p, B_2))]$, where the factor of 2 accounts for the two ways in which we can choose $B_1$ and $B_2$ ($B_1$ then $B_2$ or $B_2$ then $B_1$). We restrict the values of $p$ and $i - p$ to biologically valid combinations. First, $0 \leq p \leq n/2$, as the $B_1$ gamete cannot carry negative numbers of mutant mitochondria, nor can it contain more mutant mitochondria than the total number of mitochondria in the gamete. Likewise, $0 \leq i - p \leq n/2$ for the $B_2$ gamete, which,
when rearranged gives \( i - n/2 \leq p \leq i \). Valid values for \( p \) lie in the range of intersection of these two inequalities, giving \( \max(0, i - n/2) \leq p \leq \min(n/2, i) \). We can thus obtain the probability of forming any given diploid cell type after random mating with the sum,

\[
P \left( M^{t,\tau_2} = (i, B_1 B_2) \right) = 2 \sum_{p=\max(0,i-n/2)}^{\min(n/2,i)} P(M^{t,\tau_1} = (p, B_1)) P \left( M^{t,\tau_1} = (i - p, B_2) \right).
\]

### A.5.2.2 Uniparental mating

Because uniparental matings between \( U_1 \) and \( B_2 \) gametes contain mitochondria from \( U_1 \) alone, \( U_1 B_2 \) cells initially have \( n/2 \) mitochondria. To restore the total complement of \( n \) mitochondria, we sample \( n/2 \) mitochondria with replacement from the \( n/2 \) mitochondria in the \( U_1 B_2 \) cell, and add the \( n/2 \) sampled mitochondria to the original set of mitochondria to form a cell with \( n \) mitochondria.

For a gamete with identity \( M^{t,\tau_1} = (p, U_1) \) to produce a diploid cell with identity \( M^{t,\tau_2} = (i, U_1 B_2) \), it must sample \( n/2 \) mitochondria containing \( i - p \) mutant mitochondria and \( n/2 - (i - p) \) wild type mitochondria. The mitochondrial state of the \( B_2 \) gamete is irrelevant because its mitochondria are discarded and we will refer to this cell as \( M^{t,\tau_1} = (r, B_2) \).

Sampling of the \( n/2 \) mitochondria follows a binomial distribution, which we denote \( T(i - p; n/2, 2p/n) \), where \( i - p \) refers to the number of mutant mitochondria that need to be sampled, \( n/2 \) refers to the number of mitochondria being sampled, and \( 2p/n \) is the probability of drawing a single mutant mitochondrion from a \( U_1 B_2 \) cell with \( p \) (out of \( n/2 \)) mutant mitochondria (where \( 2p/n \) is obtained by rearranging \( p/(n/2) \)). The probability of sampling \( i - p \) mutant mitochondria (and \( n/2 - (i - p) \) wild type mitochondria) is given by

\[
T\left(i - p; \frac{n}{2}, \frac{2p}{n}\right) = \binom{n/2}{i - p} \left(\frac{2p}{n}\right)^{i-p} \left(1 - \frac{2p}{n}\right)^{n/2-i-p}.
\]

(A.1)

The restrictions on \( p \) and \( i - p \) are the same as those in biparental mating. Since \( U_1 \) will form the same initial \( U_1 B_2 \) cell regardless of the \( B_2 \) gamete with which it mates, the probability of producing each type of \( U_1 \) gamete is multiplied by the probability of selecting any \( B_2 \) gamete. The probability of forming a given \( U_1 B_2 \) cell after random mating is determined by
\[ P \left( M^{t,\tau_2} = (i, U_1 B_2) \right) \]
\[ = 2 \left[ \sum_{p=\max(0,i-n/2)}^{\min(n/2,i)} P \left( M^{t,\tau_1} = (p, U_1) \right) T \left( i - p, \frac{n}{2}, \frac{2p}{n} \right) \sum_{r=0}^{n/2} P \left( M^{t,\tau_1} = (r, B_2) \right) \right] . \]

(A.2)

### A.5.3 Mutation

We denote the post-mutation states of cells as \( M^{t,\tau_3} = (i, G) \), where \( \tau_3 \) indicates the post-mutation life cycle stage. If we define the number of wild type mitochondria that mutate to mutant mitochondria to be \( a \) and the number of mutant mitochondria that mutate to wild type mitochondria as \( b \), a post-mutation cell in state \( M^{t,\tau_3} = (i, G) \) must be derived from a pre-mutation cell in state \( M^{t,\tau_2} = (i - a + b, G) \) (because the pre-mutation cell gains \( a \) mutant mitochondria and loses \( b \) mutant mitochondria to form the post-mutation cell). Similarly, if the post-mutation cell has \( j \) wild type mitochondria, then the pre-mutation cell must have \( j + a - b \) wild type mitochondria, where \( j = n - i \).

First, we must work out the probability that a cell mutates \( a \) of its wild type mitochondria to mutant mitochondria. We define \( Y(a; n - i + a - b, \mu) \) as the probability that a pre-mutation cell has \( a \) mutations in its \( n - i + a - b \) wild type mitochondria, given that each mitochondrion mutates with probability \( \mu \). The accumulation of mutations is binomially distributed such that

\[
Y(a; n - i + a - b, \mu) = \binom{n - i + a - b}{a} \mu^a (1 - \mu)^{n - i - b}.
\]

Similarly we define \( Y(b; i - a + b, \mu_b) \) to be the probability that a pre-mutation cell acquires \( b \) mutations in its \( i - a + b \) mutant mitochondria, given that each mitochondrion mutates with probability \( \mu_b \). This probability is given by

\[
Y(b; i - a + b, \mu_b) = \binom{i - a + b}{a} \mu_b^b (1 - \mu_b)^{i - a}.
\]

For any combination of values for \( a, b, \) and \( i \), multiplying \( Y(a; n - i + a - b, \mu) \) by \( Y(b; i - a + b, \mu_b) \) gives the probability of a particular transition from a pre-mutation cell with identity \( M^{t,\tau_2} = (i - a + b, G) \) to a post-mutation cell with identity \( M^{t,\tau_3} = (i, G) \). To get the overall probability that such a transition occurs, we multiply the probability of the transition by the proportion of pre-mutation cells in the population. To produce the post-mutation population, we sum all possible transitions between
pre-mutation and post-mutation cells. All valid transitions must satisfy $0 \leq a \leq i$ (because the post-mutation cell cannot receive more than $i$ mutant mitochondria) and $0 \leq b \leq n - i$ (because the post-mutation cell cannot receive more than $n - i$ wild type mitochondria). Thus, we can determine the post-mutation population by

$$P (M^{t,\tau_3} = (i, G)) = \sum_{a=0}^{i} \sum_{b=0}^{n-i} Y (a; n - i + a - b, \mu) Y (b; i - a + b, \mu_b) P (M^{t,\tau_2} = (i - a + b, G)).$$

In the neutral scenario, $\mu = \mu_b$ (i.e. the rate of mutation from wild type to mutant is equal to the rate of mutation from mutant to wild type).

### A.5.4 Selection

The relative fitness of a cell, $\omega(i)$, is a measure of how likely a cell type is to survive and reproduce, and we assume that cells carrying multiple mitochondrial types have lower fitness. For the first fitness function, the relative fitness of a cell with $i$ mutant mitochondria is determined according to the following piecewise concave function

$$\omega(i) = \begin{cases} 1 - c_h \left( \frac{i}{n/2} \right)^2 & \text{for } 0 \leq i < n/2, \\ 1 - c_h \left( \frac{n-i}{n/2} \right)^2 & \text{for } n/2 \leq i \leq n. \end{cases}$$

(A.3)

This function is valid for even values of $n$ and $0 \leq c_h \leq 1$, where $c_h$ is the cost of heteroplasmy. In this function, a cell containing $n/2$ mutant and $n/2$ wild type mitochondria has minimum relative fitness. The post-selection population of each cell type is then given by $P (M^{t,\tau_4} = (i, G)) = \omega(i) P (M^{t,\tau_3} = (i, G))$. We also make use of two alternative fitness functions. The first of these is the piecewise linear function

$$\omega(i) = \begin{cases} 1 - c_h \left( \frac{i}{n/2} \right) & \text{for } 0 \leq i < n/2, \\ 1 - c_h \left( \frac{n-i}{n/2} \right) & \text{for } n/2 \leq i \leq n. \end{cases}$$

The third fitness function is the piecewise convex function

$$\omega(i) = \begin{cases} 1 - c_h \sqrt{\frac{i}{n/2}} & \text{for } 0 \leq i < n/2, \\ 1 - c_h \sqrt{\frac{n-i}{n/2}} & \text{for } n/2 \leq i \leq n. \end{cases}$$

(A.4)

We normalize the post-selection population by
A.5. Model description

\[ P(M^{t,\tau_5} = (i,G)) = \frac{P(M^{t,\tau_4} = (i,G))}{\sigma}, \]

where

\[ \sigma = \sum_{i=0}^{n} \left[ P(M^{t,\tau_4} = (i,U_1B_2)) + P(M^{t,\tau_4} = (i,B_1B_2)) \right], \]

so that the sum of the proportions of the population equals 1.

A.5.5 Meiosis

The cell must first duplicate its chromosomes and double its mitochondrial complement (from \( n \) to \( 2n \)). This cell with \( 2n \) mitochondria then produces gametes with \( n/2 \) mitochondria. Meiosis occurs in two steps. First, we sample \( n \) mitochondria with replacement from a cell containing \( n \) mitochondria and add the set of sampled mitochondria to the original set of mitochondria to form a cell containing \( 2n \) mitochondria (this is the same process that occurs in uniparental mating only with \( n \) mitochondria rather than \( n/2 \) mitochondria). We let \( M^{t,\tau_6} = (l,2G) \) represent the cell with doubled mitochondria and nuclear genotype, where \( l \) takes values in \( \{0,1,\ldots,2n\} \) and \( 2G \) takes values in \( \{U_1U_1B_2B_2,B_1B_1B_2B_2\} \).

For a cell to contain \( l \) mutant mitochondria after duplication of its mitochondria, it must sample \( l - i \) mutant mitochondria. We denote the probability of sampling \( l - i \) mutant mitochondria from \( M^{t,\tau_5} = (i,G) \) as \( F(l - i; n, i/n) \). Sampling follows a binomial distribution such that

\[ F(l - i; n, i/n) = \binom{n}{l-i} \left( \frac{i}{n} \right)^{l-i} \left( 1 - \frac{i}{n} \right)^{n-l+i}. \]

We obtain \( M^{t,\tau_6} = (l,2G) \) by

\[ P(M^{t,\tau_6} = (l,2G)) = \sum_{i=\max(0,l-n)}^{\min(l,n)} F(l - i; n, i/n) P(M^{t,\tau_5} = (i,G)). \]

During the second step of meiosis, the cells with \( 2n \) mitochondria produce gametes with \( n/2 \) mitochondria. Biologically, this occurs in two steps. In meiosis 1, the homologous chromosomes are pulled apart to produce two haploid cells that contain two identical nuclear alleles (sister chromatids) and \( n \) mitochondria. In meiosis 2, the two cells divide to produce four gametes, each with a single nuclear allele and \( n/2 \) mitochondria. Since mitochondria segregate independently of nuclear alleles during cell partitioning, we model this as a single step. We define \( S(p;2n,l,n/2) \) to be
A.5. Model description

the probability of obtaining \( p \) mutant mitochondria in \( n/2 \) draws from a cell in state \( (M^{t,\tau_6} = (l, 2G)) \). Here, sampling is without replacement and follows a hypergeometric distribution, giving

\[
S\left(p; 2n, l, \frac{n}{2}\right) = \frac{\binom{l}{p} \binom{2n-l}{n/2-p}}{\binom{2n}{n/2}}.
\] (A.5)

Gametes produced by meiosis are represented by \( M^{t+1,\tau_1} = (p, g) \). We determine the probability of obtaining a particular gamete using

\[
P(M^{t+1,\tau_1} = (p, U_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, U_1B_1B_2B_2)\right) \right],
\]

\[
P(M^{t+1,\tau_1} = (p, B_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, B_1B_1B_2B_2)\right) \right],
\]

and

\[
P(M^{t+1,\tau_1} = (p, B_2)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, U_1U_1B_2B_2)\right) \right],
\]

\[
+ \frac{1}{2} \left[ \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, B_1B_1B_2B_2)\right) \right].
\]

Factors of 1/2 in the above three equations take into account that half of the gametes produced from parent cells with nuclear genotype \( U_1B_2 \) will carry the \( U_1 \) allele and the other half will carry the \( B_2 \) allele (with similar logic applied for gametes produced from parent cells with nuclear genotype \( B_1B_2 \)). Meiosis completes a single generation of the life cycle.

A.5.6 Relative fitness of cells

The relative fitness of \( U_1B_2 \) cells is given by

\[
\bar{\omega}_{U_1B_2} = \frac{\sum_{i=0}^{n} P(M^{t,\tau_3} = (i, U_1B_2)) \omega(i)}{\sum_{i=0}^{n} P(M^{t,\tau_3} = (i, U_1B_2))},
\]
while the relative fitness of $B_1B_2$ cells is

$$
\bar{\omega}_{B_1B_2} = \frac{\sum_{i=0}^{n} P(M^{i,\tau_3} = (i, B_1B_2)) \omega(i)}{\sum_{i=0}^{n} P(M^{i,\tau_3} = (i, B_1B_2))}.
$$

### A.5.7 Relative fitness of gametes

Although gametes are not subject to selection in our model, and thus do not technically have fitness values, it is informative to track the relative fitness of gametes throughout the simulation. We define a gamete’s relative fitness as the fitness that a diploid cell would have if it had the same mitochondrial composition as the gamete. Since gametes contain $n/2$ mitochondria, they will have minimum fitness when they carry $n/4$ wild type and $n/4$ mutant mitochondria. To rescale the fitness function, we substitute $n/2$ for $n$ in the diploid cell fitness functions. For example, Equation A.3 becomes

$$
\omega(i)_g = \begin{cases} 
1 - c_h \left( \frac{i}{n/4} \right)^2 & \text{for } 0 \leq i < n/4, \\
1 - c_h \left( \frac{n/2-i}{n/4} \right)^2 & \text{for } n/4 \leq i \leq n/2.
\end{cases}
$$

Once the fitness function is scaled to gametes, we can determine the relative fitness of the $U_1$, $B_1$, and $B_2$ alleles by

$$
\bar{\omega}_{U_1} = \frac{\sum_{i=0}^{n/2} P(M^{i,\tau_1} = (i, U_1)) \omega_{U_1}(i)}{\sum_{i=0}^{n/2} P(M^{i,\tau_1} = (i, U_1))},
$$

$$
\bar{\omega}_{B_1} = \frac{\sum_{i=0}^{n/2} P(M^{i,\tau_1} = (i, B_1)) \omega_{B_1}(i)}{\sum_{i=0}^{n/2} P(M^{i,\tau_1} = (i, B_1))},
$$

and

$$
\bar{\omega}_{B_2} = \frac{\sum_{i=0}^{n/2} P(M^{i,\tau_1} = (i, B_2)) \omega_{B_2}(i)}{\sum_{i=0}^{n/2} P(M^{i,\tau_1} = (i, B_2))}.
$$
A.6 Non-neutral models

A.6.1 Deleterious model

Let $s_d$ be the selection coefficient of the deleterious mutation (where the fitness of a cell that is homoplasmic for mutant mitochondria is $1 - s_d$). Assuming that the cost imposed by mutant mitochondria can be modeled by a concave function, the fitness of a cell with $i$ mutant mitochondria is given by

$$\omega_d(i) = \omega(i) \left(1 - s_d \left(\frac{i}{n}\right)^2\right),$$

where $\omega(i)$ is determined by either Equation A.3 or Equation A.4.

A.6.2 Advantageous model

Again, we generate $\omega(i)$ using Equation A.3 or Equation A.4. We let $s_a$ be the selection coefficient of the deleterious mutation (where the fitness of a cell that is homoplasmic for wild type mitochondria is $1 - s_a$). Since there are no data as to how fitness increases as advantageous mitochondria accumulate, we model the increase in fitness using a concave function and a convex function. If we assume that advantageous mutations convey fitness benefits consistent with a concave function, the fitness of a cell with $i$ mutant mitochondria is given by

$$\omega_a(i) = \omega(i) \left(1 - s_a \left(\frac{n - i}{n}\right)^2\right).$$

If we assume that fitness accumulates as a convex function of the number of mutant mitochondria, then the fitness of a cell with mutant mitochondria is given by

$$\omega_a(i) = \omega(i) \left(1 - s_a \sqrt{\frac{n - i}{n}}\right).$$
A.7 Mating types with recombination

In this scenario, there are four gametes ($U_1$, $U_2$, $B_1$, and $B_2$) and four genotypes ($U_1U_2$, $U_1B_2$, $U_2B_1$, and $B_1B_2$).

A.7.1 Random mating

A.7.1.1 Biparental mating ($B_1B_2$ cells)

The probability of producing a $B_1B_2$ cell type after random mating is given by

$$P(M_t,\tau_2 = (i,B_1B_2)) = 2 \left[ \sum_{p=\max(0,i-n/2)}^{\min(n/2,i)} P(M_t,\tau_1 = (p,B_1))P(M_t,\tau_1 = (i-p,B_2)) \right].$$

A.7.1.2 Biparental mating ($U_1U_2$ cells)

The probability of producing a $U_1U_2$ cell, when we assume that $U_1 \times U_2$ matings are biparental, is given by

$$P(M_t,\tau_2 = (i,U_1U_2)) = 2 \left[ \sum_{p=\max(0,i-n/2)}^{\min(n/2,i)} P(M_t,\tau_1 = (p,U_1))P(M_t,\tau_1 = (i-p,U_2)) \right].$$

A.7.1.3 Uniparental mating ($U_1B_2$ and $U_2B_1$ cells)

The probability of forming a $U_2B_1$ cell is

$$P(M_t,\tau_2 = (i,U_2B_1)) = 2 \left[ \sum_{p=\max(0,i-n/2)}^{\min(n/2,i)} P(M_t,\tau_1 = (p,U_2)) T \left( i - p; \frac{n}{2}, \frac{2p}{n} \right) \sum_{r=0}^{n/2} P(M_t,\tau_1 = (r,B_1)) \right].$$

For $U_1B_2$, see Equation A.2.

A.7.1.4 Uniparental mating ($U_1U_2$ cells)

The probability of producing a $U_1U_2$ cell, when we assume that $U_1U_2$ matings are uniparental, is
A.7. Mating types with recombination

\[ P(M^{t,\tau_2} = (i, U_1 U_2)) \]
\[ = 2 \left[ \sum_{p=\max(0, i-\frac{n}{2})}^{\min(n/2, i)} P(M^{t,\tau_1} = (p, U_1)) T \left( i - p, \frac{n}{2}, \frac{2p}{n} \right) \sum_{r=0}^{\frac{n}{2}} P\left(M^{t,\tau_1} = (r, U_2)\right) \right]. \]

(Note that we also ran simulations where \( U_2 \), rather than \( U_1 \), was the mitochondrial donor but this did not affect our results.) As before, mutation, selection and normalization are the same as the general model.

A.7.2 Meiosis

During meiosis 1, homologous chromosomes line up and may undergo recombination. The probability of recombination, \( P_r \), cannot exceed 0.5 because, at most, only two of the four chromatids can recombine. \( U_2 \) gametes are produced when \( U_1 B_2 \) cells undergo recombination, giving rise to a \( U_2 B_1 \) cell. \( U_2 B_1 \) cells may also undergo recombination to give \( U_1 B_2 \) cells. If we let \( P_r \) be the probability that the mating type and inheritance loci recombine, then the probability of producing a \( U_1 \) gamete is given by

\[ P\left(M^{t+1,\tau_1} = (p, U_1)\right) \]
\[ = \frac{1}{2} \left[ (1 - P_r) \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, U_1 U_1 B_2 B_2)\right) \right] \]
\[ + \frac{1}{2} P_r \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, U_2 U_2 B_1 B_1)\right) \]
\[ + \frac{1}{2} \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, U_1 U_1 U_2 U_2)\right) \]

where \( S\left(p; 2n, l, \frac{n}{2}\right) \) is given by Equation A.5. The probability of producing a \( U_2 \) gamete is
\[
P (M^{t+1, \tau_1} = (p, U_2)) \\
= \frac{1}{2} (1 - P_r) \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, U_2U_1B_1B_1)) \right] \\
+ \frac{1}{2} P_r \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, U_1U_1B_2B_2)) \right] \\
+ \frac{1}{2} \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, U_1U_1U_2U_2)) \right],
\]
the probability of producing a \(B_1\) gamete is

\[
P (M^{t+1, \tau_1} = (p, B_1)) \\
= \frac{1}{2} (1 - P_r) \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, U_2U_2B_1B_1)) \right] \\
+ \frac{1}{2} P_r \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, U_1U_1B_2B_2)) \right] \\
+ \frac{1}{2} \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, B_1B_1B_2B_2)) \right],
\]
and the probability of producing a \(B_2\) gamete is

\[
P (M^{t+1, \tau_1} = (p, B_2)) \\
= \frac{1}{2} (1 - P_r) \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, U_1U_1B_2B_2)) \right] \\
+ \frac{1}{2} P_r \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, U_2U_2B_1B_1)) \right] \\
+ \frac{1}{2} \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P (M^{t, \tau_6} = (l, B_1B_1B_2B_2)) \right].
\]
A.8 No mating types

In this version of the model, there are two possible alleles (U and B) and three possible genotypes (UU, UB, and BB).

A.8.1 Initialization

We introduce the U allele into homoplasmic wild type gametes at a proportion of 0.01 and remove 0.01 from the B gametes. All other details of initialization remain the same as the general model.

A.8.2 Random mating

A.8.2.1 Biparental mating (BB cells)

The probability of producing a BB cell type after random mating is given by

$$P\left(M^{i,\tau_2} = (i, BB)\right) = \sum_{p = \max(0, i-n/2)}^{\min(n/2, i)} P(M^{i,\tau_1} = (p, B))P\left(M^{i,\tau_1} = (i-p, B)\right).$$

A.8.2.2 Biparental mating (UU cells)

The probability of producing a UU cell, when we assume that U \times U matings are biparental, is given by

$$P\left(M^{i,\tau_2} = (i, UU)\right) = \sum_{p = \max(0, i-n/2)}^{\min(n/2, i)} P(M^{i,\tau_1} = (p, U))P\left(M^{i,\tau_1} = (i-p, U)\right).$$

A.8.2.3 Uniparental mating (UB cells)

The probability of forming a cell by random mating is determined by

$$P\left(M^{i,\tau_2} = (i, UB)\right)$$

$$= 2 \left[ \sum_{p = \max(0, i-n/2)}^{\min(n/2, i)} P\left(M^{i,\tau_1} = (p, U)\right) T\left(i - p, \frac{n}{2}, \frac{2p}{n}\right) \sum_{r = 0}^{n/2} P\left(M^{i,\tau_1} = (r, B)\right) \right],$$

where $T\left(i - p, \frac{n}{2}, \frac{2p}{n}\right)$ is given by Equation A.1.
A.8.2.4 Uniparental mating (UU cells)

The probability of producing a UU cell, when we assume that $U \times U$ matings are uniparental, is

\[
P \left( M^{t,\tau_2} = (i, UU) \right) = \sum_{p=\max(0,i-n/2)}^{\min(n/2,i)} P \left( M^{t,\tau_1} = (p, U) \right) T \left( i-p; \frac{n}{2}, \frac{2p}{n} \right) \sum_{r=0}^{n/2} P \left( M^{t,\tau_1} = (r, U) \right)
\]

The mutation, selection and normalization stages are the same as the general model (although there are now three genotypes instead of two).

A.8.3 Meiosis

The process by which cells in state $M^{t,\tau_5} = (i, G)$ become cells in state $M^{t,\tau_6} = (l, 2G)$ does not change. Thus, the probability of producing a U gamete is

\[
P \left( M^{t+1,\tau_1} = (p, U) \right) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P \left( M^{t,\tau_6} = (l, UUBB) \right) \right]
\]

\[+ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P \left( M^{t,\tau_6} = (l, UUUU) \right),
\]

while the probability of producing a B gamete is

\[
P \left( M^{t+1,\tau_1} = (p, B) \right) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P \left( M^{t,\tau_6} = (l, UUBB) \right) \right]
\]

\[+ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P \left( M^{t,\tau_6} = (l, BBBB) \right).
\]

$S \left( p; 2n, l, \frac{n}{2} \right)$ is given by Equation A.5.
A.9 Mitosis

In this scenario, we add mitosis to the life-cycle. Mitosis mimics the continual turnover of mitochondria that occurs within a cell. (Mitochondria form interconnected networks that continually undergo fission/fusion.) We examine how the model behaves when mitosis is inserted before selection and when it is inserted after selection. To model mitosis, we sample \( n \) mitochondria with replacement from a cell that contains \( i \) mutant (out of \( n \)) mitochondria. We denote the probability of sampling \( i' \) mutant mitochondria with replacement from a cell with \( i \) mutant mitochondria as \( X(i'; n, i/n) \). When mitosis occurs before selection, the post-mitosis population is determined by

\[
P(M^{t,\tau_4} = (i', G)) = \sum_{i=0}^{n} X(i; n, i/n) P(M^{t,\tau_3} = (i, G)).
\]

When mitosis occurs after selection, the post-mitosis population is given by

\[
P(M^{t,\tau_6} = (i', G)) = \sum_{i=0}^{n} X(i; n, i/n) P(M^{t,\tau_5} = (i, G)).
\]
A.10 Frequencies of \( U \) and \( B \) alleles that maximize the level of uniparental inheritance

Here we determine the frequencies of \( U \) and \( B \) that maximize the level of uniparental inheritance, assuming that \( U \times U \) matings have biparental inheritance with probability \( P_b \) (0 \( \leq \) \( P_b \) \( \leq \) 1) and uniparental inheritance with probability \( 1 - P_b \). Writing \( U \) as the total proportion of the population with the \( U \) allele and \( B \) as the total proportion of the population with the \( B \) allele, the pre-mating population of gametes satisfies \( U + B = 1 \), and the post-mating population satisfies

\[
P_bU^2 + (1 - P_b)U^2 + 2UB + B^2 = 1. \tag{A.6}
\]

We define the part of Equation A.6 that leads to uniparental inheritance as \( f = (1 - P_b)U^2 + 2UB \). Thus we can rearrange Equation A.6 to give

\[
f = 1 - P_bU^2 - B^2. \tag{A.7}
\]

We substitute \( U = 1 - B \) into Equation A.7 to give \( f(B) = 1 - P_b(1 - B)^2 - B^2 \), which upon rearrangement gives

\[
f(B) = (1 - P_b) + 2P_bB - (P_b + 1)B^2. \tag{A.8}
\]

Differentiating Equation A.8 with respect to \( B \) gives

\[
\frac{df}{dB} = 2P_b - 2(P_b + 1)B. \tag{A.9}
\]

Local optima of \( f(B) \) satisfy \( \frac{df}{dB} = 0 \). Therefore, \( 2P_b = 2(P_b + 1)B \), which leads to

\[
B = \frac{P_b}{P_b + 1}. \tag{A.10}
\]

The frequency of the \( U \) allele at equilibrium is then \( U = 1 - B \). Differentiating Equation A.9 with respect to \( B \) gives

\[
\frac{d^2f}{dB^2} = -2(P_b + 1),
\]

which is less than 0 for all \( P_b \in [0, 1] \). Hence the optima at \( B = \frac{P_b}{P_b + 1} \) is a local maximum. At the local maximum, \( f(B_{max}) = \frac{1}{P_b + 1} \). To determine if the local maximum also indicates the maximum value of \( f(B) \) we must also check the values of \( f(B) \) at the end points of the line \( U + B = 1 \). \( f(1) = 0 \) and \( f(0) = 1 - P_b \). For all \( P_b \in [0, 1] \),
f(0) ≤ f(B_{max}), and therefore the maximum frequency of uniparental inheritance occurs when \( B = \frac{P_b}{P_b + 1} \).

We checked the predictions of Equation A.10 against our simulation results when \( U \times U \) matings are possible. When fitness is linear or convex, Equation A.10 predicts the equilibrium state every time (i.e. the equilibrium state is such that the frequency of uniparental inheritance is at its maximum possible value). When fitness is concave, however, uniparental inheritance is only maximized under certain values of \( P_b \) (Table A.22 and Table A.23).
A.11 Model assuming three mitochondrial types

Diploid cell types are described by the vector $M_{t,\tau} = (i, j, G)$, where $i$ corresponds to the number of type I mitochondria and takes values in $\{0, 1 \ldots n\}$, $j$ represents the number of type J mitochondria and takes values in $\{0, 1 \ldots n - i\}$, $t$ indicates the generation, and $\tau$ indicates the stage of the life cycle. If $i$ and $j$ are specified, the number of type K mitochondria is fixed as $k = n - i - j$. $G$ indicates the nuclear genotype and takes values in $\{U_1 B_2, B_1 B_2\}$. Gametes are described by the vector $M_{t,\tau} = (p, q, g)$, where $p$ is the number of type I mitochondria and takes values in $\{0, 1 \ldots n/2\}$, and $q$ is the number of type J mitochondria and takes values in $\{0, 1 \ldots n/2 - p\}$. $g$ represents the nuclear allele and takes values in $\{U_1, B_1, B_2\}$. The probability of obtaining a particular diploid cell type is written as $P(M_{t,\tau} = (i, j, G))$, and the probability of obtaining a particular gamete is written as $P(M_{t,\tau} = (p, q, g))$. There are $\frac{(n+1)(n+2)}{2}$ total mitochondrial states for diploid cells and $\frac{(n/2+1)(n/2+2)}{2}$ possible mitochondrial states for haploid cells.

A.11.1 Initialization

The starting population is evenly split between $B_1$ and $B_2$ gametes, and all gametes contain type K mitochondria (i.e. $P(M^0,\tau_1 = (0, 0, B_1)) = 0.5$, $P(M^0,\tau_1 = (0, 0, B_2)) = 0.5$, and $P(M^0,\tau_1 = (p, q, B_1)) = 0, \forall p > 0, q > 0$ and $g = U_1$). After 100 generations, we identify the mitochondrial state of the $B_1$ gamete that makes up the greatest proportion of the population (denoted $M_{101,\tau_1} = (p^*, q^*, B_1)$). We reduce the proportion of this cell type by 0.01 (i.e. $P(M_{101,\tau_1} = (p^*, q^*, B_1)) = P(M_{101,\tau_1} = (p^*, q^*, B_1)) - 0.01$) and then we introduce $U_1$ gametes by setting $P(M_{101,\tau_1} = (p^*, q^*, U_1)) = 0.01$. Unless $\mu$ is very high, these conditions lead to $U_1$ being introduced where $p^* = q^* = 0$ (i.e. homoplasmic for type K).

A.11.2 Random mating

A.11.2.1 Biparental mating

Consider a biparental mating of a gamete in state $M_{t,\tau_1} = (p, q, B_1)$, where $\tau_1$ is the gamete stage of the life cycle. For this gamete to produce a diploid cell with type $M_{t,\tau_2} = (i, j, B_1 B_2)$, where $\tau_2$ is the diploid stage of the life cycle that precedes mutation, it must mate with a gamete of type $M_{t,\tau_1} = (i - p, j - q, B_2)$. The probability of this mating is $2P(M_{t,\tau_1} = (p, q, B_1))P(M_{t,\tau_1} = (i - p, j - q, B_2))$. 


However, not all combinations of $p$, $q$, $i-p$, and $j-q$ lead to valid matings; thus, we must restrict these values to biologically valid combinations.

First, $0 \leq p \leq n/2$, as the $B_1$ gamete cannot carry negative numbers of type $I$ mitochondria, nor can it contain more type $I$ mitochondria than the total number of mitochondria in the gamete. Likewise, $0 \leq i - p \leq n/2$ for the $B_2$ gamete, which rearranged gives $i - n/2 \leq p \leq i$. Valid values for $p$ lie in the range of intersection of these two inequalities, giving $\max(0, i - n/2) \leq p \leq \min(n/2, i)$.

The first restriction for $q$ is $0 \leq q \leq n/2 - p$ because the $B_1$ gamete already contains $p$ type $I$ mitochondria and cannot contain more than $n/2$ mitochondria. $q$ is also restricted by $0 \leq j - q \leq n/2 - (i - p)$ because the $B_2$ gamete cannot contain more than $n/2$ mitochondria and already contains $i - p$ type $I$ mitochondria, which gives $i + j - (n/2) - p \leq q \leq j$ when rearranged. The intersection of these inequalities gives $\max(0, i + j - (n/2) - p) \leq q \leq \min(n/2 - p, j)$. Thus,

$$P \left( M_{t,\tau_2} = (i, j, B_1B_2) \right)$$

$$= 2 \left[ \sum_{p=\max(0, i-n/2)}^{\min(n/2, i)} \sum_{q=\max(0, i+j-(n/2)-p)}^{\min(n/2-p, j)} P \left( M_{t,\tau_1} = (p, q, B_1) \right) P \left( M_{t,\tau_1} = (i - p, j - q, B_2) \right) \right].$$

A.11.2.2 Uniparental mating

Since uniparental matings between $U_1$ and $B_2$ gametes contain mitochondria from $U_1$ alone, $U_1B_2$ cells initially have $n/2$ mitochondria. To restore the total complement of $n$ mitochondria, we sample $n/2$ mitochondria with replacement from the $n/2$ mitochondria in the $U_1B_2$ cell, and add the $n/2$ sampled mitochondria to the original set of mitochondria to form a cell with $n$ mitochondria.

For a gamete with identity $M_{t,\tau_1} = (p, q, U_1)$ to produce a diploid cell with identity $M_{t,\tau_2} = (i, j, U_1B_2)$, it must sample $n/2$ mitochondria containing $i - p$ type $I$ mitochondria and $j - q$ type $J$ mitochondria. The mitochondrial state of the $B_2$ gamete is irrelevant because its mitochondria are discarded and we will refer to this cell as $M_{t,\tau_1} = (r, s, B_2)$.

Sampling of the $n/2$ mitochondria follows a multinomial distribution, which we denote $T(i - p, j - q; n/2, 2p/n, 2q/n)$, where $i - p$ and $j - q$ refer to the number of type $I$ and $J$ mitochondria that need to be sampled, $n/2$ refers to the number of mitochondria being sampled, and $2p/n$ and $2q/n$ refer to the probability of drawing type $I$ and $J$ mitochondria respectively from a $U_1B_2$ cell that contains $p$ type $I$ and $q$ type $J$ mitochondria (where $2p/n$ is obtained by rearranging $p/(n/2)$ and $2q/n$ is
obtained by rearranging \( q/(n/2) \). The probability of sampling \( i - p \) and \( j - q \) type I and \( J \) mitochondria respectively (and \( n/2 - (i - p) - (j - q) \) type \( K \) mitochondria) is given by

\[
T \left( i - p, j - q; \frac{n}{2}, \frac{2p}{n}, \frac{2q}{n} \right) = \frac{n/2!}{(i - p)! (j - q)! \left( \frac{n}{2} - (i - p) - (j - q) \right)!} \left( \frac{2p}{n} \right)^{i-p} \left( \frac{2q}{n} \right)^{j-q} \left( \frac{2 \left( \frac{n}{2} - p - q \right)}{n} \right)^{n/2 - (i-p)-(j-q)}.
\]

The restrictions on \( p, q, i-p, \) and \( j-q \) are the same as those in biparental mating. Since \( U_1 \) will form the same initial \( U_1B_2 \) cell regardless of the \( B_2 \) gamete with which it mates, the probability of producing each type of \( U_1 \) gamete is multiplied by the probability of selecting each \( B_2 \) gamete. The probability of forming a given \( U_1B_2 \) cell after random mating is

\[
P \left( M^{t, \tau_2} = (i, j, U_1B_2) \right) = 2 \sum_{p=\max(0,i-n/2)}^{\min(n/2,i)} \sum_{q=\max(0,i+j-(n/2)-p)}^{\min(n/2-p,j)} \sum_{r=0}^{n/2} \sum_{s=0}^{n/2-r} \sum P \left( M^{t, \tau_3} = (r, s, B_2) \right) T \left( i - p, j - q; \frac{n}{2}, \frac{2p}{n}, \frac{2q}{n} \right) \sum P \left( M^{t, \tau_1} = (p, q, U_1) \right) T \left( i - p, j - q; \frac{n}{2}, \frac{2p}{n}, \frac{2q}{n} \right) \sum P \left( M^{t, \tau_3} = (r, s, B_2) \right).
\]

**A.11.3 Mutation**

As each mitochondrion can mutate into either of the other two mitochondrial types, there can be many different transitions between pre-mutation cell types and a single specified post-mutation cell type, which we denote as \( M^{t, \tau_3} = (i, j, G) \), where \( \tau_3 \) refers to the post-mutation life cycle stage. We introduce six mutation variables \( (a, b, c, x_i, x_j, \text{ and } x_k) \) to describe the way in which a pre-mutation cell can mutate to a post-mutation cell (Table A.28).

Cells lose a type \( I \) mitochondria due to mutation from type \( I \) into other mitochondrial types, but concurrently gain \( x_j + x_k \) type \( I \) mitochondria that have mutated away from type \( J \) and \( K \). Thus, post-mutation cells with \( i \) type \( I \) mitochondria come from pre-mutation cells with \( i + a - x_j - x_k \) type mitochondria. Similarly, post-mutation cells with \( j \) type \( J \) mitochondria are derived from pre-mutation cells with \( j + b - x_i - (c - x_k) \) type \( J \) mitochondria, and (implicitly) post-mutation cells with \( k = n - i - j \) type \( K \) mitochondria are derived from pre-mutation cells with
A.11. Model assuming three mitochondrial types

$k + c - (a - x_i) - (b - x_j)$ type $K$ mitochondria. Thus, post-mutation cells in state $M^{t,\tau_3} = (i, j, G)$ are derived from pre-mutation cells in state

\[ M^{t,\tau_2} = (i + a - x_j - x_k, j + b - x_i - (c - x_k), G) = (it, jt, G). \]

We identify all valid combinations of mutation variables that lead from pre-mutation cells to a particular post-mutation cell in state $M^{t,\tau_3} = (i, j, G)$.

To determine allowable ranges for $a$, $b$, and $c$ we use the following conceptual approach: the number of mutations away from a particular type must be less than or equal to the total number of mitochondria that can be received by the other two mitochondrial types. Thus, $a$, the number of type $I$ mitochondria that mutate into type $J$ or type $K$ mitochondria, must be less than or equal to the sum of $j$ type $J$ mitochondria and $k$ type $K$ mitochondria. Therefore, $a \leq j + k$, where $j + k = j + (n - i - j) = n - i$, giving $0 \leq a \leq n - i$. Before we can determine $b$ (mutations away from type $J$), however, we must know how many of the $a$ mutations in the type $I$ mitochondria become type $J$ ($x_i$), as $x_i$ affects the restrictions on $b$.

Our conceptual approach for $x_i$, $x_j$, and $x_k$ is as follows. Each of these variables has three restrictions. First, by definition, these variables are less than or equal to the number of mutations in a particular mitochondrial type (e.g. $x_i$ must satisfy $0 \leq x_i \leq a$). The second and third restrictions ensure that the number of mutations to a particular mitochondrial type is less than or equal to the number of mitochondria required by that mitochondrial type. Valid values of these variables are found within the range of intersection of these three restrictions.

The first restriction for $x_i$, the number (out of $a$) of type $I$ mitochondria that mutate to type $J$, is $0 \leq x_i \leq a$. The second restriction is $x_i \leq j$ (number of mutations to type $J$ is less than or equal to the number of mutations required by type $J$) and the third restriction is $a - x_i \leq k$ (number of mutations to type $K$ is less than or equal to the number of mutations required by type $K$), which gives $x_i \geq a - n + i + j$. The intersection of these restrictions gives $\max(0, a - n + i + j) \leq x_i \leq \min(j, a)$.

Now we move to $b$, the number of type $J$ mitochondria that mutate into type $I$ or type $K$ mitochondria. The mutations in type $J$ cannot be greater than the number of type $I$ and type $K$ mitochondria required by the cell post-mutation. Thus, $b \leq i + k - (a - x_i)$, where $i$ is the number of required type $I$ mitochondria and $k - (a - x_i)$ is the number of type $K$ mitochondria still required (because the cell
receives $a - x_i$ type $K$ from mutations in type $I$). $i + k - (a - x_i) = n - a - j + x_i$, giving $0 \leq b \leq n - a - j + x_i$.

The first restriction on $x_j$, the number (out of $b$) of type $J$ mitochondria that mutate to type $I$, is $0 \leq x_j \leq b$. The second restriction is $x_j \leq i$ (the number of mutations that go to type $I$ cannot be larger than the number of required type $I$—type $I$ to receive any mutations from other types) and the third restriction is $b - x_j \leq k - (a - x_i)$ (the number of mutations that go to type $K$ must be less than or equal to the number of type $K$ still needed by the cell), which when rearranged gives $x_j \geq i - n + a + b + j - x_i$. The intersection of these inequalities is $\max(0, i - n + a + b + j - x_i) \leq x_j \leq \min(i, b)$.

$c$, the number of type $K$ mitochondria that mutate into type $I$ or type $J$ mitochondria, cannot be greater than the remaining complement of type $I$ (now needs $i - x_j$ mitochondria after receiving $x_j$ from type $J$), and $J$ (still needs $j - x_i$ mitochondria). Thus, $c \leq (i - x_i) + (j - x_i)$, which gives $0 \leq c \leq i + j - x_i - x_j$.

Finally, the first restriction on $x_k$, the number of $c$ mutations in type $K$ mitochondria that mutate to type $I$, is $0 \leq x_k \leq c$, the second restriction is $x_k \leq i - x_j$, and the third restriction is $c - x_k \leq j - x_i$, which can be rearranged to give $x_k \geq c - j - x_i$. The intersection of these inequalities gives $\max(0, c - j + x_i) \leq x_k \leq \min(i - x_j, c)$.

Once we have determined all valid combinations of mutation variables, we must determine the probability of each transition (a single combination of mutation variables) from a pre-mutation cell to our specified post-mutation cell. To determine the probability of a single transition, we multiply the probabilities associated with each mutation variable by the probability of the pre-mutation cell.

$Y(a; i + a - x_j - x_k, \mu) = Y(a; i, \mu)$ represents the probability of obtaining $a$ mutations in pre-mutation cells that contain $i$ type $I$ mitochondria. The accumulation of $a$ mutations follows a binomial distribution,

$$Y(a; i, \mu) = \binom{i}{a} \mu^a (1 - \mu)^{i - a},$$

where $\mu$ is the probability of a mitochondrion being chosen for mutation. $Z(x_i; a, P_{ij})$ represents the probability that $x_i$ mutations (out of $a$) become $J$ mitochondria and $a - x_i$ mutations become $K$ mitochondria, where each mutation becomes type $J$ with probability $P_{ij}$. This follows a binomial distribution,

$$Z(x_i; a, P_{ij}) = \binom{a}{x_i} P_{ij}^{x_i} (1 - P_{ij})^{a - x_i}. \quad (A.11)$$
We let $Y(b; j + b - x_i - (c - x_k), \mu) = Y(b; j_t, \mu)$ represent the probability of obtaining $b$ mutations in pre-mutation cells that contain $j_t$ type $J$ mitochondria, where

$$ Y(b; j_t, \mu) = \binom{j_t}{b} \mu^b (1 - \mu)^{j_t - b - c + x_k}. $$

Likewise, $Y(c; k + c - (a - x_i) - (b - x_j), \mu) = Y(c; k_t, \mu)$ is the probability of obtaining $c$ mutations in the pre-mutation cells that contain $k_t$ type $K$ mitochondria, where

$$ Y(c; k_t, \mu) = \binom{k_t}{c} \mu^c (1 - \mu)^{k_t - (a - x_i) - (b - x_j)}. $$

$Z(x_j; b, P_{ji})$ is the probability that $x_j$ (out of $b$) mutations become $I$ mitochondria and $b - x_j$ mutations become $K$ mitochondria, given that each mutation becomes type $I$ with probability $P_{ji}$, where

$$ Z(x_j; b, P_{ji}) = \binom{b}{x_j} P_{ji}^{x_j} (1 - P_{ji})^{b-x_j}. \quad \text{(A.12)} $$

$Z(x_k; c, P_{ki})$ is the probability that $x_k$ (out of $c$) mutations become $I$ mitochondria and $b - x_j$ mutations become $J$ mitochondria, where mutations become type $I$ with probability $P_{ki}$, where

$$ Z(x_k; c, P_{ki}) = \binom{c}{x_k} P_{ki}^{x_k} (1 - P_{ki})^{c-x_k}. \quad \text{(A.13)} $$

When there are no fitness differences between the mitochondrial types, $P_{ij} = P_{ji} = P_{ki} = 0.5$ (i.e. there is no bias in mutation between $I$, $J$, and $K$). Putting everything together, we can determine the probability of any post-mutation cell via

$$ P(M^{i,\tau_3} = (i, j, G)) = \frac{\sum_{a=0}^{n-i} \sum_{b=0}^{n-a-j+x_i} \sum_{c=0}^{n-a-j-x_i} \binom{b}{x_j} \mu^b (1 - \mu)^{j_t - b - c + x_k} \binom{c}{x_k} \mu^c (1 - \mu)^{k_t - (a - x_i) - (b - x_j)} \times Y(a; i_t, \mu) Y(b; j_t, \mu) Y(c; k_t, \mu) Z(x_i; a, P_{ij}) Z(x_j; b, P_{ji}) Z(x_k; c, P_{ki}) \sigma (M^{i,\tau_2} = (i_t, j_t, G))}{\sigma}, $$

where

$$ \sigma = \sum_{i=0}^{n} \sum_{j=0}^{n-i} P(M^{i,\tau_3} = (i, j, U_1 B_2)) + P(M^{i,\tau_3} = (i, j, B_1 B_2)). $$
so that the sum of the proportions of the population equals 1.

A.11.4 Selection

The fitness function is substantially more complicated when we consider three mitochondrial types. We generated two fitness functions that are similar to the concave and convex fitness functions in the main model Figure A.14. The fitness shape is given by

$$\omega'(i,j) = \frac{e^{-tu} - 0.5}{1 + e^{-tu}}$$  \hspace{1cm} \text{(A.14)}$$

where

$$u(i,j) = v \sqrt{(i-y)^2 + (j-y)^2 + (k-y)^2}.$$  

t, v, and y are variables that alter the shape and compression of the fitness function. The values of these variables were chosen to generate a three-dimensional equivalent to the concave fitness function and the convex fitness function (Table A.29 and Figure A.14). We normalize the fitness shape so that maximum fitness equals 1 and minimum fitness equals a pre-determined value, $h$ (Table A.29). The cost of heteroplasmy, $c_h$, is given by $c_h = 1 - h$.

$$\omega(i,j) = \frac{(1-h)(\omega'(i,j) - \min(\omega'))}{\max(\omega') - \min(\omega')} + h,$$

where $\min(\omega')$ and $\max(\omega')$ are the minimum and maximum values of $\omega'$ over the domain of valid values for $i$ and $j$. The state $P(M_{t,\tau}^6 = (l,m,2G))$ represents cells after selection, which we determine via $P(M_{t,\tau}^6 = (i,j,G)) = P(M_{t,\tau}^4 = (i,j,G)) \omega(i,j)$.

A.11.5 Meiosis

As in the main model, we sample $n$ mitochondria with replacement from a cell containing $n$ mitochondria and add the set of sampled mitochondria to the original set of mitochondria to form a cell containing $2n$ mitochondria. We let $M_{t,\tau}^6 = (l,m,2G)$ represent the cell with doubled mitochondria and nuclear genotype, where $l$ takes values in $\{0,1,\ldots,2n\}$, $m$ takes values in $\{0,1,\ldots,2n-l\}$ and $2G$ takes values in $\{U_1U_1B_2B_2, B_1B_1B_2B_2\}$.

We denote the probability of sampling $l-i$ type $I$ mitochondria and $m-j$ type $J$ mitochondria from $M_{t,\tau}^6 = (i,j,G)$ as $F(l-i,m-j; n, i/n, j/n)$. Sampling follows a multinomial distribution, giving
A.11. Model assuming three mitochondrial types

\[ F(l - i, m - j; n, i/n, j/n) = \frac{n!}{(l - i)! (m - j)! (n - (l - i) - (m - j))!} \left( \frac{i}{n} \right)^{l - i} \left( \frac{j}{n} \right)^{m - j} \left( \frac{k}{n} \right)^{n - (l - i) - (m - j)}. \]

We obtain \( M^{l,\tau_6} = (l, m, 2G) \) by

\[
P(M^{l,\tau_6} = (l, m, 2G)) = \min(l,n) \min(m,n-i) \sum_{i = \max(0, l-n)} \sum_{j = \max(0, m+l-n-i)} F(l - i, m - j; n, i/n, j/n) P(M^{l,\tau_6} = (i, j, G)).
\]

During the second step of meiosis, the cells with \( 2n \) mitochondria produce gametes with \( n/2 \) mitochondria. We define \( S(p, q; 2n, l, m, n/2) \) to be the probability of obtaining \( p \) type I and \( q \) type J mitochondria in \( n/2 \) draws from the \( M^{l,\tau_6} = (l, m, 2G) \) cell that contains \( l \) type I and \( m \) type J mitochondria (out of \( 2n \) total mitochondrial). Here, sampling is without replacement and follows a multivariate hypergeometric distribution, giving

\[
S(p, q; 2n, l, m, n/2) = \frac{\binom{l}{p} \binom{m}{q} \binom{2n-l-m}{n/2-p-q}}{\binom{2n}{n/2}}.
\]

Gametes produced by meiosis are represented by \( M^{l+1,\tau_1} = (p, q, g) \). We determine the probability of obtaining a particular gamete using

\[
P(M^{l+1,\tau_1} = (p, q, U_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S(p, q; 2n, l, m, n/2) P(M^{l,\tau_6} = (l, m, U_1U_2B_2)) \right],
\]

\[
P(M^{l+1,\tau_1} = (p, q, B_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S(p, q; 2n, l, m, n/2) P(M^{l,\tau_6} = (l, m, B_1B_2B_2)) \right],
\]

and
A.11. Model assuming three mitochondrial types

\[ P \left( M_{t+1,\tau}^{t+1} = (p, q, B_2) \right) \]
\[ = \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S \left( p, q; 2n, l, m, \frac{n}{2} \right) P \left( M_{t,\tau}^{t+1} = (l, m, U_1 U_1 B_2 B_2) \right) \right] \]
\[ + \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S \left( p, q; 2n, l, m, \frac{n}{2} \right) P \left( M_{t,\tau}^{t+1} = (l, m, B_1 B_1 B_2 B_2) \right) \right]. \]

A.11.6 Deleterious mutations

We alter the fitness function slightly to account for non-neutral mutations. First, we determine \( \omega' \) using Equation A.14 as before. We reduce the fitness of types \( I \) and \( J \) mitochondria via

\[ \omega''(i, j) = \omega'(i, j) \left( 1 - \left( \frac{i \phi}{n} \right)^2 \right) \left( 1 - \left( \frac{j \phi}{n} \right)^2 \right), \]

where \( \phi \) controls the severity of the deleterious mutation (Table A.30; column 1). Fitness is normalized as before by

\[ \omega_d(i, j) = \frac{(1 - h) (\omega''(i, j) - \min(\omega''))}{\max(\omega'') - \min(\omega'')} + h. \]

We choose a value of \( \phi \) that leads to cells that are homoplasmic for type \( I \) or \( J \) having a fitness of \( 1 - s_d \) (i.e. \( \omega_d(n, 0) = \omega_d(0, n) = 1 - s_d \)). We also alter Equation A.11, Equation A.12, and Equation A.13. Under the deleterious scenario, \( P_{ij} = 0.99 \), \( P_{ji} = 0.99 \), and \( P_{ki} = 0.01 \). The probability that a deleterious type mutates to a neutral type is lower than the probability that a deleterious type mutates to another deleterious type (as the former is effectively an advantageous mutation).

A.11.7 Advantageous mutations

Again, we determine \( \omega' \) using Equation A.14 but now type \( I \) mitochondria have an advantage, determined via

\[ \omega''(i, j) = \omega'(i, j) \left( 1 + \left( \frac{i \zeta}{n} \right)^2 \right), \]

where \( \zeta \) scales the benefit of the advantageous mutation (Table A.30; column 2). Fitness is normalized so that maximum fitness is 1,
\[ \omega_a(i, j) = \frac{(1 - h)(\omega''(i, j) - \min(\omega''))}{\max(\omega'') - \min(\omega'')} + h. \]

We choose a value of \( \zeta \) so that cells homoplasmic for type I have a fitness of 1, while cells homoplasmic for type J and K have a fitness of \( 1 - s_a \) (i.e. \( \omega_a(n, 0) = 1 \) and \( \omega_a(0, 0) = \omega_a(0, n) = 1 - s_a \)). Again, we also alter Equation A.11, Equation A.12, and Equation A.13. Under the advantageous scenario, \( P_{ij} = 0.5, P_{ji} = 0.01, \) and \( P_{ki} = 0.01. \) This accounts for the fact that advantageous mutations are less common than deleterious mutations (i.e. mutations from advantageous to neutral).
Bibliography


Appendix B

Chapter 3: Supplementary Material

B.1 Supplementary figures
B.1. Supplementary figures

Figure B.1: **Time to accumulate a beneficial substitution.** Each plot shows the number of generations to accumulate a beneficial substitution (number of generations before each cytoplasmic genome carries at least $\gamma = 5$ substitutions divided by the mean substitutions per genome in that generation). Parameter values for \textbf{A–B}: $N = 1000$, $n = 50$, $\mu_b = 10^{-8}$, and $b = 25$ (relaxed transmission bottleneck) or $b = 5$ (tight transmission bottleneck). \textbf{A}. Selection coefficient of 0.1. \textbf{B}. Selection coefficient of 0.01. Parameter values for \textbf{C} (unless otherwise stated on the x-axis): $N = 1000$, $n = 50$, $\mu_b = 10^{-8}$, $s_b = 0.1$, a linear fitness function for beneficial substitutions, and $b = n/2$ (relaxed transmission bottleneck) or $b = n/10$ (tight transmission bottleneck). Error bars are standard error of the mean.
Expected separation

Beneficial sweep  Deleterious sweep

A

B

C

D

φ << 1
φ ~ 1
φ > 1
To calculate the genetic hitchhiking index ($\phi$), we compare the number of generations separating beneficial and deleterious sweeps to the number of generations we expect if the two events are uncorrelated. We examine all beneficial sweeps except those involving genomes with > 5 beneficial substitutions (to maintain consistency between the different fitness functions). We map each beneficial sweep to a single deleterious sweep but do not limit the number of times a single deleterious sweep can be mapped to (e.g. B and C). The expected separation between beneficial and deleterious sweeps for this hypothetical example is shown at the top of the figure. See below for details of how the index is calculated.

A. When beneficial sweeps are closely followed by deleterious sweeps, $\phi < 1$ and we infer that genetic hitchhiking has occurred.

B. When the mean of the number of generations separating beneficial and deleterious sweeps are as expected, $\phi \approx 1$ and we infer that the beneficial sweep does not affect the deleterious sweep.

C. When deleterious sweeps follow beneficial sweeps later than expected, $\phi > 1$ and we infer that genetic hitchhiking is suppressed.

D. When a beneficial sweep is followed by a deleterious sweep, we call it a “paired” sweep. In some instances, the simulation terminates before a deleterious sweep can follow a beneficial sweep (an “unpaired” sweep; e.g. the last beneficial sweep in D). For unpaired sweeps, we add the number of generations separating the beneficial sweep and the end of the simulation. To calculate the mean generations separating the sweeps, however, we only divide by the number of paired sweeps. Thus, the equation for the index is $\phi = \left[ \frac{\sum_{i=1}^{n_p} (g_d(i) - g_b(i)) + \sum_{j=1}^{n_u} (g_t - g_b(j))}{n_p} \right] / E[s]$. $n_p$ is the total number of paired sweeps, $g_d(i)$ is the generation in which the $i$th paired deleterious sweep occurred, and $g_b(i)$ is the generation in which the $i$th paired beneficial sweep occurred. $n_u$ is the total number of unpaired sweeps, $g_t$ is the number of generations in each run (10000), and $g_b(j)$ is the generation in which the $j$th unpaired beneficial sweep occurred. $E[s]$ is the expected separation in generations and given by $E[s] = \left[ \frac{\sum_{k=1}^{r} g_d(k) / d(k)}{r} \right] - 1$, where $d(k)$ is the number of deleterious sweeps we considered in the $k$th simulation, $g_d(k)$ is the generation at which the $d(k)$th deleterious sweep occurred in the $k$th simulation, and $r$ is the number of runs for each set of parameter values (500). We subtract 1 because the deleterious sweeps can occur in the same generation as the beneficial sweep.
Figure B.3: Genetic hitchhiking when beneficial mutations are rare. Parameters: \( N = 1000, n = 50, \mu_b = 10^{-9}, \mu_d = 10^{-7}, \) and \( b = 25 \) (relaxed transmission bottleneck) or \( b = 5 \) (tight transmission bottleneck). A shows a selection coefficient of 0.1 while B shows a selection coefficient of 0.01. The plots show the overall level of genetic hitchhiking in each population, measured by our genetic hitchhiking index (see Figure B.2 for details). When \( \phi < 1 \), it indicates the presence of genetic hitchhiking. Error bars are ± standard error of the mean. Note that this figure depicts a beneficial mutation rate 10 times smaller than shown in Figure 3.5 (\( \mu_b = 10^{-9} \) versus \( \mu_b = 10^{-8} \)).
Figure B.4: Ratio of beneficial to deleterious substitutions accumulated under the two inheritance modes. Parameters: $N = 1000$, $n = 50$, $\mu_d = 10^{-7}$, and $b = 25$ (relaxed transmission bottleneck) or $b = 5$ (tight transmission bottleneck). Panels A and B show selection coefficients of $s_b = s_d = 0.01$, while panels C and D show selection coefficients of $s_b = s_d = 0.1$. For panels A and C, the beneficial mutation rate is $\mu_b = 10^{-8}$, while for panels B and D the beneficial mutation rate is $\mu_b = 10^{-9}$. In all cases, uniparental inheritance has a higher ratio of beneficial to deleterious substitutions than biparental inheritance. Error bars are ± standard error of the mean. See Figure 3.7 for details of how we calculate the ratio of beneficial to deleterious substitutions.
## B.2 Supplementary tables

Table B.1: Benchmarking the genetic hitchhiking index using randomly simulated data

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>inheritance</td>
<td>$b$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 25$</td>
</tr>
<tr>
<td>UPI</td>
<td>$b = 5$</td>
</tr>
<tr>
<td>BPI</td>
<td>$b = 5$</td>
</tr>
</tbody>
</table>

Parameters: $N = 1000$, $n = 50$. $\phi \pm sd$ shows the genetic hitchhiking index for randomly simulated datasets $\pm$ standard deviation. For each set of parameter values, we determined the expected distance between beneficial and deleterious sweeps. (The expected distance separating beneficial sweeps is $E[d_b] = \left(\sum_{i=1}^{r} g_b(i)/n_b(i)\right)/r$, where $n_b(i)$ is the number of beneficial sweeps we considered in the $i$th simulation, $g_b(i)$ is the generation at which the $n_b(i)$th beneficial sweep occurred in the $i$th simulation, and $r$ is the number of runs for each set of parameter values (500).
The expected distance separating deleterious sweeps is \( E[d_d] = \left( \frac{\sum_{i=1}^{r} g_d(i) / n_d(i)}{r} \right) / r, \)
where \( n_d(i) \) is the number of deleterious sweeps we considered in the \( i \)th simulation, \( g_d(i) \) is the generation at which the \( n_d(i) \)th deleterious sweep occurred in the \( i \)th simulation, and \( r \) is the number of runs for each set of parameter values. We used these expected values to generate 500 randomly simulated runs, and for each one, used binomial sampling to generate a random number of beneficial and deleterious sweeps. (The number of beneficial sweeps is given by the random variable \( R_b^i \) and the number of deleterious sweeps by the random variable \( R_d^i \), where \( i \) is the number of the simulated run (out of 500). To obtain \( R_b^i \) and \( R_d^i \), we used the R function \texttt{rbinom} with parameters \( n = 1 \), \texttt{size} = 10000, and \( \text{prob} = 1/E[d_b] \) for beneficial sweeps or \( \text{prob} = 1/E[d_d] \) for deleterious sweeps.) For each run, we uniformly sampled \( R_b^i \) beneficial and \( R_d^i \) deleterious sweeps over 10,000 generations to get the locations of our random beneficial and deleterious sweeps. We then calculated \( \phi \) in the same way as our model-generated data (Figure B.2). For each set of parameter values, we repeated this process 100 times, giving us 100 estimates of \( \phi \). The fifth column shows the mean and standard deviation of these 100 estimates. As can be seen, when beneficial and deleterious sweeps are uncorrelated, \( \phi \approx 1. \)
B.3 Beneficial mutation model

The model is an individual-based model, in which we track all cells in the population (and their gametes). It is written in R version 3.1.2 [1]. For each set of parameter values, we ran 500 Monte Carlo simulations. These Monte Carlo simulations were run using packages that enable R code to be run in parallel (doMC and foreach [2, 3]) and produce reproducible output doRNG [4]). We ran our simulations on High Performance Computing clusters at The University of Sydney (“Artemis”) and National Computational Infrastructure, Australia (“Raijin”).

We store the population of cells in a matrix called $C_{T,G}^{t,\tau,\xi}$ that has $N$ rows (each representing an individual cell) and $n$ columns (each representing a cytoplasmic genome). We will use the terminology $C_{T,G}^{t,\tau,\xi}(i,\ast)$ to refer to the $i$th row in $C_{T,G}^{t,\tau,\xi}$ (equivalently the $i$th cell in the population). $G$ represents the inheritance mode and takes values in $\{U,B\}$, where $U$ denotes a cell with uniparental inheritance and $B$ denotes a cell with biparental inheritance. The generation is given by $t$, while the stage of the life cycle is given by $\tau$. Thus,

$$C_{T,G}^{t,\tau,\xi} = \begin{bmatrix}
    C_{T,G}^{t,\tau,\xi}(1,1) & C_{T,G}^{t,\tau,\xi}(1,2) & \cdots & C_{T,G}^{t,\tau,\xi}(1,n) \\
    C_{T,G}^{t,\tau,\xi}(2,1) & C_{T,G}^{t,\tau,\xi}(2,2) & \cdots & C_{T,G}^{t,\tau,\xi}(2,n) \\
    \vdots & \vdots & \ddots & \vdots \\
    C_{T,G}^{t,\tau,\xi}(N,1) & C_{T,G}^{t,\tau,\xi}(N,2) & \cdots & C_{T,G}^{t,\tau,\xi}(N,n)
\end{bmatrix}$$

where $C_{T,G}^{t,\tau,\xi}(i,j) = \alpha$ represents $\alpha$ beneficial substitutions in the $j$th cytoplasmic genome of individual $i$. Cytoplasmic genomes have $l$ bases, each of which can mutate from a neutral site to a beneficial site. Initially, all genomes have $\alpha = 0$ beneficial substitutions. The first stage of the life cycle is mutation.

B.3.1 Mutation

We only consider forward mutation (i.e. genomes can gain beneficial mutations but cannot lose beneficial mutations). We assume that the $j$th cytoplasmic genome in the $i$th cell receives $m_{ij}^{b,t}$ new beneficial mutations in generation $t$, where $m_{ij}^{b,t}$ takes values in $\{0, 1, 2, 3, 4, 5\}$. The probability that a cytoplasmic genome receives 5 mutations in a single generation is equal to the probability that a genome receives 5 or more mutations (when $\mu_b = 10^{-8}$ and $l = 20000$, the probability that a cytoplasmic genome receives more than 5 mutations in a single generation is calculated by R as 0, so this is a very accurate approximation).
The probability that a genome mutates depends on the mutation rate per base per generation \((\mu_b)\), on the number of base pairs available to be mutated \((l - \alpha)\), and on the number of mutations that occur \((m_{i,j}^{b,t})\). To store these probabilities, we generate a matrix, \(M\), with \(l + 1\) rows (\(\alpha\) can take values in \(\{0, 1, \ldots l\}\)) and 5 columns. Thus, 

\[
M = \begin{bmatrix}
M(0,0) & M(0,1) & M(0,2) & M(0,3) & M(0,4) \\
M(1,0) & M(1,1) & M(1,2) & M(1,3) & M(1,4) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
M(l,0) & M(l,1) & M(l,2) & M(l,3) & M(l,4)
\end{bmatrix}.
\]

Each generation, we generate a uniformly random number between 0 and 1, \(r_{i,j}^{b,t}\), which determines the number of mutations gained by the \(j\)th cytoplasmic genome in the \(i\)th cell in generation \(t\) (i.e. \(r_{i,j}^{b,t}\) is matched to \(C_{G}^{l,t,\tau}(i,j)\)). \(r_{i,j}^{b,t}\) causes \(m_{i,j}^{b,t}\) mutations in a genome that already carries \(\alpha\) substitutions according to

\[
m_{i,j}^{b,t} = 5 \text{ if } r_{i,j}^{b,t} < M(\alpha, 0),
\]

\[
m_{i,j}^{b,t} = 5 - x \text{ if } M(\alpha, x - 1) \leq r_{i,j}^{b,t} < M(\alpha, x) \text{ for } 1 \leq x \leq 4
\]

\[
m_{i,j}^{b,t} = 0 \text{ if } r_{i,j}^{b,t} \geq M(\alpha, 4)
\]

The entries of \(M\) are given by

\[
M(\alpha, 0) = 1 - \sum_{m_{i,j}^{b,t} = 0}^{4} \binom{l - \alpha}{m_{i,j}^{b,t}} \mu_b^{m_{i,j}^{b,t}} (1 - \mu_b)^{l - \alpha - m_{i,j}^{b,t}}
\]

and

\[
M(\alpha, x) = 1 - \sum_{m_{i,j}^{b,t} = 0}^{4} \binom{l - \alpha}{m_{i,j}^{b,t}} \mu_b^{m_{i,j}^{b,t}} (1 - \mu_b)^{l - \alpha - m_{i,j}^{b,t}}
\]

\[
+ \sum_{y=5-x}^{4} \binom{l - \alpha}{y} \mu_b^{y} (1 - \mu_b)^{l - \alpha - y} \text{ for } 1 \leq x \leq 4.
\]
For the $j$th cytoplasmic genome in the $i$th cell, we add the $m_{ij}^{b,t}$ new mutations to the existing $\alpha$ substitutions according to

$$C_G^{t,\tau_2}(i,j) = C_G^{t,\tau_1}(i,j) + m_{ij}^{b,t}$$

### B.3.2 Selection

The next life cycle stage is selection. Here, each cell is assigned a fitness value based on the number of beneficial cytoplasmic substitutions they carry. The number of beneficial substitutions carried by the $i$th cell is given by $\beta(i)$, where

$$\beta(i) = \sum_{j=1}^{n} C_G^{t,\tau_2}(i,j).$$

We examine three fitness functions: concave up, linear, and concave down. The fitness of the $i$th cell under the concave up fitness function is given by

$$\omega_{u,b}(\beta(i)) = 1 + s_b \left[ \left( \frac{\beta(i)}{n\gamma} \right)^2 - 1 \right],$$

the fitness of the $i$th cell under the linear fitness function by

$$\omega_{l,b}(\beta(i)) = 1 + s_b \left[ \frac{\beta(i)}{n\gamma} - 1 \right],$$

and the fitness of the $i$th cell under the concave down fitness function by

$$\omega_{d,b}(\beta(i)) = 1 + s_b \left[ \sqrt{\frac{\beta(i)}{n\gamma}} - 1 \right],$$

where $\gamma$ is the number of beneficial substitutions each cytoplasmic genome must accumulate before the simulation terminates, $n$ is the number of cytoplasmic genomes in each cell, and $s_b$ is the beneficial selection coefficient. We then normalize each cell’s fitness so that they sum to 1. The 1-by-$N$ vector $S'_G$ stores the normalized fitness of the population, where $S'_G(i)$ gives the relative fitness of the $i$th cell in the population. To generate $S'_G$, we first generate a temporary 1-by-$N$ vector, $S''_G$, where

$$S''_G(i) = \omega_{f,b}(\beta(i)).$$

where $f$ represents the fitness function used. To generate $S'_G$, we normalize this vector according to
\[ S^t_G(i) = \frac{S'_{Gz}(i)}{N \sum_{z=1} S'_{Gz}(z)} \]

Finally, we feed these probabilities into a multinomial distribution (function \texttt{rmultinomial} in the \texttt{multinomRob} package \cite{5}) to generate \( N \) new cells for the population. Cells can thus die, replace themselves, or produce multiple copies of themselves. We pass the \texttt{rmultinomial} function the arguments \( N \) and the probability vector \( S^t_G \), which generates a 1-by-\( N \) vector, \( O^t_G \), whose sum is \( N \) and whose \( i \)th entry represents the number of “offspring” left by the \( i \)th cell in the pre-selection population described by \( C^t_{G2} \). We then use these offspring to reform the post-selection population described by \( C^t_{G3} \), assuming that each offspring is a perfect copy of its parent. For example, if \( O^t_G(i) = 2 \) then in \( C^t_{G3} \) there will be two copies of \( C^t_{G2}(i,*) \).

\section*{B.3.3 Meiosis}

Each cell produces two gametes: one with mating type \( A \) and the other with mating type \( a \).

\subsection*{B.3.3.1 Biparental inheritance}

To choose which cytoplasmic genomes are passed on, for each mating type we generate a matrix, \( H^t_g(i,d) = Y \) with \( N \) rows and \( b \) columns populated with uniformly random positive integers \( (Y) \) in the set \( \{1,2,...n\} \), where \( g \) represents the nuclear allele of the gamete and when inheritance is biparental takes values in \( \{B_A,B_a\} \). \( H^t_g(i,d) = Y \) denotes that the \( d \)th genome chosen for the new gamete of type \( g \) is derived from the \( Y \)th cytoplasmic genome of the \( i \)th cell. Sampling is with replacement and gametes are stored in a matrix, \( G^t_{Bg} \), which has \( N \) rows and \( b \) columns. \( G^t_{BA}(i,d) \) is produced by

\[ G^t_{BA}(i,d) = C^t_{BA}(i,H^t_{BA}(i,d) = Y). \]

\( G^t_{Ba}(i,d) \) is produced by

\[ G^t_{Ba}(i,d) = C^t_{Ba}(i,H^t_{Ba}(i,d) = Y). \]
B.3. Beneficial mutation model

B.3.3 Uniparental inheritance

When inheritance is uniparental, \( g \) takes values in \( \{U_A, U_a\} \). \( G^{t,\tau_4}_{U_A}(i, d) \) is produced by

\[
G^{t,\tau_4}_{U_A}(i, d) = C^{t}_{U}(i, H^{t}_{U_A}(i, d) = Y),
\]

and \( G^{t,\tau_4}_{U_a}(i, d) \) is produced by

\[
G^{t,\tau_4}_{U_a}(i, d) = C^{t}_{U}(i, H^{t}_{U_a}(i, d) = Y).
\]

B.3.4 Random mating

B.3.4.1 Biparental inheritance

Biparental inheritance simply combines the cytoplasmic genomes of both gametes. For each of the \( B_A^- \) and \( B_a^- \)-carrying gametes, we generate a 1-by-\( N \) vector, \( T_g(i) = Z \) that contains a random ordering (without replacement) of positive integers from the set \( \{1, 2, \ldots, N\} \). We use these vectors to pair up gametes according to

\[
C^{t+1,\tau_1}_{B}(i, *) = G^{t,\tau_4}_{B_A^-}(T^{t}_{B_A^-}(i) = Z, *) \parallel G^{t,\tau_4}_{B_A^-}(T^{t}_{B_a^-}(i) = Z, *),
\]

where \( \parallel \) indicates that the two vectors are concatenated. \( C^{t+1,\tau_1}_{B} \) is a temporary matrix (to be replaced by \( C^{t+1,\tau_1}_{B} \)), which contains 2\( b \) columns (representing the 2\( b \) genomes). Since 2\( b < n \) when we impose a transmission bottleneck, the final step for each cell is to sample \( n \) genomes with replacement from these 2\( b \) genomes (we include this step even when the transmission bottleneck is relaxed and 2\( b = n \)). This sampling follows the same approach as described in meiosis, but now instead of choosing \( b \) genomes from a cell with \( n \) genomes, we choose \( n \) genomes from a cell with 2\( b \) genomes. We generate a matrix, \( F_{B}^{t}(i, j) = Q \) with \( N \) rows and \( n \) columns populated with uniformly random positive integers sampled with replacement from the set \( \{1, 2, \ldots, 2b\} \), which we use to sample the new genomes according to

\[
C^{t+1,\tau_1}_{B}(i, j) = C^{t+1,\tau_1}_{B}(i, F_{G}^{t}(i, j) = Q).
\]

B.3.4.2 Uniparental inheritance

Under uniparental inheritance, only the gamete with mating type \( A \) passes on its cytoplasmic genomes. Thus, to pair up gametes we only need to generate one 1-by-\( N \)
vector, $T_{UA}^t(i) = Z$ that contains a random ordering (without replacement) of positive integers in the set \{1, 2, ...N\}, giving

$$C_{UA}^{t+1, \tau_1}(i, *) = G_{UA}^{t, \tau_2}(T_{UA}^t(i) = Z, *) .$$

(Note, randomly ordering the $UA$ gametes is not strictly necessary, but we do it to be consistent with the model of biparental inheritance.) Now $C_{UA}^{t+1, \tau_1}(i, *)$ only contains $b$ columns (representing $b$ genomes), so for each cell we sample $n$ genomes with replacement from these $b$ genomes. We generate a matrix, $F_{UA}^t(i, j) = Q$ with $N$ rows and $n$ columns populated with uniformly random positive integers sampled with replacement from the set \{1, 2, ...b\}. We use this to sample the new genomes according to

$$C_{UA}^{t+1, \tau_1}(i, j) = C_{UA}^{t+1, \tau_1}(i, F_{UA}^t(i, j) = Q) .$$
B.4 Deleterious mutation model

This model differs from the previous model in how it deals with selection. Mutations are now deleterious, not beneficial. Each cell is assigned a fitness value based on the number of deleterious cytoplasmic substitutions it carries. The number of deleterious substitutions carried by the $i$th cell is given by $\rho(i)$, where

$$\rho(i) = \sum_{j=1}^{n} C_{G}^{l,\tau_{2}}(i, j).$$

For deleterious mutations, we examine the concave down (decreasing) fitness function. The fitness of the $i$th cell is given by

$$\omega_{d,d}(\rho(i)) = 1 - s_{d} \left( \frac{\rho(i)}{n\gamma} \right)^{2},$$

where $n$ is the number of cytoplasmic genomes in each cell, and $s_{d}$ is the deleterious selection coefficient. To maintain consistency with the model that considers only beneficial mutations, $\gamma$ is set to the same value as in the first model. If $\omega_{d,d}(\rho(i)) < 0$ we set $\omega_{d,d}(\rho(i)) = 0$ (as fitness cannot be negative). Everything else proceeds as detailed in section B.3.2.
B.5 Beneficial and deleterious mutation model

In this version of the model, we store the population of cells in a matrix called $C_{t,τ}^G$ that has $2N$ rows and $n$ columns. $C_{t,τ}^G(i, j)$ stores the number of beneficial substitutions in the $j$th genome of the $i$th cell, while $C_{t,τ}^G(i + N, j)$ stores the number of deleterious substitutions in the $j$th genome of the $i$th cell. As before, $G$ represents the inheritance mode and takes values in $\{U, B\}$. The generation is given by $t$, while the stage of the life cycle is given by $τ$. Thus,

$$C_{t,τ}^G = \begin{bmatrix}
C_{t,τ}^{G,1}(1,1) & C_{t,τ}^{G,1}(1,2) & \ldots & C_{t,τ}^{G,1}(1,n) \\
C_{t,τ}^{G,2}(1,1) & C_{t,τ}^{G,2}(1,2) & \ldots & C_{t,τ}^{G,2}(1,n) \\
\vdots & \vdots & \ddots & \vdots \\
C_{t,τ}^{G,2N}(1,1) & C_{t,τ}^{G,2N}(1,2) & \ldots & C_{t,τ}^{G,2N}(1,n)
\end{bmatrix},$$

where $C_{t,τ}^{G,1}(i, j) = \alpha$ and $C_{t,τ}^{G,1}(i + N, j) = \kappa$ represent $\alpha$ beneficial substitutions and $\kappa$ deleterious substitutions respectively in the $j$th cytoplasmic genome of individual $i$. Cytoplasmic genomes have $l$ bases, each of which can change from a neutral site to a beneficial or deleterious substitution. Initially, all genomes have $\alpha = 0$ beneficial substitutions and $\kappa = 0$ deleterious substitutions. The first stage of the life cycle is mutation.

B.5.1 Mutation

We assume that the $j$th cytoplasmic genome in the $i$th cell gains $m_{ij}^{b,t}$ new beneficial mutations in generation $t$, and $m_{ij}^{d,t}$ new deleterious mutations in generation $t$, where both $m_{ij}^{b,t}$ and $m_{ij}^{d,t}$ take values in $\{0, 1, 2, 3, 4, 5\}$. We store the probabilities of gaining $m_{ij}^{b,t}$ beneficial mutations in a matrix, $M_b$, with $l + 1$ rows (representing the possible states that a cytoplasmic genome can take) and 5 columns. Thus,

$$M_b = \begin{bmatrix}
M_b(0, 0) & M_b(0, 1) & M_b(0, 2) & M_b(0, 3) & M_b(0, 4) \\
M_b(1, 0) & M_b(1, 1) & M_b(1, 2) & M_b(1, 3) & M_b(1, 4) \\
M_b(2, 0) & M_b(2, 1) & M_b(2, 2) & M_b(2, 3) & M_b(2, 4) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
M_b(l, 0) & M_b(l, 1) & M_b(l, 2) & M_b(l, 3) & M_b(l, 4)
\end{bmatrix}.$$

Likewise, we store the probabilities of gaining $m_{ij}^{d,t}$ deleterious mutations in a matrix, $M_d$, given by
B.5. Beneficial and deleterious mutation model

\[ M_d = \begin{bmatrix}
M_d(0,0) & M_d(0,1) & M_d(0,2) & M_d(0,3) & M_d(0,4) \\
M_d(1,0) & M_d(1,1) & M_d(1,2) & M_d(1,3) & M_d(1,4) \\
M_d(2,0) & M_d(2,1) & M_d(2,2) & M_d(2,3) & M_d(2,4) \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
M_d(l,0) & M_d(l,1) & M_d(l,2) & M_d(l,3) & M_d(l,4)
\end{bmatrix}. \]

Each generation, we generate two uniformly random numbers between 0 and 1, \( r_{b,t}^{i,j} \) and \( r_{d,t}^{i,j} \), where \( r_{b,t}^{i,j} \) determines the number of beneficial mutations gained by the \( j \)th cytoplasmic genome in the \( i \)th cell in generation \( t \) and \( r_{d,t}^{i,j} \) determines the number of deleterious mutations gained by the \( j \)th cytoplasmic genome in the \( i \)th cell in generation \( t \) (i.e. \( r_{b,t}^{i,j} \) is matched to \( C_{G}^{i,j}(N + i, j) \) and \( r_{d,t}^{i,j} \) is matched to \( C_G^{i,j}(N + i, j) \)). \( r_{b,t}^{i,j} \) causes \( m_{b,t}^{i,j} \) beneficial mutations in the \( j \)th genome of the \( i \)th cell, which already carries \( \alpha + \kappa \) mutations according to

\[ m_{b,t}^{i,j} = 5 \text{ if } r_{b,t}^{i,j} < M_b(\alpha + \kappa, 0), \]

\[ m_{b,t}^{i,j} = 5 - x \text{ if } M_b(\alpha + \kappa, x - 1) \leq r_{b,t}^{i,j} < M_b(\alpha + \kappa, x) \text{ for } 1 \leq x \leq 4, \]

\[ m_{b,t}^{i,j} = 0 \text{ if } r_{b,t}^{i,j} \geq M_b(\alpha + \kappa, 4). \]

The entries of \( M_b \) are given by

\[ M_b(\alpha + \kappa, 0) = 1 - \sum_{m_{b,t}^{i,j}=0}^{4} \binom{l - \alpha - \kappa}{m_{b,t}^{i,j}} \mu_b^{m_{b,t}^{i,j}} (1 - \mu_b)^{l - \alpha - \kappa - m_{b,t}^{i,j}} \]

and

\[ M_b(\alpha + \kappa, x) = 1 - \sum_{m_{b,t}^{i,j}=0}^{4} \binom{l - \alpha - \kappa}{m_{b,t}^{i,j}} \mu_b^{m_{b,t}^{i,j}} (1 - \mu_b)^{l - \alpha - \kappa - m_{b,t}^{i,j}} + \sum_{y=5-x}^{4} \binom{l - \alpha - \kappa}{y} \mu_b^{y} (1 - \mu_b)^{l - \alpha - \kappa - y} \text{ for } 1 \leq x \leq 4. \]
$r_{ij}^{d,t}$ causes $m_{ij}^{d,t}$ deleterious mutations in the $j$th genome of the $i$th cell, which already carries $\alpha + \kappa$ mutations according to

$$m_{ij}^{d,t} = 5 \text{ if } r_{ij}^{d,t} < M_d(\alpha + \kappa, 0),$$

$$m_{ij}^{d,t} = 5 - x \text{ if } M_d(\alpha + \kappa, x - 1) \leq r_{ij}^{d,t} < M_d(\alpha + \kappa, x) \text{ for } 1 \leq x \leq 4,$$

$$m_{ij}^{d,t} = 0 \text{ if } r_{ij}^{d,t} \geq M_d(\alpha + \kappa, 4).$$

The entries of $M_d$ are given by

$$M_d(\alpha + \kappa, 0) = 1 - \sum_{m_{ij}^{d,t} = 0}^{4} \left( \begin{array}{c} l - \alpha - \kappa \\ m_{ij}^{d,t} \\ m_{ij}^{d,t} \end{array} \right) \mu_d^{m_{ij}^{d,t}} (1 - \mu_d)^{l - \alpha - \kappa - m_{ij}^{d,t}}$$

and

$$M_d(\alpha + \kappa, x) = 1 - \sum_{m_{ij}^{d,t} = 0}^{4} \left( \begin{array}{c} l - \alpha - \kappa \\ m_{ij}^{d,t} \\ m_{ij}^{d,t} \end{array} \right) \mu_d^{m_{ij}^{d,t}} (1 - \mu_d)^{l - \alpha - \kappa - m_{ij}^{d,t}}$$

$$+ \sum_{y=5-x}^{4} \left( \begin{array}{c} l - \alpha - \kappa \\ y \\ y \end{array} \right) \mu_d^{y} (1 - \mu_d)^{l - \alpha - \kappa - y} \text{ for } 1 \leq x \leq 4.$$

For the $j$th cytoplasmic genome in the $i$th cell, we add the $m_{ij}^{b,t}$ new beneficial mutations to the existing $\alpha$ beneficial mutations and the $m_{ij}^{d,t}$ new deleterious mutations to the existing $\kappa$ beneficial mutations according to

$$C_{G}^{t,\tau_2}(i, j) = C_{G}^{t,\tau_1}(i, j) + m_{ij}^{b,t},$$

and

$$C_{G}^{t,\tau_2}(i + N, j) = C_{G}^{t,\tau_1}(i + N, j) + m_{ij}^{d,t}.$$
B.5.2 Selection

The next life cycle stage is selection. Here, each cell is assigned a fitness value based on the number of beneficial and deleterious substitutions they carry. The number of beneficial substitutions carried by the $i$th cell is given by $\beta(i)$ and the number of deleterious substitutions carried by the $i$th cell is $\rho(i)$, where

$$\beta(i) = \sum_{j=1}^{n} C_{G}^{\text{U},\tau_{2}}(i, j),$$

and

$$\rho(i) = \sum_{j=1}^{n} C_{G}^{\text{D},\tau_{2}}(i + N, j).$$

We examine concave down fitness (decreasing) for deleterious substitutions, and concave up, linear, and concave down fitness functions for beneficial substitutions. The fitness of the $i$th cell, which carries $\beta(i)$ beneficial substitutions and $\rho(i)$ deleterious substitutions under the concave up fitness function for beneficial substitutions is given by

$$\omega_{u,\text{bd}}(\beta(i), \rho(i)) = 1 + s_{b} \left( \frac{\beta(i)}{n\gamma} - 1 \right) - s_{d} \left( \frac{\rho(i)}{n\gamma} \right)^{2},$$

its fitness under the linear fitness function for beneficial substitutions is given by

$$\omega_{l,\text{bd}}(\beta(i), \rho(i)) = 1 + s_{b} \left( \frac{\beta(i)}{n\gamma} - 1 \right) - s_{d} \left( \frac{\rho(i)}{n\gamma} \right)^{2},$$

and its fitness under the concave down fitness function for beneficial substitutions is given by

$$\omega_{d,\text{bd}}(\beta(i), \rho(i)) = 1 + s_{b} \left( \sqrt{\frac{\beta(i)}{n\gamma}} - 1 \right) - s_{d} \left( \frac{\rho(i)}{n\gamma} \right)^{2},$$

where $n$ is the number of cytoplasmic genomes in each cell, $s_{b}$ is the beneficial selection coefficient and $s_{d}$ is the deleterious selection coefficient. To maintain consistency with the first two models, $\gamma$ is set to the same value as in the model with beneficial mutations only. If $\omega_{f,\text{bd}}(\beta(i), \rho(i)) < 1$ we set $\omega_{f,\text{bd}}(\beta(i), \rho(i)) = 0$ (as fitness cannot be negative).

The 1-by-$N$ vector $S_{G}^{t}$ stores the normalized fitness of the population, where $S_{G}^{t}(i)$ gives the relative fitness of the $i$th cell in the population. To generate $S_{G}^{t}$, we
first generate a temporary 1-by-$N$ matrix, $S'_G$ where $S'_G(i) = \omega_{f,bd}(\beta(i), \rho(i))$. To generate $S'_G$, we normalize this vector according to

$$S'_G(i) = \frac{S'_G(i)}{\sum_{z=1}^{N} S'_G(z)}.$$ 

Finally, we use the probabilities in $S'_G$ to generate $N$ new cells for the population, using the process described in section B.3.2.

### B.5.3 Meiosis

#### B.5.3.1 Biparental inheritance

To choose which cytoplasmic genomes are passed on, for each mating type we generate a matrix, $H_g^t(i, d) = Y$ with $N$ rows and $b$ columns populated with uniformly random positive integers ($Y$) in the set $\{1, 2, \ldots n\}$, where $g$ represents the nuclear allele of the gamete and when inheritance is biparental takes values in $\{B_A, B_a\}$. $H_g^t(i, d) = Y$ denotes that the $d$th genome chosen for the new gamete of type $g$ is derived from the $Y$th cytoplasmic genome of the $i$th cell. Sampling is with replacement and gametes are stored in a matrix, $G_{t,\tau}^{\delta,\tau_4}$ which has $2N$ rows and $b$ columns. Since the beneficial substitutions of the $d$th genome of the $i$th gamete is stored in $G_{g}^{t,\tau_4}(i,d)$ and the deleterious substitutions of the $d$th genome of the $i$th gamete are stored in $G_{g}^{t,\tau_4}(i+N,d)$, both must segregate together. $G_{BA}^{t,\tau_4}(i,d)$ is produced by

$$G_{BA}^{t,\tau_4}(i+d) = C_{BA}^{t,\tau_3}(i, H_{BA}^t(i, d) = Y),$$

and

$$G_{BA}^{t,\tau_4}(i+N,d) = C_{BA}^{t,\tau_3}(i+N, H_{BA}^t(i, d) = Y).$$

$G_{Ba}^{t,\tau_4}(i,d)$ is produced by

$$G_{Ba}^{t,\tau_4}(i,d) = C_{Ba}^{t,\tau_3}(i, H_{Ba}^t(i, d) = Y),$$

and

$$G_{Ba}^{t,\tau_4}(i+N,d) = C_{Ba}^{t,\tau_3}(i+N, H_{Ba}^t(i, d) = Y).$$
B.5.3.2 Uniparental inheritance

When inheritance is uniparental, $G_{tA}^{t_2 \tau_4}(i, d)$ is produced by

$$G_{tA}^{t_2 \tau_4}(i, d) = C_{U}^{t_2 \tau_3}(i, H_{tA}^{t}(i, d) = Y),$$

and

$$G_{tA}^{t_2 \tau_4}(i + N, d) = C_{U}^{t_2 \tau_3}(i + N, H_{tA}^{t}(i, d) = Y).$$

$G_{ut}^{t_2 \tau_4}(i, d)$ is produced by

$$G_{ut}^{t_2 \tau_4}(i, d) = C_{U}^{t_2 \tau_3}(i, H_{ut}^{t}(i, d) = Y),$$

and

$$G_{ut}^{t_2 \tau_4}(i + N, d) = C_{U}^{t_2 \tau_3}(i + N, H_{ut}^{t}(i, d) = Y).$$

B.5.4 Random mating

B.5.4.1 Biparental inheritance

Biparental inheritance simply combines the cytoplasmic genomes of both gametes. For each of the $B_A$- and $B_a$-carrying gametes, we generate a 1-by-$N$ vector, $T_B^t(i) = Z$ that contains a random ordering (without replacement) of positive integers from the set $\{1, 2, ..., N\}$. We use these vectors to pair up gametes according to

$$C_B^{t+1, \tau_1}(i, *) = G_{BA}^{t_2 \tau_4}(T_B^t(i) = Z, *) \| G_B^{t_2 \tau_4}(T_B^t(i) = Z, *),$$

and

$$C_B^{t+1, \tau_1}(i + N, *) = G_{BA}^{t_2 \tau_4}((T_B^t(i) = Z) + N, *) \| G_B^{t_2 \tau_4}((T_B^t(i) = Z) + N, *).$$

$\|$ indicates that the two vectors are concatenated. $C_B^{t+1, \tau_1}$ is a temporary matrix (to be replaced by $C_B^{t+1, \tau_1}$), which contains $2b$ columns (representing $2b$ genomes). Since $2b < n$ when we impose a transmission bottleneck, the final step for each cell is to sample $n$ genomes with replacement from these $2b$ genomes. This sampling follows the same approach as described in meiosis, but now instead of choosing $b$ genomes from a cell with $n$ genomes, we choose $n$ genomes from a cell with $2b$ genomes. We generate a matrix, $F_B^t(i, j) = Q$ with $N$ rows and $n$ columns populated with uniformly random
positive integers sampled with replacement from the set \( \{1, 2, \ldots, 2b\} \), which we use to sample the new genomes according to

\[
C^{t+1, \tau_1}_B(i, j) = C^{t+1, \tau_1}_B(i, F_B'(i, j) = Q),
\]

and

\[
C^{t+1, \tau_1}_B(i + N, j) = C^{t+1, \tau_1}_B(i + N, F_B'(i, j) = Q).
\]

### B.5.4.2 Uniparental inheritance

Under uniparental inheritance, only the gamete with mating type \( A \) passes on its cytoplasmic genomes. Thus, to pair up gametes we only need to generate one 1-by-\( N \) vector, \( T_{UA}^t(i) = Z \) that contains a random ordering (without replacement) of positive integers in the set \( \{1, 2, \ldots, N\} \), giving

\[
C^{t+1, \tau_1}_U(i, *) = G^{t, \tau_4}_{UA}(T_{UA}^t(i) = Z, *) ,
\]

and

\[
C^{t+1, \tau_1}_U(i + N, *) = G^{t, \tau_4}_{UA}((T_{UA}^t(i) = Z) + N, *).
\]

Now \( C^{t+1, \tau_1}_U(i, *) \) only contains \( b \) columns (representing \( b \) genomes), so for each cell we sample \( n \) genomes with replacement from these \( b \) genomes. We generate a matrix, \( F_U^t(i, j) = Q \) with \( N \) rows and \( n \) columns populated with uniformly random positive integers sampled with replacement from the set \( \{1, 2, \ldots, b\} \). We use this to sample the new genomes according to

\[
C^{t+1, \tau_1}_U(i, j) = C^{t+1, \tau_1}_U(i, F_U'(i, j) = Q),
\]

and

\[
C^{t+1, \tau_1}_U(i + N, j) = C^{t+1, \tau_1}_U(i + N, F_U'(i, j) = Q).
\]
B.6 Free-living genome model

In our model of free-living genomes, we store the population of cells in a 1-by-$N \times n$ vector (or 1-by-$2(N \times n)$ vector for the model with both beneficial and deleterious mutations). In the model that only considers beneficial mutations, $C^{t,\tau}(i) = \alpha$ indicates that the $i$th free-living cell carries $\alpha$ substitutions. In the model that only considers deleterious mutations, $C^{t,\tau}(i) = \kappa$ indicates that the $i$th free-living cell carries $\kappa$ substitutions. In the model that considers both beneficial and deleterious mutations, $C^{t,\tau}(i) = \alpha$ and $C^{t,\tau}(i + Nn) = \kappa$ indicates that the $i$th free-living cell carries $\alpha$ beneficial and $\kappa$ deleterious substitutions.

There are two stages to the free-living life cycle: mutation and selection. Mutation proceeds in the same way as it does in the model of cytoplasmic genomes (but now the uniformly random number $r^{t}_{i}$ is matched to the $i$th cell in the population). Selection now acts directly on free-living genomes rather than on host cells that carry multiple cytoplasmic genomes. For example, the fitness of the $i$th cell ($C^{t,\tau}(i) = \alpha$) under the linear fitness function in the model that considers beneficial mutations only is

$$\omega_{l,b}(C^{t,\tau}(i)) = 1 + s_{b} \left[ \frac{\alpha}{n^{\gamma}} - 1 \right].$$

Based on these fitness values, we generate a 1-by-$Nn$ normalized fitness vector, which we use to choose $Nn$ cells by multinomial sampling for the new population, as described in section B.3.2.
Bibliography


Appendix C

Chapter 4: Supplementary Material

C.1 Supplementary tables
Table C.1: The frequency of uniparental inheritance at equilibrium when $n = 20$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_b$</td>
<td>fitness</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.1</td>
<td>linear</td>
</tr>
<tr>
<td>0.1</td>
<td>linear</td>
</tr>
<tr>
<td>0.1</td>
<td>linear</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.5</td>
<td>linear</td>
</tr>
<tr>
<td>0.5</td>
<td>linear</td>
</tr>
<tr>
<td>0.5</td>
<td>linear</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. up</td>
</tr>
<tr>
<td>0.9</td>
<td>linear</td>
</tr>
<tr>
<td>0.9</td>
<td>linear</td>
</tr>
<tr>
<td>0.9</td>
<td>linear</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. down</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. down</td>
</tr>
</tbody>
</table>

---

1 The frequency of uniparental inheritance at equilibrium  
2 The proportion of mitochondrial haplotypes that carry a beneficial mutation at equilibrium  
3 The number of generations to reach equilibrium
Table C.2: The frequency of uniparental inheritance at equilibrium when $n = 50$

<table>
<thead>
<tr>
<th>$s_b$</th>
<th>fitness</th>
<th>$\mu$</th>
<th>UPI freq.</th>
<th>prop. bene.</th>
<th>gen.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>conc. up</td>
<td>$10^{-3}$</td>
<td>0.25</td>
<td>1</td>
<td>338</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. up</td>
<td>$10^{-6}$</td>
<td>0.27</td>
<td>1</td>
<td>399</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. up</td>
<td>$10^{-9}$</td>
<td>0.27</td>
<td>1</td>
<td>465</td>
</tr>
<tr>
<td>0.1</td>
<td>linear</td>
<td>$10^{-3}$</td>
<td>0.17</td>
<td>1</td>
<td>393</td>
</tr>
<tr>
<td>0.1</td>
<td>linear</td>
<td>$10^{-6}$</td>
<td>0.18</td>
<td>1</td>
<td>432</td>
</tr>
<tr>
<td>0.1</td>
<td>linear</td>
<td>$10^{-9}$</td>
<td>0.18</td>
<td>1</td>
<td>497</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. down</td>
<td>$10^{-3}$</td>
<td>0.10</td>
<td>1</td>
<td>594</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. down</td>
<td>$10^{-6}$</td>
<td>0.11</td>
<td>1</td>
<td>596</td>
</tr>
<tr>
<td>0.1</td>
<td>conc. down</td>
<td>$10^{-9}$</td>
<td>0.11</td>
<td>1</td>
<td>661</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. up</td>
<td>$10^{-3}$</td>
<td>0.71</td>
<td>1</td>
<td>107</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. up</td>
<td>$10^{-6}$</td>
<td>0.72</td>
<td>1</td>
<td>107</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. up</td>
<td>$10^{-9}$</td>
<td>0.72</td>
<td>1</td>
<td>108</td>
</tr>
<tr>
<td>0.5</td>
<td>linear</td>
<td>$10^{-3}$</td>
<td>0.56</td>
<td>1</td>
<td>145</td>
</tr>
<tr>
<td>0.5</td>
<td>linear</td>
<td>$10^{-6}$</td>
<td>0.57</td>
<td>1</td>
<td>144</td>
</tr>
<tr>
<td>0.5</td>
<td>linear</td>
<td>$10^{-9}$</td>
<td>0.57</td>
<td>1</td>
<td>144</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. down</td>
<td>$10^{-3}$</td>
<td>0.39</td>
<td>1</td>
<td>198</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. down</td>
<td>$10^{-6}$</td>
<td>0.41</td>
<td>1</td>
<td>197</td>
</tr>
<tr>
<td>0.5</td>
<td>conc. down</td>
<td>$10^{-9}$</td>
<td>0.41</td>
<td>1</td>
<td>197</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. up</td>
<td>$10^{-3}$</td>
<td>0.91</td>
<td>1</td>
<td>72</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. up</td>
<td>$10^{-6}$</td>
<td>0.91</td>
<td>1</td>
<td>72</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. up</td>
<td>$10^{-9}$</td>
<td>0.91</td>
<td>1</td>
<td>72</td>
</tr>
<tr>
<td>0.9</td>
<td>linear</td>
<td>$10^{-3}$</td>
<td>0.78</td>
<td>1</td>
<td>101</td>
</tr>
<tr>
<td>0.9</td>
<td>linear</td>
<td>$10^{-6}$</td>
<td>0.79</td>
<td>1</td>
<td>101</td>
</tr>
<tr>
<td>0.9</td>
<td>linear</td>
<td>$10^{-9}$</td>
<td>0.78</td>
<td>1</td>
<td>101</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. down</td>
<td>$10^{-3}$</td>
<td>0.53</td>
<td>1</td>
<td>144</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. down</td>
<td>$10^{-6}$</td>
<td>0.60</td>
<td>1</td>
<td>142</td>
</tr>
<tr>
<td>0.9</td>
<td>conc. down</td>
<td>$10^{-9}$</td>
<td>0.62</td>
<td>1</td>
<td>142</td>
</tr>
</tbody>
</table>

1 The frequency of uniparental inheritance at equilibrium
2 The proportion of mitochondrial haplotypes that carry a beneficial mutation at equilibrium
3 The number of generations to reach equilibrium
Table C.3: The frequency of uniparental inheritance at equilibrium when $n = 100$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
<th>UPI freq.(^1)</th>
<th>prop. bene.(^2)</th>
<th>gen.(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_b$ conc. up $10^{-3}$</td>
<td></td>
<td>0.26</td>
<td>1</td>
<td>403</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-6}$</td>
<td></td>
<td>0.27</td>
<td>1</td>
<td>422</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-9}$</td>
<td></td>
<td>0.27</td>
<td>1</td>
<td>486</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-3}$</td>
<td></td>
<td>0.17</td>
<td>1</td>
<td>505</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-6}$</td>
<td></td>
<td>0.19</td>
<td>1</td>
<td>501</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-9}$</td>
<td></td>
<td>0.19</td>
<td>1</td>
<td>521</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-3}$</td>
<td></td>
<td>0.10</td>
<td>1</td>
<td>669</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-6}$</td>
<td></td>
<td>0.12</td>
<td>1</td>
<td>651</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-9}$</td>
<td></td>
<td>0.12</td>
<td>1</td>
<td>687</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-3}$</td>
<td></td>
<td>0.71</td>
<td>1</td>
<td>149</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-6}$</td>
<td></td>
<td>0.72</td>
<td>1</td>
<td>149</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-9}$</td>
<td></td>
<td>0.72</td>
<td>1</td>
<td>149</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-3}$</td>
<td></td>
<td>0.57</td>
<td>1</td>
<td>204</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-6}$</td>
<td></td>
<td>0.58</td>
<td>1</td>
<td>203</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-9}$</td>
<td></td>
<td>0.58</td>
<td>1</td>
<td>203</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-3}$</td>
<td></td>
<td>0.40</td>
<td>1</td>
<td>281</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-6}$</td>
<td></td>
<td>0.42</td>
<td>1</td>
<td>280</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-9}$</td>
<td></td>
<td>0.42</td>
<td>1</td>
<td>280</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-3}$</td>
<td></td>
<td>0.91</td>
<td>1</td>
<td>102</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-6}$</td>
<td></td>
<td>0.91</td>
<td>1</td>
<td>102</td>
</tr>
<tr>
<td>$s_b$ conc. up $10^{-9}$</td>
<td></td>
<td>0.91</td>
<td>1</td>
<td>102</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-3}$</td>
<td></td>
<td>0.81</td>
<td>1</td>
<td>143</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-6}$</td>
<td></td>
<td>0.81</td>
<td>1</td>
<td>143</td>
</tr>
<tr>
<td>$s_b$ linear $10^{-9}$</td>
<td></td>
<td>0.79</td>
<td>1</td>
<td>144</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-3}$</td>
<td></td>
<td>0.59</td>
<td>1</td>
<td>204</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-6}$</td>
<td></td>
<td>0.65</td>
<td>1</td>
<td>202</td>
</tr>
<tr>
<td>$s_b$ conc. down $10^{-9}$</td>
<td></td>
<td>0.65</td>
<td>1</td>
<td>202</td>
</tr>
</tbody>
</table>

\(^1\) The frequency of uniparental inheritance at equilibrium
\(^2\) The proportion of mitochondrial haplotypes that carry a beneficial mutation at equilibrium
\(^3\) The number of generations to reach equilibrium
### Table C.4: The frequency of uniparental inheritance at equilibrium when \( n = 200 \)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>( s_b )</td>
<td></td>
</tr>
<tr>
<td>0.1 conc. up</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.1 conc. up</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.1 conc. up</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.1 linear</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.1 linear</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.1 linear</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.1 conc. down</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.1 conc. down</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.1 conc. down</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.5 conc. up</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.5 conc. up</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.5 conc. up</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.5 linear</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.5 linear</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.5 linear</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.5 conc. down</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.5 conc. down</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.5 conc. down</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.9 conc. up</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.9 conc. up</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.9 conc. up</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.9 linear</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.9 linear</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.9 linear</td>
<td>10(^{-9})</td>
</tr>
<tr>
<td>0.9 conc. down</td>
<td>10(^{-3})</td>
</tr>
<tr>
<td>0.9 conc. down</td>
<td>10(^{-6})</td>
</tr>
<tr>
<td>0.9 conc. down</td>
<td>10(^{-9})</td>
</tr>
</tbody>
</table>

\(^1\) The frequency of uniparental inheritance at equilibrium

\(^2\) The proportion of mitochondrial haplotypes that carry a beneficial mutation at equilibrium

\(^3\) The number of generations to reach equilibrium
Table C.5: Comparison of the single-locus and two-locus models ($n = 20$ and $\mu = 10^{-6}$)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UPI freq. single-locus</td>
</tr>
<tr>
<td>$s_b$</td>
<td>Up</td>
</tr>
<tr>
<td>0.1 conc. up</td>
<td>0.26</td>
</tr>
<tr>
<td>0.1 linear</td>
<td>0.18</td>
</tr>
<tr>
<td>0.1 conc. down</td>
<td>0.11</td>
</tr>
<tr>
<td>0.5 conc. up</td>
<td>0.71</td>
</tr>
<tr>
<td>0.5 linear</td>
<td>0.56</td>
</tr>
<tr>
<td>0.5 conc. down</td>
<td>0.40</td>
</tr>
<tr>
<td>0.9 conc. up</td>
<td>0.90</td>
</tr>
<tr>
<td>0.9 linear</td>
<td>0.77</td>
</tr>
<tr>
<td>0.9 conc. down</td>
<td>0.51</td>
</tr>
</tbody>
</table>

1. The frequency of uniparental inheritance at equilibrium in the single-locus model
2. The frequency of uniparental inheritance at equilibrium in the two-locus model
Table C.6: The frequency of uniparental inheritance at equilibrium: mating types vs. no mating types ($\mu = 10^{-6}$)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
<th>mating types</th>
<th>no mating types</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$s_b$</td>
<td>fitness</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>conc. up</td>
<td>0.26</td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>linear</td>
<td>0.18</td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>conc. down</td>
<td>0.11</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>conc. up</td>
<td>0.71</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>linear</td>
<td>0.56</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.40</td>
</tr>
<tr>
<td>20</td>
<td>0.9</td>
<td>conc. up</td>
<td>0.90</td>
</tr>
<tr>
<td>20</td>
<td>0.9</td>
<td>linear</td>
<td>0.77</td>
</tr>
<tr>
<td>20</td>
<td>0.9</td>
<td>conc. down</td>
<td>0.51</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>conc. up</td>
<td>0.27</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>linear</td>
<td>0.18</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>conc. down</td>
<td>0.11</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>conc. up</td>
<td>0.72</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>linear</td>
<td>0.57</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.41</td>
</tr>
<tr>
<td>50</td>
<td>0.9</td>
<td>conc. up</td>
<td>0.91</td>
</tr>
<tr>
<td>50</td>
<td>0.9</td>
<td>linear</td>
<td>0.79</td>
</tr>
<tr>
<td>50</td>
<td>0.9</td>
<td>conc. down</td>
<td>0.61</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>conc. up</td>
<td>0.27</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>linear</td>
<td>0.19</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>conc. down</td>
<td>0.12</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>conc. up</td>
<td>0.72</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>linear</td>
<td>0.58</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.42</td>
</tr>
<tr>
<td>100</td>
<td>0.9</td>
<td>conc. up</td>
<td>0.91</td>
</tr>
<tr>
<td>100</td>
<td>0.9</td>
<td>linear</td>
<td>0.81</td>
</tr>
<tr>
<td>100</td>
<td>0.9</td>
<td>conc. down</td>
<td>0.65</td>
</tr>
<tr>
<td>200</td>
<td>0.1</td>
<td>conc. up</td>
<td>0.27</td>
</tr>
<tr>
<td>200</td>
<td>0.1</td>
<td>linear</td>
<td>0.19</td>
</tr>
<tr>
<td>200</td>
<td>0.1</td>
<td>conc. down</td>
<td>0.12</td>
</tr>
<tr>
<td>200</td>
<td>0.5</td>
<td>conc. up</td>
<td>0.73</td>
</tr>
<tr>
<td>200</td>
<td>0.5</td>
<td>linear</td>
<td>0.58</td>
</tr>
<tr>
<td>200</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.43</td>
</tr>
<tr>
<td>200</td>
<td>0.9</td>
<td>conc. up</td>
<td>0.92</td>
</tr>
<tr>
<td>200</td>
<td>0.9</td>
<td>linear</td>
<td>0.82</td>
</tr>
<tr>
<td>200</td>
<td>0.9</td>
<td>conc. down</td>
<td>0.69</td>
</tr>
</tbody>
</table>

1 The fifth column assumes $U \times U$ matings lead to uniparental inheritance while the sixth column assumes $U \times U$ matings lead to biparental inheritance.
Table C.7: The spread of uniparental inheritance under fluctuating selection pressures ($\mu = 10^{-6}$)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
<th>UPI freq.</th>
<th>generations</th>
<th>perturbations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>$s_b$</td>
<td>fitness</td>
<td>UPI freq.</td>
<td>generations</td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>conc. up</td>
<td>1.00</td>
<td>15101</td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>linear</td>
<td>1.00</td>
<td>122000</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.17</td>
<td>1000000</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>conc. up</td>
<td>1.00</td>
<td>8027</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>linear</td>
<td>1.00</td>
<td>21021</td>
</tr>
<tr>
<td>20</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.69</td>
<td>1000000</td>
</tr>
<tr>
<td>20</td>
<td>0.9</td>
<td>conc. up</td>
<td>1.00</td>
<td>5010</td>
</tr>
<tr>
<td>20</td>
<td>0.9</td>
<td>linear</td>
<td>1.00</td>
<td>10008</td>
</tr>
<tr>
<td>20</td>
<td>0.9</td>
<td>conc. down</td>
<td>1.00</td>
<td>104000</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>conc. up</td>
<td>1.00</td>
<td>14154</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>linear</td>
<td>1.00</td>
<td>120000</td>
</tr>
<tr>
<td>50</td>
<td>0.1</td>
<td>conc. down</td>
<td>0.16</td>
<td>1000000</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>conc. up</td>
<td>1.00</td>
<td>8028</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>linear</td>
<td>1.00</td>
<td>20028</td>
</tr>
<tr>
<td>50</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.67</td>
<td>1000000</td>
</tr>
<tr>
<td>50</td>
<td>0.9</td>
<td>conc. up</td>
<td>1.00</td>
<td>5011</td>
</tr>
<tr>
<td>50</td>
<td>0.9</td>
<td>linear</td>
<td>1.00</td>
<td>9010</td>
</tr>
<tr>
<td>50</td>
<td>0.9</td>
<td>conc. down</td>
<td>1.00</td>
<td>96014</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>conc. up</td>
<td>1.00</td>
<td>14142</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>linear</td>
<td>1.00</td>
<td>119000</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>conc. down</td>
<td>0.16</td>
<td>1000000</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>conc. up</td>
<td>1.00</td>
<td>8032</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>linear</td>
<td>1.00</td>
<td>20028</td>
</tr>
<tr>
<td>100</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.66</td>
<td>1000000</td>
</tr>
<tr>
<td>100</td>
<td>0.9</td>
<td>conc. up</td>
<td>1.00</td>
<td>5012</td>
</tr>
<tr>
<td>100</td>
<td>0.9</td>
<td>linear</td>
<td>1.00</td>
<td>9009</td>
</tr>
<tr>
<td>100</td>
<td>0.9</td>
<td>conc. down</td>
<td>1.00</td>
<td>86023</td>
</tr>
<tr>
<td>200</td>
<td>0.1</td>
<td>conc. up</td>
<td>1.00</td>
<td>14151</td>
</tr>
<tr>
<td>200</td>
<td>0.1</td>
<td>linear</td>
<td>1.00</td>
<td>119000</td>
</tr>
<tr>
<td>200</td>
<td>0.1</td>
<td>conc. down</td>
<td>0.16</td>
<td>1000000</td>
</tr>
<tr>
<td>200</td>
<td>0.5</td>
<td>conc. up</td>
<td>1.00</td>
<td>8037</td>
</tr>
<tr>
<td>200</td>
<td>0.5</td>
<td>linear</td>
<td>1.00</td>
<td>20030</td>
</tr>
<tr>
<td>200</td>
<td>0.5</td>
<td>conc. down</td>
<td>0.65</td>
<td>1000000</td>
</tr>
<tr>
<td>200</td>
<td>0.9</td>
<td>conc. up</td>
<td>1.00</td>
<td>5013</td>
</tr>
<tr>
<td>200</td>
<td>0.9</td>
<td>linear</td>
<td>1.00</td>
<td>8012</td>
</tr>
<tr>
<td>200</td>
<td>0.9</td>
<td>conc. down</td>
<td>1.00</td>
<td>78026</td>
</tr>
</tbody>
</table>

1 The frequency of uniparental inheritance at equilibrium
2 The number of generations to reach equilibrium (unless this did not occur within 1,000,000 generations)
3 The number of environmental perturbations that occurred
C.2 Single-locus model assuming mating types

Our model tracks the distribution of cell types through each stage of the life cycle across multiple generations. The redistribution of cell types is based on probability theory, but the model itself is deterministic. We assume that the population is effectively infinite and unaffected by genetic drift, as is regularly assumed in models such as ours [1–3]. Consequently, the probability that a cell takes a particular state equates to the proportion of that cell type in the population.

Diploid cell types are described by the random variable $M^{t,\tau_\alpha} = (i, G)$, where $i$ corresponds to the number of beneficial mitochondria and takes values in $\{0, 1...n\}$, $t$ indicates the generation, and $\tau_\alpha$ indicates the stage of the life cycle. Since cells carry $i$ beneficial mitochondria, the number of wild type mitochondria is given by $n-i$. The nuclear genotype is given by $G$, which takes values in $\{U_1B_2, B_1B_2\}$. Gametes are described by the random variable $M^{t,\tau_\alpha} = (p, g)$, where $p$ is the number of beneficial mitochondria and takes values in $\{0, 1...n/2\}$ and $g$ represents the nuclear allele and takes values in $\{U_1, B_1, B_2\}$. The probability of obtaining a particular diploid cell type is written as $P(M^{t,\tau_\alpha} = (i, G))$, and the probability of obtaining a particular gamete is written as $P(M^{t,\tau_\alpha} = (p, g))$. Since we assume an infinite population, these probabilities are equivalent to the proportion of the population with that particular cell or gamete type. We wrote our model in MATLAB® (version 2015a) and ran simulations to determine the frequency of genotypes and alleles at equilibrium.

C.2.1 Initialization

The starting population contains the $B_1$ allele at 49%, the $B_2$ allele at 50%, and the $U_1$ allele at 1%. All gametes initially carry wild type mitochondria. This population of gametes then enters the first stage of the life cycle: random mating.

C.2.2 Random mating

Gametes with $n/2$ mitochondria randomly mate with the opposite mating type to produce diploid cells containing $n$ mitochondria. In effect, this is random mating in which all matings between the same mating type (i.e. $U_1U_1$, $B_1B_1$, $U_1B_1$ and $B_2B_2$) are lethal, and the only viable genotypes are $U_1B_2$ and $B_1B_2$. 
C.2. Single-locus model assuming mating types

C.2.2 Biparental mating

Consider a biparental mating of a gamete with identity $M^{t,\tau_1} = (p, B_1)$, where $\tau_1$ is the stage of the life cycle in which gametes are present. For this gamete to produce a diploid cell with type $M^{t,\tau_2} = (i, B_1B_2)$, where $\tau_2$ is the stage of the life cycle immediately after random mating, it must mate with a gamete with identity $M^{t,\tau_1} = (i-p, B_2)$. The probability of this mating is $2 \left[ P(M^{t,\tau_1} = (p, B_1)) \cdot P(M^{t,\tau_1} = (i-p, B_2)) \right]$, where the factor of 2 accounts for the two permutations of $B_1$ and $B_2$. We restrict the values of $p$ and $q$ to biologically valid combinations. For the $B_1$ gamete, $0 \leq p \leq n/2$ applies, as the $B_1$ gamete cannot carry negative numbers of beneficial mitochondria, nor can it contain more beneficial mitochondria than the total number of mitochondria in the gamete. Likewise, $0 \leq i-p \leq n/2$ applies to the $B_2$ gamete, which gives $i-n/2 \leq p \leq i$ when rearranged. Valid values for $p$ lie in the range of intersection of these two inequalities, giving $\max(0, i-n/2) \leq p \leq \min(n/2, i)$. The probability of forming any given diploid cell type after random mating is thus given by

$$P(M^{t,\tau_2} = (i, B_1B_2)) = 2 \sum_{p=\max(0, i-n/2)}^{\min(n/2, i)} P(M^{t,\tau_1} = (p, B_1))P(M^{t,\tau_1} = (i-p, B_2)).$$

C.2.2.2 Uniparental mating

Because uniparental matings between $U_1$ and $B_2$ gametes contain mitochondria from $U_1$ alone, $U_1B_2$ cells carry $n/2$ mitochondria immediately after mating. To restore the total complement of $n$ mitochondria, we sample $n/2$ mitochondria with replacement from the $n/2$ mitochondria in the $U_1B_2$ cell, adding the $n/2$ sampled mitochondria to the original set of mitochondria to form a cell with $n$ mitochondria.

For a gamete with identity $M^{t,\tau_1} = (p, U_1)$ to produce a diploid cell with identity $M^{t,\tau_2} = (i, U_1B_2)$, it must sample $n/2$ mitochondria with $i-p$ beneficial mitochondria and $n/2 - (i-p)$ wild type mitochondria. The mitochondrial state of the $B_2$ gamete is irrelevant because its mitochondria are discarded and we will refer to this cell as $M^{t,\tau_1} = (r, B_2)$.

Sampling of the $n/2$ mitochondria follows a binomial distribution, which we denote $T(i-p; n/2, 2p/n)$, where $i-p$ refers to the number of beneficial mitochondria required, $n/2$ refers to the number of mitochondria in the gamete, and $2p/n$ is the probability of drawing a single mutant mitochondrion from a $U_1B_2$ cell with $p$ mutant mitochondria (where $2p/n$ is obtained by rearranging $p/(n/2)$). The probability of
C.2. Single-locus model assuming mating types

sampling \( i - p \) mutant mitochondria (and \( n/2 - (i - p) \) wild type mitochondria) is given by

\[
T(i - p; n/2, 2p/n) = \left( \frac{n/2}{i - p} \right) \left( \frac{2p}{n} \right)^{i-p} \left( 1 - \frac{2p}{n} \right)^{n/2-i-p}.
\] (C.1)

The restrictions on \( p \) and \( i - p \) are the same as those in biparental mating. Since \( U_1 \) will form the same initial \( U_1B_2 \) cell regardless of the \( B_2 \) gamete with which it mates, the probability of producing each type of \( U_1 \) gamete is multiplied by the probability of selecting any \( B_2 \) gamete. The probability of forming a given \( U_1B_2 \) cell after random mating is determined by

\[
P(M^{t,\tau_2} = (i, U_1B_2))
\]

\[
= 2 \left[ \sum_{p = \max(0, i-n/2)}^{\min(n/2, i)} P(M^{t,\tau_1} = (p, U_1)) T(i - p; n/2, 2p/n) \sum_{r=0}^{n/2} P(M^{t,\tau_1} = (r, B_2)) \right].
\]

C.2.2.3 Normalisation

Since mating types are self-incompatible, the sum of all cell types is less than 1 after random mating. To keep the sum of the proportions of the population at 1, we normalise the population after random mating. The normalised population is given by

\[
P(M^{t,\tau_3} = (i, G)) = \frac{P(M^{t,\tau_2} = (i, G))}{k},
\]

where

\[
k = \sum_{i=0}^{n} \left[ P(M^{t,\tau_2} = (i, U_1B_2)) + P(M^{t,\tau_2} = (i, B_1B_2)) \right].
\]

C.2.3 Mutation

The identity of a cell immediately after mutation is \( M^{t,\tau_4} = (i, G) \), (where \( \tau_4 \) is the life cycle stage immediately after mutation). If \( a \) is the number of wild type mitochondria that mutate to beneficial mitochondria, then a cell with identity \( M^{t,\tau_4} = (i, G) \) after mutation must be derived from a pre-mutation cell in state \( M^{t,\tau_3} = (i - a, G) \) (because the pre-mutation cell gains \( a \) beneficial mitochondria to form the post-mutation cell). Similarly, if the post-mutation cell has \( n - i \) wild type mitochondria, then the pre-mutation cell must have \( n - i + a \) wild type mitochondria.
First, we must work out the probability that a cell mutates $a$ of its wild type mitochondria to beneficial mitochondria. We define $Y(a; n - i + a, \mu)$ as the probability that a pre-mutation cell has $a$ mutations in its $n - i + a$ wild type mitochondria, given that each mitochondrion mutates with probability $\mu$. The accumulation of mutations is binomially distributed such that

$$Y(a; n - i + a, \mu) = \binom{n - i + a}{a} \mu^a (1 - \mu)^{n-i}.$$ 

The number of beneficial mutations, $a$, must satisfy $0 \leq a \leq i$, as the post-mutation cell cannot receive a negative number of mutations nor can it receive more than $i$ beneficial mutations. The probability that a pre-mutation cell with identity $M_{\tau}^{i, \gamma_3} = (i - a, G)$ becomes a post-mutation cell with identity $M_{\tau}^{i, \gamma_4} = (i, G)$ is given by

$$M_{\tau}^{i, \gamma_4} = (i, G) = \sum_{a=0}^{i} Y(a; n - i + a, \mu) P(M_{\tau}^{i, \gamma_3} = (i - a, G)).$$

### C.2.4 Selection

The relative fitness of a cell, $\omega_{\gamma}(i)$, is a measure of how likely a cell type is to survive and reproduce, where $\gamma$ denotes the type of fitness function. We assume that beneficial mutations improve cell fitness additively. Thus, fitness monotonically increases as a function of the number of beneficial mitochondria. The relative fitness of a cell with $i$ beneficial mitochondria, where $0 \leq i \leq n$, is given by

$$\omega_{u}(i) = 1 + s_b \left[ \left( \frac{i}{n} \right)^2 - 1 \right],$$

for the concave up fitness function,

$$\omega_{l}(i) = 1 + s_b \left( \frac{i}{n} - 1 \right),$$

for the linear fitness function, and

$$\omega_{d}(i) = 1 + s_b \left[ \sqrt[2]{\frac{i}{n}} - 1 \right],$$

for the concave down fitness function. The proportion of a cell type remaining post-selection is given by

$$P (M_{\tau}^{i, \gamma_5} = (i, G)) = \omega_{\gamma}(i) P (M_{\tau}^{i, \gamma_4} = (i, G)).$$
C.2.5 Meiosis

The cell first duplicates its chromosomes and double its mitochondrial complement (from \(n\) to \(2n\)). A cell with \(2n\) mitochondria then produces gametes with \(n/2\) mitochondria. Meiosis occurs in two steps. First, we sample \(n\) mitochondria with replacement from a cell containing \(n\) mitochondria and add the set of sampled mitochondria to the original set of mitochondria to form a cell containing \(2n\) mitochondria. We let \(M^{t,\tau_5} = (l, 2G)\) represent the cell with doubled mitochondria and nuclear genotype, where \(l\) takes values in \(\{0, 1, \ldots, 2n\}\) and \(2G\) takes values in \(\{U_1 U_2 B_2, B_1 B_2 B_2\}\).

For a cell to contain \(l\) beneficial mitochondria after its mitochondria are duplicated, it must sample \(l - i\) beneficial mitochondria. We denote the probability of sampling mutant mitochondria from \(M^{t,\tau_5} = (i, G)\) as \(F(l - i; n, i/n)\). Sampling follows a binomial distribution such that

\[
F(l - i; n, i/n) = \binom{n}{l - i} \left( \frac{i}{n} \right)^{l - i} \left( 1 - \frac{i}{n} \right)^{n - l + i}.
\]

After doubling of the nuclear genome and mitochondria, cells are given by

\[
P(M^{t,\tau_6} = (l, 2G)) = \sum_{i = \max(0, l-n)}^{\min(l, n)} F(l - i; n, i/n) P(M^{t,\tau_5} = (i, G)).
\]

During the second step of meiosis, the cells with \(2n\) mitochondria produce gametes with \(n/2\) mitochondria. Biologically, this occurs in two steps. In meiosis 1, the homologous chromosomes are pulled apart to produce two haploid cells that contain two identical nuclear alleles (sister chromatids) and \(n\) mitochondria. In meiosis 2, the two cells divide to produce four gametes, each with a single nuclear allele and \(n/2\) mitochondria. Since mitochondria segregate independently of nuclear alleles during cell partitioning, we model this as a single step.

We define \(S(p; 2n, l, n/2)\) to be the probability of obtaining \(p\) mutant mitochondria in \(n/2\) draws from a cell in state \((M^{t,\tau_6} = (l, 2G))\). Here, sampling is without replacement and follows a hypergeometric distribution, giving

\[
S(p; 2n, l, n/2) = \binom{l}{p} \binom{2n-l}{n/2-p} / \binom{2n}{n/2}.
\]  
(C.5)

The gametes produced by meiosis are represented by \(M^{t+1,\tau_1} = (p, g)\). The probability of obtaining a \(U_1\) gamete with \(p\) beneficial mutations is given by
C.2. Single-locus model assuming mating types

\[ P(M^{t+1, \tau_1} = (p, U_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S(p; 2n, l, \frac{n}{2}) P(M^{t, \tau_6} = (l, U_1U_1B_2B_2)) \right], \]

a \( B_1 \) gamete with \( p \) beneficial mutations by

\[ P(M^{t+1, \tau_1} = (p, B_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S(p; 2n, l, \frac{n}{2}) P(M^{t, \tau_6} = (l, B_1B_1B_2B_2)) \right], \]

and a \( B_2 \) gamete with \( p \) beneficial mutations by

\[
P(M^{t+1, \tau_1} = (p, B_2)) \]
\[
= \frac{1}{2} \left[ \sum_{l=0}^{2n} S(p; 2n, l, \frac{n}{2}) P(M^{t, \tau_6} = (l, U_1U_1B_2B_2)) \right] + \frac{1}{2} \left[ \sum_{l=0}^{2n} S(p; 2n, l, \frac{n}{2}) P(M^{t, \tau_6} = (l, B_1B_1B_2B_2)) \right].
\]

The factors of \( 1/2 \) in the above three equations take into account that half of the gametes produced from parent cells with nuclear genotype \( U_1B_2 \) will carry the \( U_1 \) allele and the other half will carry the \( B_2 \) allele (with the same approach applied for gametes produced from parent cells with nuclear genotype \( B_1B_2 \)). Meiosis completes a single generation of the life cycle.

C.2.6 Relative fitness of gametes

Although gametes are not subject to selection in our model, and thus do not technically have fitness values, it is informative to track the relative fitness of gametes throughout the simulation (see Figure 4.3). We define a gamete’s relative fitness as the fitness that a diploid cell would have if it had the same mitochondrial composition as the gamete. Since gametes contain \( n/2 \) mitochondria, while cells carry \( n \) mitochondria, we replace \( n \) with \( n/2 \) in the fitness functions for the cells, where \( 0 \leq i \leq n/2 \).

Equation C.2 (concave up fitness) becomes

\[
\omega^g(i) = 1 + s_b \left[ \left( \frac{i}{n/2} \right)^2 - 1 \right],
\]

Equation C.3 (linear fitness) becomes
C.2. Single-locus model assuming mating types

\[ \omega_i^g(i) = 1 + s_b \left[ \left( \frac{i}{n/2} \right)^2 - 1 \right], \]

and Equation C.4 (concave down fitness) becomes

\[ \omega_i^d(i) = 1 + s_b \left[ \sqrt{\frac{i}{n/2}} - 1 \right]. \]

Once the fitness function is scaled to gametes, we can determine the relative fitness of the \(U_1\), \(B_1\), and \(B_2\) alleles by

\[ \bar{\omega}^U_1 = \frac{\sum_{i=0}^{n/2} P(M^{t,\tau_1} = (i, U_1)) \omega_i^U_1(i)}{\sum_{i=0}^{n/2} P(M^{t,\tau_1} = (i, U_1))}, \]

\[ \bar{\omega}^B_1 = \frac{\sum_{i=0}^{n/2} P(M^{t,\tau_1} = (i, B_1)) \omega_i^B_1(i)}{\sum_{i=0}^{n/2} P(M^{t,\tau_1} = (i, B_1))}, \]

and

\[ \bar{\omega}^B_2 = \frac{\sum_{i=0}^{n/2} P(M^{t,\tau_1} = (i, B_2)) \omega_i^B_2(i)}{\sum_{i=0}^{n/2} P(M^{t,\tau_1} = (i, B_2))}. \]
C.3 Two-locus model

In the two-locus model, there are three types of mitochondrial haplotypes: wild type, one beneficial mutation, and two beneficial mutations. Diploid cell types are described by the random variable $M_{t,\tau} = (i, j, G)$, where $i$ corresponds to the number of mitochondrial haplotypes with one beneficial mutation and takes values in \{0, 1, ..., n\}, and $j$ represents the number of mitochondrial haplotypes with two beneficial mitochondria and takes values in \{0, 1, ..., n - i\}. As in the single-locus model, $t$ indicates the generation, and $\tau$ indicates the stage of the life cycle. The number of wild type mitochondria is $n - i - j$. $G$ indicates the nuclear genotype and takes values in \{U, B_1, B_2\}. Gametes are described by the random variable $M_{t,\tau} = (p, q, g)$, where $p$ is the number of mitochondrial haplotypes with one beneficial mutation and takes values in \{0, 1, ..., n/2\}, and $q$ is the number of mitochondrial haplotypes with two beneficial mutations and takes values in \{0, 1, ..., n/2 - p\}. $g$ represents the nuclear allele and takes values in \{U, B_1, B_2\}. Initialization is the same as the single-locus model.

C.3.1 Random mating

C.3.1.1 Biparental mating

For a gamete with identity $M_{t,\tau_1} = (p, q, B_1)$ to undergo biparental mating and produce a diploid cell of identity $M_{t,\tau_2} = (i, j, B_1B_2)$, it must mate with a gamete of type $M_{t,\tau_1} = (i - p, j - q, B_2)$. Next, we must restrict values of $p$, $q$, $i$, and $j$ to biologically valid combinations.

First, $0 \leq p \leq n/2$, as the $B_1$ gamete cannot carry negative numbers of mitochondria with 1 beneficial mutation, nor can it contain more mitochondria with one beneficial mutation than the total number of mitochondria in the gamete. Likewise, $0 \leq i - p \leq n/2$ for the $B_2$ gamete, which gives $i - n/2 \leq p \leq i$ when rearranged. Valid values for $p$ lie in the range of intersection of these two inequalities, giving $\max(0, i - n/2) \leq p \leq \min(n/2, i)$.

The first restriction for $q$ is $0 \leq q \leq n/2 - p$ because the $B_1$ gamete already contains $p$ mitochondria with one beneficial mutation and cannot contain more than $n/2$ mitochondria. There is another restriction, $0 \leq j - q \leq n/2 - (i - p)$, because the $B_2$ gamete cannot contain more than $n/2$ mitochondria and already contains $i - p$ mitochondria with 1 beneficial mitochondria. This inequality gives $i + j - p - n/2 \leq q \leq j$ when
C.3. Two-locus model

rearranged. The intersection of these inequalities gives \(0, i + j - p - n/2 \leq q \leq \min(n/2 - p, j)\). Thus,

\[
P \left( M^{l, r_2} = (i, j, B_1 B_2) \right) = 2 \left[ \sum_{p = \max(0, i - n/2)}^{\min(n/2, i)} \sum_{q = \max(0, i + j - p - n/2)}^{\min(n/2 - p, j)} P \left( M^{l, r_1} = (p, q, B_1) \right) P \left( M^{l, r_1} = (i - p, j - q, B_2) \right) \right].
\]

C.3.2 Uniparental mating

For a gamete with identity \( M^{l, r_1} = (p, q, U_1) \) to produce a diploid cell with identity \( M^{l, r_2} = (i, j, U_1 B_2) \), it must sample \( n/2 \) mitochondria containing \( i - p \) mitochondria with one beneficial mutation and \( j - q \) mitochondria with two beneficial mutations. The mitochondrial state of the \( B_2 \) gamete is irrelevant because its mitochondria are discarded and we will refer to this cell as \( M^{l, r_1} = (r, s, B_2) \).

Sampling of the \( n/2 \) mitochondria follows a multinomial distribution, which we denote \( T (i - p, j - q; n/2, 2p/n, 2q/n) \), where \( i - p \) and \( j - q \) refer to the number of mitochondria with one and two beneficial mutations that need to be sampled, and \( n/2 \) is the number of mitochondria being sampled. \( 2p/n \) and \( 2q/n \) refer to the probabilities of drawing mitochondria with one beneficial mutation and two beneficial mutations respectively from a \( U_1 B_2 \) cell that contains \( p \) mitochondria with one beneficial mutation and \( q \) mitochondria with two beneficial mutations (where \( 2p/n \) is obtained by rearranging \( p/(n/2) \) and \( 2q/n \) is obtained by rearranging \( q/(n/2) \)).

The probability of sampling \( i - p \) and \( j - q \) mitochondria with one and two beneficial mutations respectively (and \( n/2 - (i - p) - (j - q) \) wild type mitochondria) is given by

\[
T (i - p, j - q; n/2, 2p/n, 2q/n) = \frac{n!}{(i - p)! (j - q)! (n/2 - (i - p) - (j - q))!} \times \left( \frac{2p}{n} \right)^{i-p} \left( \frac{2q}{n} \right)^{j-q} \left( \frac{2(n/2 - p - q)}{n} \right)^{n/2 - (i-p) - (j-q)}.
\]

The restrictions on \( p, q, i, \) and \( j \) are the same as those in biparental mating. Because \( U_1 \) will form the same initial \( U_1 B_2 \) cell regardless of the \( B_2 \) gamete with which it mates, the probability of each \( U_1 \) gamete is multiplied by the probability of selecting each \( B_2 \) gamete. The probability of forming a given \( U_1 B_2 \) cell after random mating is
C.3. Two-locus model

\[ P \left( M^{t,\tau_2} = (i, j, U_1 B_2) \right) = 2 \left[ \sum_{p=\max(0, i-n/2)}^{\min(n/2, i)} \sum_{q=\max(0, i+j-p-n/2)}^{\min(n/2-p, j)} P \left( M^{t,\tau_1} = (p, q, U_1) \right) \times T \left( i - p, j - q; \frac{n}{2}, \frac{2p}{n}, \frac{2q}{n} \right) \sum_{r=0}^{n/2} \sum_{s=0}^{n/2-r} P \left( M^{t,\tau_1} = (r, s, B_2) \right) \right]. \]

C.3.2.1 Normalisation

To keep the sum of the proportions of the population at 1 after random mating, we normalize the population. The normalized population is given by

\[ P \left( M^{t,\tau_3} = (i, j, G) \right) = \frac{P \left( M^{t,\tau_2} = (i, j, G) \right)}{\beta}, \]

where

\[ \beta = \sum_{i=0}^{n} \sum_{j=0}^{n-i} \left[ P \left( M^{t,\tau_2} = (i, j, U_1 B_2) \right) + P \left( M^{t,\tau_2} = (i, j, B_1 B_2) \right) \right]. \]

C.3.3 Mutation

Wild type mitochondria can mutate into mitochondria with one beneficial mutation, and mitochondria with one beneficial mutation can mutate into mitochondria with two beneficial mutations. The probability of both types of mutation is given by \( \mu \). We ignore back mutation (e.g. mutation from a mitochondrion with one beneficial mutation to a wild type mitochondrion) and ignore the probability that a wild type mitochondrion accumulates two beneficial mutations in a single generation.

If a post-mutation cell is given by \( M^{t,\tau_4} = (i, j, G) \), then it must be derived from a pre-mutation cell with identity \( M^{t,\tau_3} = (i - a + b, j - b, G) \), where \( a \) is the number of mutations in wild type mitochondria (which become mitochondria with one beneficial mutation) and \( b \) is the number of mutations in mitochondria with one beneficial mutation (which become mitochondria with two beneficial mutations).

We let \( Y(a, n - i - j + a, \mu) \) represent the probability of \( a \) mutations in the \( n - i - j + a \) wild type mitochondria in the pre-mutation cell. The accumulation of \( a \) mutations follows a binomial distribution, giving

\[ Y(a, n - i - j + a, \mu) = \binom{n - i - j + a}{a} \mu^a (1 - \mu)^{n-i-j}. \]
We let $Y(b, i - a + b, \mu)$ represent the probability of $b$ mutations in the $i - a + b$ mitochondria with one beneficial mutation in the pre-mutation cell. The accumulation of $b$ mutations follows a binomial distribution, giving

$$Y(b, i - a + b, \mu) = \binom{i - a + b}{b} \mu^b (1 - \mu)^{i-a}.$$ 

Valid values of $a$ are given by $0 \leq a \leq i$ because $a$ cannot be negative, and the number of mutations in wild type mitochondria (in the pre-mutation cell) cannot exceed the number of mitochondria with one beneficial mutation (in the post-mutation cell). Likewise, valid values of $b$ are given by $0 \leq b \leq j$ because the number of mutations in mitochondria with one beneficial mutation (in the pre-mutation cell) cannot exceed the number of mitochondria with two beneficial mutations (in the post-mutation cell). Thus, the probability of a post-mutation cell is given by

$$P(M_{t,\tau_4} = (i, j, G)) = \sum_{a=0}^{i} \sum_{b=0}^{j} Y(a; n - i - j + a, \mu) Y(b, i - a + b, \mu) M_{t,\tau_3} = (i - a + b, j - b, G).$$

### C.3.4 Selection

As with the single-locus model, we examine three fitness functions. A mitochondrion with two beneficial mutations contributes twice the fitness benefits of a mitochondrion with one beneficial mutation. The relative fitness of a cell, $\omega_\gamma(i, j)$, carrying $i$ mitochondria with one beneficial mutation and $j$ mitochondria with two beneficial mutations is given by

$$\omega_u(i, j) = 1 + s_b \left[ \left( \frac{\frac{1}{2}i + j}{n} \right)^2 - 1 \right],$$

for the concave up fitness function,

$$\omega_l(i, j) = 1 + s_b \left[ \left( \frac{\frac{1}{2}i + j}{n} \right) - 1 \right],$$

for the linear fitness function, and

$$\omega_d(i, j) = 1 + s_b \left[ \sqrt{\frac{\frac{1}{2}i + j}{n}} - 1 \right],$$
C.3. Two-locus model

for the concave down fitness function. The proportion of a cell type remaining post-selection is given by

\[ P(\mathbf{M}^{l,\tau_5} = (i, j, G)) = \omega_{\gamma}(i, j) P(\mathbf{M}^{l,\tau_4} = (i, j, G)). \]

C.3.5 Meiosis

As in the single-locus model, we sample \( n \) mitochondria with replacement from a cell containing \( n \) mitochondria and add the set of sampled mitochondria to the original set of mitochondria to form a cell containing \( 2n \) mitochondria. We let \( \mathbf{M}^{l,\tau_6} = (l, m, 2G) \) represent the cell with doubled mitochondria and nuclear genotype, where \( l \) takes values in \( \{0, 1, \ldots, 2n\} \), \( m \) takes values in \( \{0, 1, \ldots, 2n - l\} \), and \( 2G \) takes values in \( \{U_1U_1B_2B_2, B_1B_1B_2B_2\} \).

The probability of sampling \( l - i \) mitochondria with one beneficial mutation and \( m - j \) mitochondria with two beneficial mutations from \( \mathbf{M}^{l,\tau_5} = (i, j, G) \) is \( F(l - i, m - j; n, i/n, j/n) \). Sampling follows a multinomial distribution, giving

\[
F(l - i, m - j; n, i/n, j/n) = \frac{n!}{(l - i)! (m - j)! (n - (l - i) - (m - j))!} \left( \frac{i}{n} \right)^{l-i} \left( \frac{j}{n} \right)^{m-j} \left( \frac{n - i - j}{n} \right)^{n-(l-i)-(m-j)}.
\]

We obtain \( \mathbf{M}^{l,\tau_6} = (l, m, 2G) \) by

\[
P(\mathbf{M}^{l,\tau_6} = (l, m, 2G)) = \sum_{i=\max(0, l-n)}^{\min(l, n)} \sum_{j=\max(0, m+l-n-i)}^{\min(m, n-i)} F(l - i, m - j; n, i/n, j/n) P(\mathbf{M}^{l,\tau_5} = (i, j, G)).
\]

During the second step of meiosis, the cells with \( 2n \) mitochondria produce gametes with \( n/2 \) mitochondria. We define \( S(p, q; 2n, l, m, n/2) \) to be the probability of obtaining \( p \) mitochondria with one beneficial mutation and \( q \) mitochondria with two beneficial mutations in \( n/2 \) draws from the \( \mathbf{M}^{l,\tau_6} = (l, m, 2G) \) cell carrying \( l \) mitochondria with one beneficial mutation and \( m \) mitochondria with two beneficial mutations. Here, sampling is without replacement and follows a multivariate hypergeometric distribution, giving

\[
S(p, q; 2n, l, m, \frac{n}{2}) = \frac{\binom{l}{p} \binom{m}{q} \binom{2n-l-m}{n/2-p-q}}{\binom{2n}{n/2}}.
\]
Gametes produced by meiosis are represented by $M^{t+1, r_1} = (p, q, g)$. The probability of obtaining a particular gamete is

$$P (M^{t+1, r_1} = (p, q, U_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S(p, q; 2n, l, m, \frac{n}{2}) P(M^{t, r_6} = (l, m, U_1U_1B_2B_2)) \right],$$

$$P (M^{t+1, r_1} = (p, q, B_1)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S(p, q; 2n, l, m, \frac{n}{2}) P(M^{t, r_6} = (l, m, B_1B_1B_2B_2)) \right] ,$$

and

$$P (M^{t+1, r_1} = (p, q, B_2)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S(p, q; 2n, l, m, \frac{n}{2}) P(M^{t, r_6} = (l, m, U_1U_1B_2B_2)) \right] ,$$

$$+ \frac{1}{2} \left[ \sum_{l=0}^{2n} \sum_{m=0}^{2n-l} S(p, q; 2n, l, m, \frac{n}{2}) P(M^{t, r_6} = (l, m, B_1B_1B_2B_2)) \right] .$$
C.4 No mating types model

As in the version with mating types, cells are described by the random variable $M^{t,\tau_\alpha} = (i, G)$, but now $G$ takes values in \{UU, UB, BB\}. Gametes are described by the random variable $M^{t,\tau_\alpha} = (p, g)$, where $g$ takes values in \{U, B\}.

C.4.1 Initialization

The starting proportions are 99% for $B$ and 1% for $U$. All other details of initialization remain the same as the general model.

C.4.2 Random mating

C.4.2.1 Biparental mating (BB cells)

The probability of producing a BB cell type after random mating is given by

$$ P \left( M^{t,\tau_2} = (i, BB) \right) = \min \left( \frac{n}{2}, i \right) \sum_{p = \max(0, i - n/2)} P \left( M^{t,\tau_1} = (p, B) \right) P \left( M^{t,\tau_1} = (i - p, B) \right). $$

C.4.2.2 Biparental mating (UU cells)

The probability of producing a UU cell, when we assume that $U \times U$ matings are biparental, is given by

$$ P \left( M^{t,\tau_2} = (i, UU) \right) = \min \left( \frac{n}{2}, i \right) \sum_{p = \max(0, i - n/2)} P \left( M^{t,\tau_1} = (p, U) \right) P \left( M^{t,\tau_1} = (i - p, U) \right). $$

C.4.2.3 Uniparental mating (UB cells)

Using Equation C.1, the probability of forming a UB cell is given by

$$ P \left( M^{t,\tau_2} = (i, UB) \right) = 2 \left[ \min \left( \frac{n}{2}, i \right) \sum_{p = \max(0, i - n/2)} P \left( M^{t,\tau_1} = (p, U) \right) T \left( i - p; \frac{n}{2}, \frac{2p}{n} \right) \sum_{r=0}^{n/2} P \left( M^{t,\tau_1} = (r, B) \right) \right]. $$
C.4.2.4 Uniparental mating (UU cells)

Again using Equation C.1, the probability of producing a UU cell, when we assume that $U \times U$ matings are uniparental, is

$$P\left(M^{t,\tau_2} = (i, UU)\right) = \sum_{p=\max(0, i-n/2)}^{\min(n/2, i)} P\left(M^{t,\tau_1} = (p, U)\right) T\left(i - p; \frac{n}{2}, \frac{2p}{n}\right) \sum_{r=0}^{n/2} P\left(M^{t,\tau_1} = (r, U)\right).$$

C.4.2.5 Normalization

Although all gametes can now randomly mate, the sum of the proportions of cell types will still deviate from 1 due to selection. Thus, once a generation we normalise the population to reset the sum of the proportions of the population to 1. The normalized population is given by

$$P\left(M^{t,\tau_3} = (i, G)\right) = \frac{P\left(M^{t,\tau_2} = (i, G)\right)}{k},$$

where

$$k = \sum_{i=0}^{n} \left[ P\left(M^{t,\tau_2} = (i, UU)\right) + P\left(M^{t,\tau_2} = (i, UB)\right) + P\left(M^{t,\tau_2} = (i, BB)\right) \right].$$

The mutation and selection stages are the same as the single-locus model that assumes mating types (although there are now three genotypes instead of two).

C.4.3 Meiosis

The process by which cells with identity $M^{t,\tau_5} = (i, G)$ become cells with identity $PM^{t,\tau_6} = (l, 2G)$ does not change. Thus, using Equation C.5 the probability of producing a $U$ gamete is

$$P\left(M^{t+1,\tau_1} = (p, U)\right) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, UUBB)\right) \right]$$

$$+ \sum_{l=0}^{2n} S\left(p; 2n, l, \frac{n}{2}\right) P\left(M^{t,\tau_6} = (l, UUU)\right),$$

while the probability of producing a $B$ gamete is
\[ P (M^{t+1, \tau_1} = (p, B)) = \frac{1}{2} \left[ \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P \left( M^{t, \tau_6} = (l, UUBB) \right) \right] + \sum_{l=0}^{2n} S \left( p; 2n, l, \frac{n}{2} \right) P \left( M^{t, \tau_6} = (l, BBBB) \right). \]
C.5 Fluctuating environment model

In this model, selection pressure changes every 1000 generations, switching the mitochondrial type that is selectively advantageous. The fluctuating environment model does not differ from the single-locus model with regard to random mating, selection, or meiosis.

C.5.1 Mutation

Since both mitochondrial types can be beneficial—depending on the direction of selection—we let mutation be bidirectional. Thus, a cell can gain beneficial mitochondria (mutations in its maladaptive haplotypes), and gain maladapted mitochondria (mutations in its adaptive haplotypes). We assume that the rate of mutation is the same in both directions (given by $\mu$).

As with the single-locus model with mating types, the identity of a cell immediately after mutation is $M^{t,\tau_4} = (i, G)$, where $i$ refers to the number of beneficial (or adapted) mitochondria carried by a cell. If $a$ is the number of beneficial mitochondria that mutate to maladapted mitochondria, and $b$ is the number of maladapted mitochondria that mutate to beneficial mitochondria, then a cell with identity $M^{t,\tau_4} = (i, G)$ after mutation must be derived from a pre-mutation cell in state $M^{t,\tau_3} = (i - a + b, G)$.

First, we must work out the probability that a cell mutates $a$ of its wild type mitochondria to beneficial mitochondria. We define $Y(a; n - i + a - b, \mu)$ as the probability that a pre-mutation cell gains $a$ beneficial mitochondria in its $n - i + a - b$ maladapted mitochondria, given that each mitochondrion mutates with probability $\mu$. The accumulation of mutations is binomially distributed such that

$$Y(a; n - i + a - b, \mu) = \binom{n - i + a - b}{a} \mu^a (1 - \mu)^{n - i - b}.$$  

Likewise, we define $Y(b; i - a + b, \mu)$ to be the probability that a pre-mutation cell acquires $b$ maladapted mitochondria in its $i - a + b$ adapted mitochondria, given a mutation rate of $\mu$. This probability is given by

$$Y(b; i - a + b, \mu) = \binom{i - a + b}{b} \mu^b (1 - \mu)^{i - a}.$$  

The number of beneficial mutations, $a$, must satisfy $0 \leq a \leq i$, and the number of maladapted mutations, $b$, must satisfy $0 \leq b \leq n - i$. The post-mutation population is given by
\[ P \left( M_{t, \tau^*} = (i, G) \right) \]
\[ = \sum_{a=0}^{i} \sum_{b=0}^{n-i} Y(a; n - i + a - b, \mu) Y(b; i - a + b, \mu) P \left( M_{t, \tau^*} = (i - a + b, G) \right). \]

### C.5.2 Environmental fluctuation

Every 1000 generations, the beneficial haplotype becomes maladapted. At the same time, the previously maladapted haplotype becomes beneficial. Thus, when the number of generations, \( t \), is a multiple of 1000, there is a special life cycle stage \( \tau^* \) that switches the beneficial mitochondria with the maladapted mitochondria in the gametes generated by cells in generation \( t - 1 \). At generation 1000, for example, the following transition occurs

\[ P \left( M^{1000, \tau^*} = (p, g) \right) = P \left( M^{1000, \tau_3} = (n/2 - p, g) \right), \]

where \( p \) is the number of beneficial mitochondria before the environmental fluctuation (and \( n/2 - p \) the number of maladapted mitochondria). The gametes with identity \( M^{1000, \tau^*} = (p, g) \) then undergo random mating to form cells with identity \( M^{1000, \tau_2} = (i, G) \), and the life cycle continues as normal until generation 2000, when another environmental fluctuation occurs (and so on).
Bibliography


Appendix D

Chapter 5: Supplementary Material

D.1 Supplementary figures
Figure D.1: **Effect of horizontal transmission on invasion of endosymbionts.** Note that the y-axes differ between plots. **A.** Mixed transmission with a horizontal transmission rate of $\mu = 10^{-2}$. **B.** Mixed transmission with $\mu = 10^{-4}$. The number of generations before endosymbionts become fixed. Simulations were terminated once 50,000 generations had passed without fixation of the endosymbiont (when endosymbionts were costly, they never invaded in the arthropod model). Parameters: $N = 1000$ and $K = 20$. For “High growth rate”, $r = 2$, and for “Low growth rate”, $r = 0.1$. $s_d = 0.5$ for “Large cost”, $s_d = 0.05$ for “Small cost”, $s = 0$ for “Neutral”, $s_b = 0.05$ for “Small benefit”, and $s_b = 0.5$ for “Large benefit”. The inoculum size is $b = 5$. 

![Supplementary figures 258](image-url)
Figure D.2: The timing of germline formation affects endosymbiont growth, selection on hosts, and discordance between soma and germline. The timing of germline formation affects the size of the germline precursor relative to the zygote (the earlier the formation, the larger the relative size of the germline precursor). In turn, the size of the germline precursor affects the distribution of endosymbiont loads between hosts. Here we examine three scenarios: (1) early germline formation ($p_s = 0.5$); (2) intermediate germline formation ($p_s = 0.25$); and (3) late germline formation $p_s = 0.1$. The figure shows the distribution of endosymbionts in the germline of hosts, starting from zygotes that carry $K/2$ endosymbionts (averaged over 10,000,000 simulations). A shows the distribution after segregation of the soma and germline, while B shows the distribution after segregation and endosymbiont growth. Parameters: $K = 20$, $s_b = 0.5$, and $s_d = 0.5$.

A. A later formation of the germline increases the variation in endosymbiont load between hosts. A later formation of the germline leads to a less negative covariance in endosymbiont load of germline and soma (early formation: Pearson’s $r = -0.53$; intermediate formation: Pearson’s $r = -0.46$; late formation: Pearson’s $r = -0.32$). However, a later germline formation increases the magnitude of the discordance between soma and germline ($d_{s,g}$; see Figure 5.6) in endosymbiont load between soma and germline (early formation: $d_{s,g} = 0.22$; intermediate formation: $d_{s,g} = 0.25$; late formation: $d_{s,g} = 0.33$). B. Earlier formation of the germline leads to faster endosymbiont growth (early formation: $\bar{r}_e = 0.37$; intermediate formation: $\bar{r}_i = 0.29$; late formation: $\bar{r}_l = 0.12$; see Figure 5.3 for a description of $\bar{r}_m$). We then separately apply positive and negative selection on the distribution in B (not depicted). The earlier the germline forms, the weaker the selection on hosts (early formation (beneficial): $\tilde{\omega}_{c,b} = 0.02$; early formation (deleterious): $\tilde{\omega}_{c,d} = -0.03$; intermediate formation (beneficial): $\tilde{\omega}_{i,b} = 0.05$; intermediate formation (deleterious): $\tilde{\omega}_{i,d} = -0.06$; late formation (beneficial): $\tilde{\omega}_{l,b} = 0.15$; late formation (deleterious): $\tilde{\omega}_{l,d} = -0.17$; see Figure 5.3 for a description of $\tilde{\omega}_{m,s}$).
Figure D.3: **Effect of horizontal transmission on reproductive manipulation by endosymbionts.** We compare cytoplasmic incompatibility (CI) and feminization (F) with no reproductive manipulation (none). **A.** The number of generations before endosymbionts become fixed under mixed transmission with a rate of horizontal transmission of $\mu = 10^{-2}$. Note that the y-axes differ between plots (numbers above bars indicate the generations). Simulations were terminated once 50,000 generations had passed without fixation of the endosymbiont. **B.** Mixed transmission with $\mu = 10^{-4}$. Parameters: $N = 1000$ and $K = 20$. For “High growth rate”, $r = 2$, and for “Low growth rate”, $r = 0.1$. $s_d = 0.5$ for “Large cost”, $s_d = 0.05$ for “Small cost”, $s = 0$ for “Neutral”, $s_b = 0.05$ for “Small benefit”, and $s_b = 0.5$ for “Large benefit”. The inoculum size is $b = 5$. 

![Figure D.3](image.png)
### Table D.1: Effect of varying $K$, $r$, and $s_b/s_d$ on endosymbiont growth and selection on hosts in the arthropod and protist models. See Figure 5.3 for a description of $\bar{r}_m$, $\omega_{m,b}$, and $\omega_{m,d}$.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$K$</td>
<td>$r$</td>
<td>$s_b/s_d$</td>
<td>$\bar{r}_m$</td>
<td>$\omega_{m,b}$</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
<td>0.1</td>
<td>0.1</td>
<td>0.027</td>
<td>0.0132</td>
</tr>
<tr>
<td>Protist</td>
<td>20</td>
<td>0.1</td>
<td>0.1</td>
<td>0.049</td>
<td>0.0013</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
<td>0.1</td>
<td>0.5</td>
<td>0.026</td>
<td>0.0808</td>
</tr>
<tr>
<td>Protist</td>
<td>20</td>
<td>0.1</td>
<td>0.5</td>
<td>0.049</td>
<td>0.0101</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
<td>1</td>
<td>0.1</td>
<td>0.258</td>
<td>0.0112</td>
</tr>
<tr>
<td>Protist</td>
<td>20</td>
<td>1</td>
<td>0.1</td>
<td>0.483</td>
<td>0.0025</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
<td>1</td>
<td>0.5</td>
<td>0.260</td>
<td>0.0617</td>
</tr>
<tr>
<td>Protist</td>
<td>20</td>
<td>1</td>
<td>0.5</td>
<td>0.483</td>
<td>0.0137</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
<td>0.1</td>
<td>0.1</td>
<td>0.045</td>
<td>0.0030</td>
</tr>
<tr>
<td>Protist</td>
<td>100</td>
<td>0.1</td>
<td>0.1</td>
<td>0.050</td>
<td>0.0005</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
<td>0.1</td>
<td>0.5</td>
<td>0.045</td>
<td>0.0165</td>
</tr>
<tr>
<td>Protist</td>
<td>100</td>
<td>0.1</td>
<td>0.5</td>
<td>0.050</td>
<td>0.0017</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
<td>1</td>
<td>0.1</td>
<td>0.446</td>
<td>0.0017</td>
</tr>
<tr>
<td>Protist</td>
<td>100</td>
<td>1</td>
<td>0.1</td>
<td>0.497</td>
<td>0.0008</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
<td>1</td>
<td>0.5</td>
<td>0.447</td>
<td>0.0099</td>
</tr>
<tr>
<td>Protist</td>
<td>100</td>
<td>1</td>
<td>0.5</td>
<td>0.497</td>
<td>0.0029</td>
</tr>
<tr>
<td>Arthropod</td>
<td>1000</td>
<td>0.1</td>
<td>0.1</td>
<td>0.049</td>
<td>0.0000</td>
</tr>
<tr>
<td>Protist</td>
<td>1000</td>
<td>0.1</td>
<td>0.1</td>
<td>0.050</td>
<td>0.0001</td>
</tr>
<tr>
<td>Arthropod</td>
<td>1000</td>
<td>0.1</td>
<td>0.5</td>
<td>0.050</td>
<td>0.0017</td>
</tr>
<tr>
<td>Protist</td>
<td>1000</td>
<td>0.1</td>
<td>0.5</td>
<td>0.050</td>
<td>0.0002</td>
</tr>
<tr>
<td>Arthropod</td>
<td>1000</td>
<td>1</td>
<td>0.1</td>
<td>0.495</td>
<td>0.0003</td>
</tr>
<tr>
<td>Protist</td>
<td>1000</td>
<td>1</td>
<td>0.1</td>
<td>0.500</td>
<td>0.0000</td>
</tr>
<tr>
<td>Arthropod</td>
<td>1000</td>
<td>1</td>
<td>0.5</td>
<td>0.495</td>
<td>0.0008</td>
</tr>
<tr>
<td>Protist</td>
<td>1000</td>
<td>1</td>
<td>0.5</td>
<td>0.500</td>
<td>0.0003</td>
</tr>
</tbody>
</table>
Table D.2: Effect of varying $K$, $r$, and $s_b/s_d$ on between-host variance in endosymbiont load, endosymbiont growth and selection on hosts in the models that remove oogamy, multicellularity, and germline. See Figure 5.3 for a description of $\bar{r}_m$, $\bar{\omega}_{m,b}$, and $\bar{\omega}_{m,d}$, and section D.4 for a description of the different models. $\sigma_g^2$ is the variance in endosymbiont load after growth (preceding selection).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>$K$</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
</tr>
<tr>
<td>No oogamy</td>
<td>20</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>20</td>
</tr>
<tr>
<td>No germline</td>
<td>20</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
</tr>
<tr>
<td>No oogamy</td>
<td>20</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>20</td>
</tr>
<tr>
<td>No germline</td>
<td>20</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
</tr>
<tr>
<td>No oogamy</td>
<td>20</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>20</td>
</tr>
<tr>
<td>No germline</td>
<td>20</td>
</tr>
<tr>
<td>Arthropod</td>
<td>20</td>
</tr>
<tr>
<td>No oogamy</td>
<td>20</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>20</td>
</tr>
<tr>
<td>No germline</td>
<td>20</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
</tr>
<tr>
<td>No oogamy</td>
<td>100</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>100</td>
</tr>
<tr>
<td>No germline</td>
<td>100</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
</tr>
<tr>
<td>No oogamy</td>
<td>100</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>100</td>
</tr>
<tr>
<td>No germline</td>
<td>100</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
</tr>
<tr>
<td>No oogamy</td>
<td>100</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>100</td>
</tr>
<tr>
<td>No germline</td>
<td>100</td>
</tr>
<tr>
<td>Arthropod</td>
<td>100</td>
</tr>
<tr>
<td>No oogamy</td>
<td>100</td>
</tr>
<tr>
<td>No multicellularity</td>
<td>100</td>
</tr>
<tr>
<td>No germline</td>
<td>100</td>
</tr>
</tbody>
</table>
D.3 Variation in endosymbiont load slows growth

Here we show that life cycles which induce higher variance in endosymbiont load between hosts slow endosymbiont growth compared to life cycles that induce lower variance between hosts. The number of endosymbionts after the endosymbiont growth phase is given by $e_c r (1 - e_c K)$, where $e_c = C^{c_{r_3}}(i)$ denotes that the $i$th host carries $e_c$ endosymbionts. As $r$ is a constant, growth is proportional to $e_c (1 - e_c K)$.

Letting $e_c = x$, below we show that the average value of the function $f(x) = x(1 - x K)$ over the uniform interval $[x - C, x + C]$ is lower than the value of the function at $x$, where $C$ is a non-zero constant. The average value of a function $f(x)$ over the interval $[a, b]$ is given by

$$f_{\text{ave}} = \frac{1}{b - a} \int_a^b f(x) \, dx.$$ 

Thus,

$$f_{\text{ave}} = \frac{1}{x + C - (x - C)} \int_{x - C}^{x + C} \left( x - \frac{x^2 K}{2} \right) \, dx$$

$$= \frac{1}{2C} \left[ \frac{x^2}{2} - \frac{x^3}{3K} \right]_{x-C}^{x+C}$$

$$= \frac{1}{2C} \left( \left( \frac{(x+C)^2}{2} - \frac{(x+C)^3}{3K} \right) - \left( \frac{(x-C)^2}{2} - \frac{(x-C)^3}{3K} \right) \right)$$

$$= \frac{1}{2C} \left( \frac{x^2 + 2Cx + C^2}{2} - \frac{(x+C)(x^2 + 2Cx + C^2)}{3K} \right.$$

$$+ \left. \frac{(x^2 - 2Cx + C^2)}{2} + \frac{(x+C)(x^2 - 2Cx + C^2)}{3K} \right)$$

$$= \frac{1}{2C} \left( \frac{x^2 + 2Cx + C^2}{2} - \frac{x^3 + 2Cx^2 + Cx^2 + 2C^2x + C^3}{3K} \right.$$

$$- \left. \frac{x^3 - 2Cx^2 + C^2x + x^2C - 2C^2x + C^3}{3K} \right)$$

$$= \frac{1}{2C} \left( \frac{4Cx}{2} + \frac{(-6Cx^2 - 2C^3)}{3K} \right)$$

$$= \frac{1}{2C} \left( 2Cx - \frac{2C^3}{K} - \frac{2C^3}{3K} \right)$$
D.3. Variation in endosymbiont load slows growth

\[ \frac{1}{2} \left( 2x - \frac{2x^2}{K} - \frac{2C^2}{3K} \right) \]

\[ = x - \frac{x^2}{K} - \frac{C^2}{3K} \]

\[ = 1 \left( x - \frac{x}{K} \right) - \frac{C^2}{3K} \]

\[ = f(x) - \frac{C^2}{3K} \]

Since \( \frac{C^2}{3K} \) is always positive (and increases with the magnitude of \( C \)), \( f_{ave} \) over the interval \([x - C, x + C]\) is always less than \( f(x) \). For the model, this means that a life cycle with high variance in endosymbiont load (i.e. larger values of \( C \)) from one generation to the next (e.g. arthropods) reduces the growth of endosymbionts compared to one that produces less variance in endosymbiont load across generations (e.g. protists).
D.4 Removing oogamy, multicellularity, and germline

In Figure 5.5 we alter the arthropod model by removing either oogamy, multicellularity, or the soma-germline separation to see how each affects within-host growth of endosymbionts and selection on hosts. In panels A and D, we remove oogamy so that endosymbionts are transmitted biparentally. We now assume that each gamete contributes equal amounts of cytoplasm (as in species with isogamy). In panels B and E, we remove multicellularity. Here we remove the parts of the model that simulate the variation involved in producing a tissue. We also remove the within-tissue variation that we simulate in the germline. Since the soma-germline separation remains, each is now comprised of a single cell. To do this, we alter two parts of the arthropod model. First, we sample the soma precursor from 50% (i.e. $K/2$ pieces of cytoplasm) of the zygote and the germline precursor from the other 50%. (Since we are removing multicellularity, there is no cell division during development; we assume that the zygote splits evenly to form the soma and germline.) To restore each cell to its full size, we sample another 50% with replacement from each precursor and add the sampled 50% to that already in the precursor. This is in contrast the arthropod model, in which we sample the entire tissue with replacement from the precursor in the arthropod model (we assume that forming an entire tissue from a precursor cell will involve more variance than forming a single cell from a precursor cell). Second, to produce eggs, we sample without replacement. Contrast this to the arthropod model, in which we sample with replacement; in the arthropod model, this simulated within-germline variation so we remove it now that the germline is no longer multicellular. In panels C and F, we remove the separation of soma and germline. The soma precursor is now formed from the entire zygote and gametes are generated directly from the soma. Since the soma remains multicellular, we sample a new soma with replacement from the soma precursor as in the arthropod model.
D.5 Protist model

In the model of single-cell protists, there is no distinction between soma and germline. We assume that mating is governed by nuclear-encoded self-incompatible mating types rather than sexes. We assume there are two mating types, $A$ and $a$, and that mating compatibility is determined by the gametic allele. Thus, a gamete carrying the $A$ allele can only mate with a gamete carrying the $a$ allele and vice versa. The only possible genotype for diploid cells is thus $Aa$ and each of these cells produces a gamete with the $A$ mating type allele and another gamete with the $a$ mating type allele.

The population has $N$ hosts, which we store in a 1-by-$N$ vector $C_{t \tau \zeta}$. $C_{t \tau \zeta}(i) = e_c$ indicates that the $i$th host in the population carries $e_c$ endosymbionts in generation $t$. $\tau \zeta$ denotes the stage of the life cycle. $i$ takes values in $\{0, 1, ..., N\}$ and $e_c$ takes values in $\{0, 1, ..., K\}$, where $K$ is the carrying capacity of the host cell. Endosymbionts replicate within the host with rate $r$, and for simplicity we let endosymbionts replicate once per host generation. An endosymbiont can have a beneficial, deleterious, or neutral effect on its host. We assume that the fitness of a host is proportional to its endosymbiont load, and that the effect of an endosymbiont on its host is linear and additive (i.e. each endosymbiont increases or decreases the fitness of its host by the same magnitude). Below we describe each model in detail.

D.5.1 Mixed transmission

In this version of the protist model, we initialize the population with zero endosymbionts. In the first stage of the life cycle, horizontal transmission, each cell has a probability of receiving $b$ endosymbionts via horizontal transmission (e.g. through contact with a free-living population of bacteria). In the second stage of the life cycle, endosymbiont growth, endosymbionts replicate within the host. In the third stage, selection, each host is assigned a fitness based on the endosymbionts it carries. The population is then reformed, and a host’s survival is proportional to its fitness. In the fourth stage, meiosis, each host produces one gamete with mating type allele $A$ and one gamete with mating type allele $a$ (recall that all hosts have the nuclear genotype $Aa$). In the final stage, mating, gametes carrying the $A$ mating type allele and gametes carrying the $a$ mating type allele randomly mate, reforming the population of diploid cells. We record the number of generations until the endosymbionts reach fixation or until 50,000 generations have passed without fixation (we consider fixation to have occurred when all hosts carry at least $K/2$ endosymbionts).
D.5.1.1 Horizontal transmission

Each host has a probability, $\mu$, of picking up endosymbionts from an external source. Each infected host gains $b$ endosymbionts. To determine which hosts will receive endosymbionts, we generate a 1-by-$N$ vector, $I^t$ of uniformly random numbers between 0 and 1 using the `runif` function in R. Thus,

$$C^{t,\tau_2}(i) = C^{t,\tau_1}(i) + b \text{ if } I^t(i) < \mu,$$

$$C^{t,\tau_2}(i) = C^{t,\tau_1}(i) \text{ if } I^t(i) \geq \mu.$$  

To ensure that hosts do not exceed the carrying capacity, $K$, we set all hosts in which $C^{t,\tau_3}(i) > K$ equal to $K$.

D.5.1.2 Endosymbiont growth

Since hosts have a carrying capacity for endosymbionts, we assume growth is logistic and depends on both the growth rate ($r$; a fixed parameter) and the number of endosymbionts in the host cell ($e_c$). For each host, we calculate a weighted growth rate, $r_{\tau'}$, according to

$$r_{\tau'}(i) = r \left(1 - \frac{e_c}{K}\right),$$

where $e_c = C^{t,\tau_2}(i)$. The expected number of endosymbionts after the growth phase in the host given by $C^{t,\tau_2}(i) = e_c$ is $r_{\tau'}(i)e_c$. Finally, we pass the `rpois` function the expected number of endosymbionts for each host to generate a random Poisson-distributed number of endosymbionts for each $C^{t,\tau_3}$ host. To ensure that hosts do not exceed the carrying capacity, $K$, we set all hosts in which $C^{t,\tau_3}(i) > K$ equal to $K$.

D.5.1.3 Selection

The fitness of a host depends solely on the number of endosymbionts it carries. Endosymbionts can have a beneficial effect, a harmful effect, or no effect on the fitness of the host. Beneficial endosymbionts have a selection coefficient given by $s_b$, while deleterious endosymbionts have a selection coefficient given by $s_d$. The function that determines the fitness of a host carrying beneficial endosymbionts is

$$\omega(e_c) = 1 + s_b \left(\frac{e_c}{K}\right) - s_b.$$  

Thus, when endosymbionts are beneficial, a host that carries zero endosymbionts will have a fitness of $1 - s_b$ while a host that carries $K$ endosymbionts will have a fitness of 1. The function that determines the fitness of a host carrying $e_c$ deleterious endosymbionts is
Thus, when endosymbionts are deleterious, a host that carries zero endosymbionts ($e_c = 0$) will have a fitness of 1 while a host that carries $K$ endosymbionts will have a fitness of $1 - s_d$. When endosymbionts have no effect on host fitness (i.e. $s = 0$), all hosts in the population have a fitness of 1.

The first step in selection is to obtain the fitness of each host ($\omega(C_{t,\tau_2}(i) = e_c)$), which is stored in a 1-by-$N$ vector $\mathbf{W}_t$. Next, $\mathbf{W}_t$ is normalized as the vector $\mathbf{W}_t'$, which sums to 1 and is generated by

$$\mathbf{W}_t'(i) = \frac{\mathbf{W}_t(i)}{\sum_{x=1}^{N} \mathbf{W}_t'(x)}.$$

We pass $\mathbf{W}_t(i)$ to the `rmultinom` function, generating a 1-by-$N$ vector $\mathbf{Y}_t$ that contains $N$ multinomially-distributed “survivors”. $\mathbf{Y}_t(i) = y$ indicates that the $i$th host leaves $y$ copies of itself. Since survivors are multinomially-distributed, each host can leave zero, one, or multiple copies of itself. We form $C_{t,\tau_3}$ from $\mathbf{Y}_t$ and $C_{t,\tau_2}$. For example, if $\mathbf{Y}_t(i) = 2$, then $C_{t,\tau_3}$ would contain two copies of $C_{t,\tau_2}(i)$.

### D.5.1.4 Meiosis

Each host (which has mating type genotype $Aa$) produces two gametes, one of which carries mating type allele $A$ and one with mating type allele $a$. Cytoplasm is equally divided between the two gametes. For the purposes of apportioning endosymbionts between gametes, we divide the host cell up into $K$ pieces of cytoplasm (where $K$ is the carrying capacity). The probability that a piece carries a endosymbiont is thus $e_c/K$ (endosymbionts carried by the cell divided by the number of cytoplasm pieces). From the diploid host with $K$ pieces of cytoplasm, we sample two gametes with $K/2$ pieces of cytoplasm and randomly assign a mating type allele to one gamete ($A$ or $a$) and assign the remaining allele to the other gamete. We sample without replacement using the `rhyper` function. For the first gamete, we use the `rhyper` arguments $nn=N$, $m=C_{t,\tau_3}$, $n=K - C_{t,\tau_3}$, $k=K/2$. The second gamete then inherits the remaining endosymbionts. Gametes are stored in the 1-by-$N$ vector $\mathbf{M}_{t,\tau_4}$, where $m$ represents the mating type allele and takes values in $\{A, a\}$. $\mathbf{M}_{t,\tau_4}(i) = e_g$ indicates the number of endosymbionts carried by the $i$th gamete with mating type $m$, where $e_g$ is the number of endosymbionts carried by a gamete and takes values in $\{0, 1...K/2\}$.
D.6. Arthropod model

D.5.1.5 Mating

In the mating phase, we randomly pair up $M^{t,\tau_4}_A$ gametes with $M^{t,\tau_4}_a$ gametes. The number of endosymbionts carried by the resulting host is the sum of the endosymbionts carried by the two gametes. For each gamete type, we use the `sample` function to generate a randomly-ordered 1-by-$N$ vector of integers from 1 to $N$ without replacement, which we call $Z^t_m$. We reform the diploid population according to

$$C^{t+1,\tau_1}(i) = M^{t,\tau_4}_A(Z^t_A(i)) + M^{t,\tau_4}_a(Z^t_a(i)).$$

D.5.2 Vertical transmission

In this version of the model, we initialize the population with $N - 1$ hosts with no endosymbionts and 1 host with $b$ endosymbionts. Aside from the initial infection of this 1 host, no other horizontal transmission takes place. A simulation stops either when all endosymbionts are lost or the endosymbiont spreads to the point that each host carries $\geq K/2$ endosymbionts (fixation). For each set of conditions, we ran 10,000 simulations and recorded the proportion in which the endosymbiont becomes fixed. The life cycle is the same as in section D.5.1 except that there is now no horizontal transmission phase.

D.6 Arthropod model

In the model of a multicellular arthropod with oogamy, there is now a distinction between soma and germline. Hosts are either male or female, and sex is determined by a XX/XY- or XX/XO-like system. Unlike the protist model, now each sex produces one type of specialized gamete (males produce sperm and females produce eggs).

As before, the population has $N$ hosts. To characterize each host, we need to know three pieces of information: the number of endosymbionts in the soma, the number of endosymbionts in the germline, and the sex of the host. We store this information in a 2-by-$N$ matrix

$$C^{t,\tau_\zeta}_S = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1N} \\ c_{21} & c_{22} & \cdots & c_{2N} \end{bmatrix}.$$  

$(C^{t,\tau_\zeta}_S(1, j) = c_{1j}) = e_s$ represents the number of endosymbionts ($e_s$) in the soma of the $j$th host in life cycle stage $\tau_\zeta$ in generation $t$. The $j$th host has sex $S$, which takes values in $\{M, F\}$ (male or female). Likewise, $(C^{t,\tau_\zeta}_S(2, j) = c_{2j}) = e_g$ indicates the number of endosymbionts in the germline of the $j$th host. Both $e_s$ and $e_g$ take values
in \( \{0, 1, \ldots, K\} \), where \( K \) is the carrying capacity of the host’s cells. \( j \) takes values in \( \{0, 1, \ldots, N\} \). We will use the terminology \( C_S^{t, \tau_1}(\bullet, j) = c_{xj} \) to refer to the \( j \)th individual. Unlike in the protist model, in the arthropod model we allow endosymbionts to distort the reproductive process of their host. We thus further divide the arthropod models into three classes: (1) no reproductive distortion; (2) cytoplasmic incompatibility; and (3) feminization.

D.6.1 Mixed transmission

In this version of the arthropod model, we initialize the population with zero endosymbionts. In the first stage of the life cycle, horizontal transmission, the single-celled zygote can pick up endosymbionts from an external source. In the second stage, soma/germline division, the zygote develops into a multicellular host with a distinct soma and germline. In the third stage, endosymbiont growth, endosymbionts replicate within a host’s soma and germline. In the fourth stage, selection, each host is assigned a fitness based on the endosymbiont load in its soma (the germline does not affect a host’s fitness). These fitness values are used to reform the population. In the fifth stage, meiosis, females produce eggs from their germline and males produce sperm from their germline. In the final stage, mating, eggs and sperm randomly pair up, reforming the population of diploid zygotes. We will first describe the model in which there is no reproductive distortion and then describe the models in which endosymbionts can cause cytoplasmic incompatibility and feminization.

D.6.1.1 Horizontal transmission

In the first stage, all hosts exist as zygotes. Zygotes are a single cell and thus do not have a distinct soma and germline. We represent a zygote as the 1-by-\( N \) vector \( Z_S^{t, \tau_1} \), where \( Z_S^{t, \tau_1}(j) = e_z \) indicates the number of endosymbionts carried by the \( j \)th zygote. Each zygote has a probability, \( \mu \), of becoming infected from an external source, and each infected cell gains \( b \) endosymbionts. To determine which hosts become infected, we generate a 1-by-\( N \) vector, \( I^t \) of uniformly random numbers between 0 and 1 using the \text{runif} function. Thus,

\[
Z_S^{t, \tau_2}(j) = \begin{cases} 
Z_S^{t, \tau_1}(j) + b & \text{if } I^t(j) < \mu, \\
Z_S^{t, \tau_1}(j) & \text{if } I^t(j) \geq \mu.
\end{cases}
\]

To ensure that no zygote carries more endosymbionts than the carrying capacity of a cell, \( K \), for all cells in which \( Z_S^{t, \tau_2}(j) > K \), we set equal to \( K \).
D.6.1.2 Soma-germline division

In this phase, the zygote develops into a multicellular individual with a distinct soma and germline. Rather than explicitly model all the cell divisions involved in the development of different tissues, we simplify the process of soma-germline division. Our simplified procedure is designed to capture several important features of multicellular development and soma-germline division. Specifically, the number of endosymbionts carried by the soma can differ from the number of endosymbionts carried by the germline, depending on the density of endosymbionts in the cytoplasm that becomes the precursor to each tissue type. Furthermore, the density of endosymbionts in the soma and germline can differ from that of the zygote, due to stochastic sampling effects. We follow two tissues: the soma and the germline. We do not take into account within-tissue variation (i.e. we assume that all somatic cells carry the same number of endosymbionts and that all germline cells carry the same number of endosymbionts). This has two implications. First, it means that we assume organism fitness depends on the mean endosymbiont load in the soma (i.e. ignoring possible epistatic interactions between different somatic tissues). Second, it means that we underestimate the variation in endosymbiont load in gametes produced by a host (since different oocytes in the germline would, in reality, vary in endosymbiont load). To help account for this latter limitation, during female meiosis, we sample an endosymbiont load for eggs with replacement from the germline (note that we sample gametes without replacement in the protist model). This introduces variance in the number of endosymbionts carried by an egg from the same individual in much the same way as if we allowed within-germline variance in oocytes and sampled without replacement from different oocytes to produce eggs.

For each zygote, we designate a proportion, $p_s$, to be the precursor of the germline. In most cases, $p_s = 0.2$, which is appropriate for multicellular animals. For example, after the first two mitotic cell divisions in *C. elegans*, 1 out of the 4 cells goes on to form the germline (during these first two cell divisions, cytoplasm is unequally apportioned towards the 3 soma cells; thus, the germline precursor cell is derived from $< 1/4$ of the zygote cytoplasm) [1].

We divide the zygote cytoplasm (which contains $e_z$ endosymbionts) into $K$ pieces. We sample $p_s K$ of these pieces without replacement to form the germline precursor of the $j$th host, while the remaining $(1 - p_s) K$ pieces become the soma precursor. To sample a germline precursor for each host, we use the rhyper function with arguments $\text{nn}=N, m=Z_S^{t,r_2}, n=K - Z_S^{t,r_2}, k=p_s K$. We store the soma and germline precursors in a 2-by-$N$ matrix $C_S^{t,r_3,t}$. The soma precursor of the $j$th individual is indicated by
To generate the germline, we use random binomial sampling to sample $K$ pieces of cytoplasm with replacement from the $p_sK$ pieces in the germline precursor (using the \texttt{rbinom} function). The probability that a piece of the germline precursor in the $j$th host contains an endosymbiont is $e_g/(p_sK)$. To generate the soma, we use random binomial sampling to sample $K$ pieces of cytoplasm with replacement from the $(1 - p_s)K$ pieces in the soma precursor. The probability that a piece of the soma precursor in the $j$th host contains an endosymbiont is $e_s/((1 - p_s)K)$. The binomially-sampled endosymbionts from $C_{t,\tau}^{t_3}S$ form the 2-by-$N$ matrix $C_{t,\tau}^{t_3}$.

### D.6.1.3 Endosymbiont growth

As in section D.5.1.2, we calculate a weighted growth rate. Since we now track two tissues, we must calculate two weighted growth rates ($r_{t,s}I$, and $r_{t,g}I$) according to

$$r_{t,s}I(j) = r \left(1 - \frac{e_s}{K}\right),$$

where $e_s = C_{t,\tau}^{t_3}(1, j)$, and

$$r_{t,g}I(j) = r \left(1 - \frac{e_g}{K}\right),$$

where $e_g = C_{t,\tau}^{t_3}(2, j)$. We generate a random Poisson-distributed number of endosymbionts for each $C_{t,\tau}^{t_3}$ host and set all hosts in which $C_{t,\tau}^{t_3}(i) > K$ equal to $K$.

### D.6.1.4 Selection

Selection proceeds as in section D.5.1.3, only now we use the endosymbiont load in the soma ($C_{t,\tau}^{t_3}(1, j) = e_s$) instead of the endosymbiont load in the single-cell protist ($C_{t,\tau}^{t_3}(i) = e_c$ in section D.5.1.3).

### D.6.1.5 Meiosis

Unlike the protist model, in which there is only one diploid cell genotype, the sex ratio need not be equal in the arthropod model (we will use $N_f$ to refer to the number of females and $N_m$ to refer to the number of males). To keep the population size constant at $N$ individuals, we choose $N$ eggs from the $N_f$ females and $N$ sperm from the $N_m$ males.
To choose which eggs will form the next generation, we sample females without replacement until we have $N$ females. Since $N_f < N$, we first randomly order the females without replacement in a 1-by-$N_f$ vector (using the `sample` function). We then concatenate that vector with itself a sufficient number of times until we obtain a vector that exceeds length $N$. Finally, we truncate that vector to length $N$. (For example, if $N = 8$, $N_f = 5$, and our random order of females is $[2, 1, 4, 3, 5, 2, 1, 4]$, then the final 1-by-$N$ vector $(M^f)$ would be $[2, 1, 4, 3, 5, 2, 1, 4]$; thus, females 1, 2 and 4 have 2 fertilized eggs while females 3 and 5 have 1 fertilized egg.) Variance in reproductive success in meiosis is thus due to chance alone (we already account for fitness in selection).

Once we have identified how many times each female reproduces, we must determine the endosymbiont load of each egg. To do this, we use random binomial sampling (function `rbinom`), where the number of “draws” is $K$ and the probability of sampling a endosymbiont in the $j$th female chosen to reproduce is given by $C_{t, \tau}^{l, \tau_1}(1, M^l_F(j))/K$. We store eggs in the 1-by-$N$ vector $G^{l, \tau_5}_F$, where $G^{l, \tau_5}_F(j) = e_e$ indicates that the $j$th egg contains $e_e$ endosymbionts.

For all models except cytoplasmic incompatibility, we do not need to track sperm, as they do not contribute endosymbionts to the zygote (see section D.6.3 for the model of cytoplasmic incompatibility).

### D.6.1.6 Mating

In the mating phase, the eggs in the vector $G^{l, \tau_5}_F$ become the diploid zygote $Z^{l+1, \tau_1}_S$. Only eggs contribute endosymbionts to the zygote; thus, the number of endosymbionts in the zygote is simply the number of endosymbionts in the egg. For each zygote, we randomly assign $S$ to male or female (we use the `runif` function to generate a number between 0 and 1; if the number is less than 0.5, the zygote develops as a male).

### D.6.2 Feminization

Endosymbionts such as *Wolbachia* have developed strategies to aid their spread by skewing the sex ratio of infected host towards females. Endosymbionts use three strategies to bias female production over males: feminization, parthenogenesis, and male-killing. For simplicity, we only model feminization. Under feminization, endosymbionts can cause genetic males to develop as phenotypic females. We assume that the feminization trait is proportional to endosymbiont load, and that it is under negative frequency-dependent selection (i.e. as the frequency of feminization increases
in the population, its penetrance reduces). In reality, this process might involve selection for traits that cause uninfected females to produce male-skewed offspring or for traits that suppress feminization [2–4]. The important dynamic to capture is negative frequency-dependent selection against the feminization trait (for the purposes of this study, the actual processes driving this phenomenon are less important). Thus, the penetrance of feminization depends on a host’s endosymbiont load and the frequency of endosymbionts in the population. The probability that the $j$th genetic male develops as a phenotypic male is given by

$$1 - \left[ \frac{\sum_{x=1}^{N} Z_{M}^{t+1,\tau_1}(x)}{NK} \right]$$

where the term $1 - \sum_{x=1}^{N} Z_{M}^{t+1,\tau_1}(x)/(NK) \approx 1$ when the endosymbiont load of the population is low (i.e. little effect of negative frequency-dependent selection) but $\approx 0$ when the endosymbiont load of the population nears its carrying capacity, $NK$ (i.e. strong effect of negative frequency-dependent selection).

D.6.3 Cytoplasmic incompatibility

Cytoplasmic incompatibility means that infected males prevent uninfected females from producing viable offspring. Males never (or at least very rarely) transmit endosymbionts. Cytoplasmic incompatibility instead occurs due to sperm modifications in the male germline, with the penetrance of cytoplasmic incompatibility being proportional to endosymbiont density (we assume that penetrance of cytoplasmic compatibility is equivalent to the probability that an zygote survives) [5, 6]. We assume that cytoplasmic incompatibility depends on the endosymbiont load in the germline in which the gametes develop. Thus, in this version of the model, we must track the number of endosymbionts in the egg and the number of endosymbionts in the germline in which the egg developed. Thus, $G^{t,\tau_5}_{F,CI}$ becomes a 2-by-$N$ matrix, where $G^{t,\tau_5}_{F,CI}(1, j)$ stores the number of endosymbionts in the $j$th egg, and $G^{t,\tau_5}_{F,CI}(2, j)$ stores the number of endosymbionts in the germline in which the $j$th egg developed. We also need to track the number of endosymbionts in the germline from which the sperm developed. We store sperm in the 1-by-$N$ matrix $G^{t,\tau_5}_{M}$, where $G^{t,\tau_5}_{M}(j) = \epsilon$ indicates that the $j$th sperm developed in a germline that contained $\epsilon$ endosymbionts.

We assume that cytoplasmic incompatibility has zero penetrance ($p_{ci}$) when the number of endosymbionts in the female germline that produced an egg is $\geq$ to the number of endosymbionts in the male germline that produced the sperm with which it mates. When this is not the case, we assume that the penetrance of cytoplasmic
incompatibility depends linearly on the difference in the number of endosymbionts in the male and female germline. The penetrance of the $j$th zygote is given by

$$p_{ci}(j) = \frac{G_{M}^{t,tn}(O_{M}(j)) - G_{F,CI}^{t,tn}(2, O_{F}(j))}{K},$$

where $O_{M}(j)$ and $O_{F}(j)$ are 1-by-$N$ vectors with a random ordering of integers from 1 to $N$ without repeats (so that eggs and sperm are randomly paired). $p_{ci}$ is a 1-by-$N$ vector that stores the penetrance of cytoplasmic incompatibility for each egg/sperm pairing. When $G_{M}^{t,tn}(O_{M}(j)) \leq G_{F,CI}^{t,tn}(2, O_{F}(j))$, $p_{ci}(j) < 0$; for all values of $j$ in which $p_{ci}(j) < 0$, we set $p_{ci}(j)$ equal to 0. Since penetrance determines the probability that a zygote survives, we normalize $p_{ci}$ so that it sums to 1 according to

$$p_{ci}'(j) = \frac{p_{ci}(j)}{\sum_{x=1}^{N} p_{ci}(x)}.$$ 

We then use random multinomial sampling to sample $N$ zygotes according to the probabilities in $p_{ci}'$.

### D.6.4 Vertical transmission

As in D.5.2, in this version of the model we initialize the population with $N - 1$ zygotes with no endosymbionts and a single zygote with $b$ endosymbionts. No other horizontal transmission takes place. A simulation stops either when all endosymbionts are lost or the endosymbiont spreads and to the point that each host carries $\geq K/2$ endosymbionts. For each set of conditions, we ran 10,000 simulations and record the proportion in which the endosymbiont becomes fixed. The life cycle is the same as in section D.6.1 except that there is now no horizontal transmission phase.
Bibliography


